BCHM 355/455: BIOCHEMISTRY

Neil Voss
Roosevelt University

LibreTexts™

BCHM 355/455 Biochemistry (Roosevelt University)

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 03/04/2026

TABLE OF CONTENTS

Licensing

Preface from the Authors

Licensing

Biochemistry Foundation Concepts and Learning Goals and Objectives

Embedded Interactive Mathematical Graphs and Models (Vcell and Minisidewider)

1: Molecules of Life

- 1.1: Cellular Foundations of Biochemistry
- 1.2: Chemical Foundations
- 1.3: Physical-Chemical Foundations
- 1.4: Genetic Foundations
- 1.5: Essential elements for life
- 1.6: The Building Blocks of Molecules
- 1.7: Types of Biological Molecules
- 1.8: Chapter 1 Questions
- 1.9: Abiotic Origins of Life

2: Chemistry of Water

- 2.1: Water
- 2.2: The multiple roles of water
- 2.3: pH and Buffers
- 2.4: Weak Acids and Bases, pH and pKa
- 2.5: Buffering against pH Changes in Biological Systems
- 2.6: Solubility in an aqueous world - noncovalent interactions in depth
- 2.7: Solubility in an aqueous world - The Hydrophobic Effect
- 2.8: Chapter 2 Questions

3: Amino Acids and Peptides

- 3.1: Amino Acids and Peptides
- 3.2: Amino Acid Charges
- 3.3: Cysteine Chemistry
- 3.4: 3.1-Structure and Property of the Naturally-Occurring Amino Acids
- 3.5: From Amino Acids to Peptides

4: Proteins- Structure and Folding

- 4.1: The Structure of Proteins- An Overview
- 4.2: Levels of Protein Structure
- 4.3: Secondary Structure
- 4.4: Protein folding
- 4.5: Analyses of Protein Structure
- 4.6: Protein function, domains and cooperativity
- 4.7: The Three-Dimensional Structure of Proteins

- 4.7: The Three-Dimensional Structure of Proteins
  - 4.7.1: Main Chain Conformations
  - 4.7.2: Secondary Structure and Loops
  - 4.7.3: Tertiary, Quaternary, and Symmetrical Structures
  - 4.7.4: Secondary Structural Motifs and Domains
  - 4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure
  - 4.7.6: Intrinsically Disordered Proteins
  - 4.7.7: Fibrillar Proteins
  - 4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics
  - 4.7.9: Protein Stability - Thermodynamics
  - 4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules
  - 4.7.11: Biomolecular Visualization - Conceptions and Misconceptions
  - 4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation
  - 4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools
  - 4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24)
  - 4.7.15: Chapter 4 Questions
- 4.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs

5: Protein Purification

- 5.1: Protein Purification
- 5.2: Cell Disruption
- 5.3: Fractionation
- 5.4: Electrophoresis
- 5.5: Electrophoresis
- 5.6: Gel Exclusion Chromatography
- 5.7: Ion Exchange Chromatography
- 5.8: Affinity Chromatography
- 5.9: High Performance Liquid Chromatography (HPLC)
- 5.10: Histidine Tagging
- 5.11: Quantification of Protein Concentration
- 5.12: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)

6: Enzyme Thermodynamics

- 6.1: Enzymes
- 6.2: Enzyme Commission Number
- 6.3: Endergonic and Exergonic Reactions
- 6.4: The Laws of Thermodynamics
- 6.5: Thermodynamics
- 6.6: Energy in Metabolism

7: Enzyme Kinetics

- 7.1: Enzyme Kinetics
- 7.2: Enzyme Activity
  - 7.2.1: How Enzymes Work
  - 7.2.2: Kinetics without Enzymes
  - 7.2.3: Kinetics with Enzymes
  - 7.2.4: Enzyme Inhibition
  - 7.2.5: Enzymatic Reaction Mechanisms
  - 7.2.6: Enzymes and Protein Regulation

- 7.2.7: Ribozymes - RNA Enzymes
- 7.2.8: Cofactors and Catalysis - A Little Help From My Friends
- 7.3: Basic Principles of Catalysis
- 7.4: Derivation of Michaelis-Menten equation
- 7.5: Binding - The First Step Towards Protein Function

8: Enzyme Regulation

- 8.1: Regulation of Enzyme Activity
- 8.2: Mechanisms of Catalysis
- 8.3: Control of Enzymatic Activity
- 8.4: Oxygen-Binding Proteins and Allosterism

9: Carbohydrates

- 9.1: Carbohydrates overview
- 9.2: Monosaccharides
- 9.3: Disaccharides
- 9.4: Polysaccharides
- 9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls
- 9.6: The Sugar Code and Lectin Decoding
- 9.7: Chapter 7 Problems
- 9.8: Chapter 7 Problems - Answer Key

10: Nucleic Acids

- 10.1: Structure and Function - Nucleic Acids
- 10.2: Nucleic Acids - Structure and Function
- 10.3: Nucleic Acids - RNA Structure and Function
- 10.4: Nucleic Acids - Comparison of DNA and RNA
- 10.5: Enzymes for Genetic modifications
- 10.6: DNA Isolation, Sequencing, and Synthesis
- 10.7: Bioinformatics

11: Lipids

- 11.1: Introduction to lipids
- 11.2: Lipids Aggregates in Water - Micelles and Liposomes
- 11.3: Structure and Function - Lipids and Membranes

12: Membranes and Membrane Proteins

- 12.1: Basic Concepts in Membranes
- 12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics
- 12.3: Working with Lipids
- 12.4: Transport in Membranes
- 12.5: Problems
- 12.6: Membranes and Membrane Proteins
- 12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion
- 12.8: Diffusion Across a Membrane - Channels
- 12.9: Diffusion Across a Membrane - Pores
- 12.10: Active Transport

13: Biochemistry of Our Senses

- 13.1: Cell Signaling
- 13.2: Ligand-gated Ion Channel Receptors
- 13.3: Nuclear Hormone Receptors
- 13.4: G-protein Coupled Receptors (GPCRs)
- 13.5: Receptor Tyrosine Kinases (RTKs)
- 13.6: Gated Ion Channels - Neural Signaling
- 13.7: Signal Transduction - Vision and Olfaction
- 13.8: Signal Transduction - Taste (Gustation)
- 13.9: Signal Transduction - Temperature
- 13.10: Signal Transduction - Pressure

Index

Glossary

Detailed Licensing

Detailed Licensing

Licensing

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

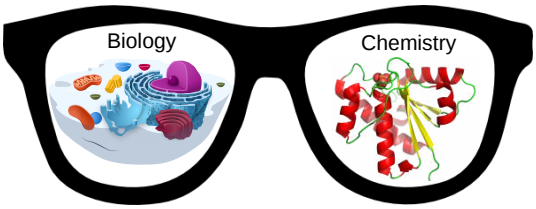
Philosophy

Writing a textbook is an enormous task. This is especially true for a biochemistry text, given the field's inherently interdisciplinary nature. Our insight into the biochemical bases for the biological world has exploded over time. Somehow that has to be presented in and integrated into a textbook, which attempts to explain life from a structure/function perspective.

Any book by default reflects the strengths, weaknesses, interests and biases of the writers. We have tried to select and explain topics that are perhaps most difficult for students to understand, but even those choices reflect in part the difficulties that authors have had to overcome throughout the years in their own understanding of the field. At the same time, we see our students struggling with the same content and concepts that we did as students. There is broad consensus among biochemists about which concepts are most difficult. Loertscher et al. have described these as threshold concepts and most are fundamental concepts in chemistry. Examples include Gibbs free energy, the difference between rapid equilibrium and the steady state, the thermodynamics of macromolecular structure formation, the physical basis of interactions, and in addition biomolecular visualization literacy.

At the same time, the biological complexities of the emergent properties of the cell, organs and organisms from a myriad of interacting molecules in space and time is truly humbling. Such complexity must increasingly be probed by mathematical models and the use of machine learning and artificial intelligence. Hence we have not shied away from the inclusion of mathematics and modeling of complex interconnected pathways. This is the future of biochemistry. Some material might be presented in too little depth for some. Again this reflects our backgrounds and experiences. Likewise, some material might be presented in too much depth for faculty adopters and students. We don’t view this as a problem since users can pick and choose the content they wish. We chose to write a text that extends students and is challenging.

No one can know all of biology and all of chemistry (an obvious statement). Each of us is predominately stronger in one field than the other. Think of yourself looking at the field of biochemistry with glasses containing a "biology-centric" lense and a "chemistry-centric" lens. We all have a dominant eye, which sees the world more clearly. Which lens is most comfortable to you as you study biochemistry?



If you have ever tried to see the world through glasses containing just one lens, the world looks distorted and incomplete. The same with biochemistry. It will become clearer to you if you correct your vision to include both a biological and chemistry lens. Adding even more lenses, including mathematics, computer sciences and physics, would offer a much better and clearer understanding (think of the compound lens of the fly)! In addition, most biochemistry texts might be considered protein centric and give less attention to other classes of molecules such as lipids and nucleic acids. Coverage of the latter topics might seem underrepresented in this book in the eyes of a lipid or nucleic acid chemist.

Fundamentals of Biochemistry is a classical textbook given its coverage and depth. It follows the arrangements and sequence of typical commercial textbooks to allow easier adoption by faculty and for ease of use by students. We have tried to not focus too much on human biochemistry, which is most likely emphasized in many biochemistry courses. Yet in reality, students who take biochemistry are often in a pre-medical track. We hope to populate chapters with problems based on the research literature to enable a more problem-solving approach.

We have deliberately chosen to incorporate graphs and some text from published materials that can be shared through Creative Common licenses. It would not have been possible to create a text without using such materials, as we did not have a publishing company supporting our work. We have done our best to attribute materials not directly made by us to the original authors. We apologize if we neglected to give attribution to those who sources we have used. The complexities of writing a book over many years in what free time we could devote to the project increases the likelihood of such omissions. In reality, we view ourselves as authors (of original materials) and compilers (of materials licensed through Creative Commons) of this textbook.

This book, as with any first edition text, will contain many mistakes. We hope these most are in the form of typos and grammatical errors. Given our backgrounds and the limitations of our own knowledge, there will also likely be conceptual mistakes. We ask for

your forgiveness for these and ask you to help correct them.

Key Features

This text covers a full year sequence of biochemistry. The book could serve traditional and ACS-certified majors, traditional and ASBMB-certified majors, as well as students taking a one-semester course. Characteristic of all LibreText books, Fundamentals of Biochemistry will be open to everyone via the internet and completely free, as all should have access to knowledge. It is comprehensive and topical, as new materials are easily added. The text is divided into five"books" and given the number of image and interactive models, it is best used online. Each chapter and section can be printed as a PDF file if needed. The LibreText collections of books have served 223 million students and saved them $31 million, an undue burden for our students and society.

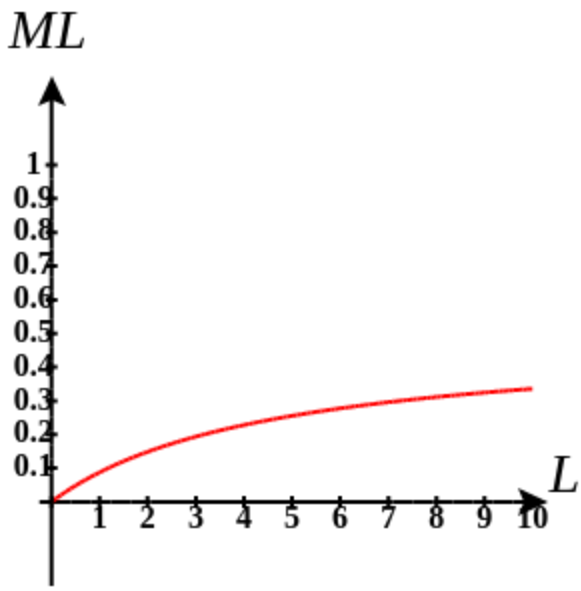
An online text has some advantages in addition to cost over hardback textbooks. Chief among these are embedded interactivity and ease of updating. Here are some of the differentiating features of Fundamentals of Biochemistry:

- Interactive 2D mathematical graphs with sliders that allow users to change mathematical constants with immediate update of graphs;
- Interactive molecular models using iCn3D (https://www.ncbi.nlm.nih.gov/Structure/icn3d/full.html) that allows students to view and interact with rendered structures;
- Embedded VCell (https://vcell.org/) models that allow users to display time course graphs for simple to complex reactions;
- Animations to accompany some of the VCell progress curve graphs;
- Chapter sections on metabolic control analyses;
- Constant addition of new materials such as RNA glycans, SARS-CoV-2 structures, etc.
- Literature-based guided assessments (described below)

Volume 4 contains two chapter not found in traditional biochemistry text books. This include chapter on the Abiotic Origins of Life and Biochemistry and Climate Change. It also contains an extensive chapter on Biosignaling - Capstone for Volume I

Interactive mathematical graphs

These offer clear advantages over static graphs (which we also choose to incorporate). Key constants in the mathematical equation can be changed using sliders with updated graphs shown in real time. For example, users can change rate constants, dissociation constants, and other parameter in way that change the apparent form of the graph. This will make the constants less abstract and give users insight into their fundamental meaning. An example is the graph for the concentration of the macromolecule:ligand complex (ML) vs free ligand (L) shown below. L_0 and M_0 are the total ligand and macromolecule concentrations, respectively.



If $L_0 \ll M_0$, then the equations simplifies to:

$$ML = \frac{(M_0)(L_0)}{K_D + L}$$

This is a graph for a hyperbola. If $K_D \gg L$ (equivalent to L << K_D) then the equation becomes

$$[\mathrm{ML}] = \left(\frac{\mathrm{M_0}}{\mathrm{K_D}}\right)\mathrm{L} \tag{1}$$

Under these conditions, ML is a linear function of L. Now change the slider for K_D in the interactive graph and see how the initial hyperbola changes into an apparently linear function when L << K_D.

Now if a reader who had fully explored this interactive graph conducted an actual binding analysis in the lab and got a linear plot of ML vs L, they might immediately recognize that they need to increase the ligand concentration to cover the full binding curve.

Use of iCn3D

iCn3D interactive molecular models are found in most chapters and sections. These models have been rendered to illustrate key points related to structure/function and are not presented just as attractive images. Users must interact with each model to enhance their understanding of the structure and function. Users can either click on the image and see an interactive popup model of the structure or use a link to go to a larger model in a separate window. Mouse instructions are given on each page of the popup window.

An example of an interactive iCn3D model presented in the book is the human nucleosome (3afa). Each iCn3D structure includes a Protein Data Bank (PDB) 4-digit alphanumeric code. We choose not to directly include iCn3D interactive models directly on the page. The size of the structure and the script necessary to produce the desired rendering make the load time for the page too long (especially when more than one model is present on a given page. Instead, users can click on the image and a popup of the fully rendered model will appear. The user can modify the rendered model by clicking the top menu bar icon (☰) and selecting from the

dropdown menu. This requires some knowledge of the iCn3D GUI. Alternatively, the user can click the link in the legend underneath the image to get an expanded new window of the model directly through the full iCn3D user webpage.

Explore the model of the nucleosome below. This structure shows how doubled-stranded DNA is packed around an octamer of pairs of 4 basic proteins called histones that are found in the eukaryotic nucleus. One member of each pair of histones is shown in cartoon rendering, while the other member of the pair is shown in the same color but in spacefill rendering. Each strand of DNA is shown in a different shade of gray.

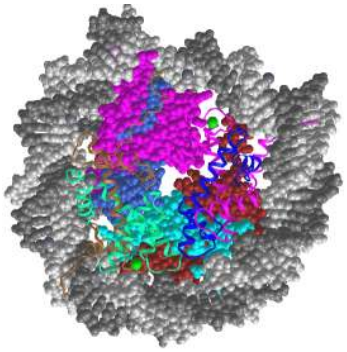


Figure **10**: Human nucleosome (3afa). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...B2SwQHYDLj4BJ6

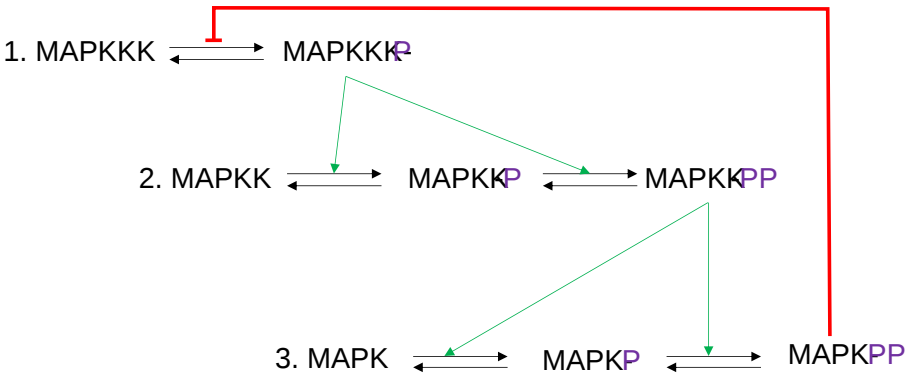
Simply viewing the image will not lead to enhanced insight into the structure. Users should hover over key groups and rotate the model to help identify interactions. Labeled and interactive computer models offer key advantages over static 2D images. Using additional menu items to further interrogate the structure over and above the rendered image offered will lead to increased insight.

A note on images: Many of the static structures of biomolecules are made with PyMOL. The rest are made with iCn3D. Efforts have been made to use appropriate colors for those who are color-blind.

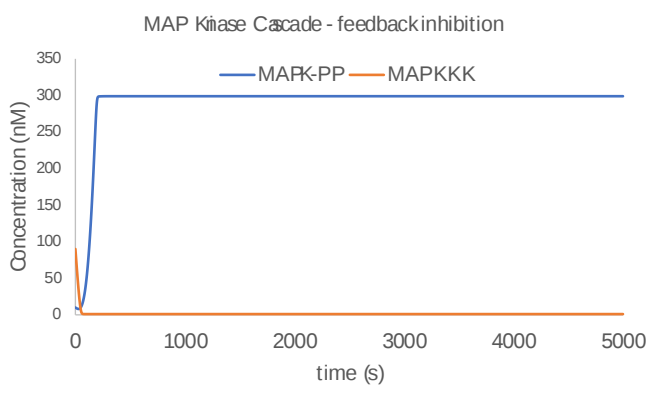
Embedded Virtual Cell (VCell) models

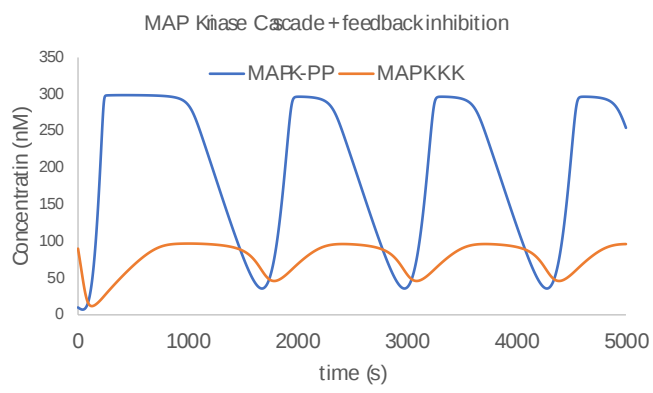
VCell (Virtual Cell) is a platform for the mathematical modeling of cell biological systems. Our use of it is to display product vs time progress curves for simple to more complicated reactions. It can do so much more than that, however. For example, it can be used to model the concentrations of all metabolites in full metabolic and signal transduction pathways. We introduce it in this book to allow users to alter the rate, kinetic, and dissociation constants, etc., and see the effects on the output of simple to more complex reactions. Given the complexity of the biological world and its interactions, mathematical modeling of biological processes is essential to fully understanding it. Most readers do not have the mathematical and computer skills to create and run models on their own. Vcell does that for you. We will explain in general the mathematical equations used in Vcell to model many reactions throughout this book. We hope that interested readers will become more familiar with Vcell and explore it for their own uses in courses and eventually in this own research.

Modeling is very important especially when our intuitions are insufficient to account for observed results. Take for example the reactions of three enzymes involved in cell signaling in the **mitogen activated protein kinase (MAPK)** cascade system as illustrated in the scheme below. (We will explore this in more detail later in the text).



A protein kinase is an enzyme that covalently transfers a terminal phosphate **(P)** from ATP to a protein. Line 1 shows one protein kinase, MAPKKK, that is phosphorylated (by another kinase not shown) to become MAPKKK-**P**. MAPKKK-**P**in turn is an active protein kinase that doubly phosphorylates MAPKK (line 2) to form MAPKK-**PP**. MAPKK-**PP** is in turn yet another downstream kinase, which doubly phosphorylates MAPK to form MAPK-**PP**. These phosphorylation reactions are shown in green in the reaction diagram. If you plot the concentration of MAPKKK (the starting enzyme with a cellular concentration set to 100 nM and MAPK-**PP** (initial concentration of MAPK is 300 nM) vs time for a specific set of conditions, you get the very predictable graph shown in the left panel below. However, if you allow the final product of the cascade (MAPK-**PP**), to inhibit the very first reaction in the cascade (Rx 1 forward, red blunt arrow), you get the remarkable oscillatory graph shown in the right panel of the figure.

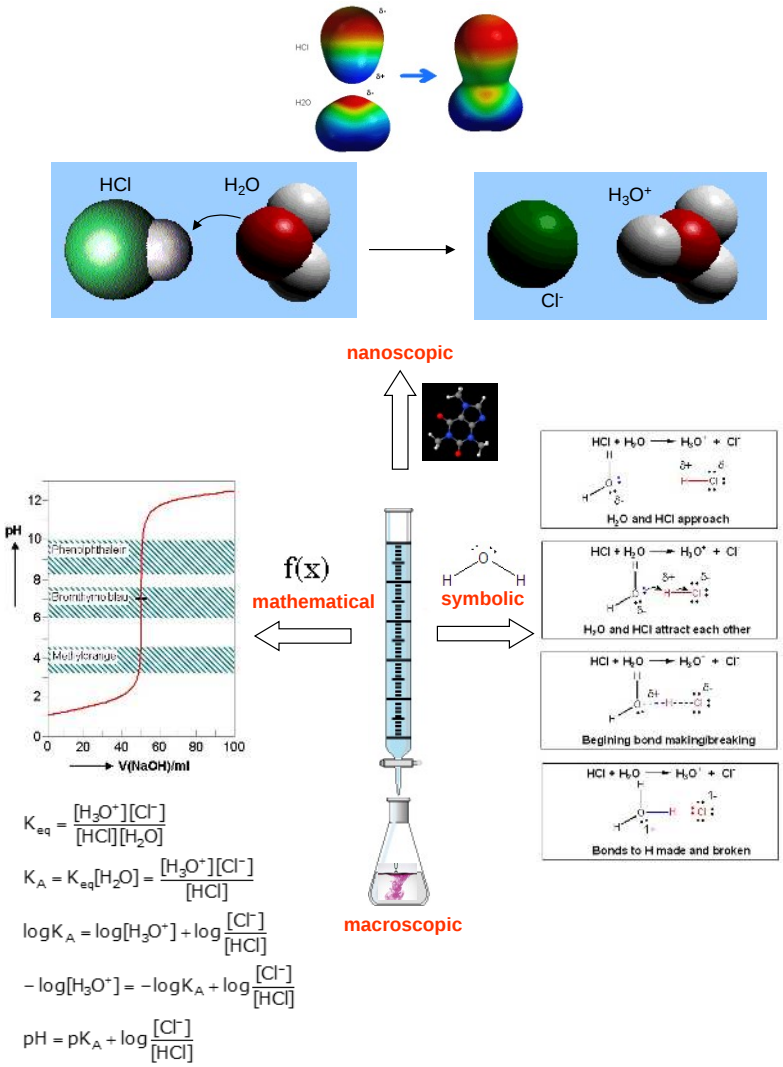




There is simply no way to predict the observed oscillatory behavior for the system from simple intuition as was possible for the reaction scheme without feedback inhibition. Vcell allows these types of calculations, which greatly informs our understanding of biological interactions in complicated biological pathways. (The Virtual Cell was developed with funding from the **National Institute of General Medical Sciences (NIGMS)** as a Biomedical Technology Research Resource at the Center for Cell Analysis and Modeling (CCAM).

Animations

Science is difficult for most students. Let's pick chemistry as an example and consider one of its simplest reactions, that of hydrogen chloride and water, to form the hydronium ion and chloride. We know such a mixture as hydrochloric acid. Chemists use every resource available to them to understand and explain reactions. These encompass the macroscopic, nanoscopic, mathematical, and symbolic realms, as shown in the figure below.



Some need words to best understand a difficult concept. Others need images. Most need multiple modalities. For that reason, we have incorporated animations of relevant physical and chemical reactions for which we also have embedded Vcell models. These animations were created by Hui Liu (Anne) and Shraddha Nayak from Janet Iwasa's lab. In addition, we have incorporated metabolic flux animations created by Nayak for key metabolic pathways.

We hope that all the different learning modalities we incorporate into Fundamentals of Biochemistry will help users comprehend this amazing yet difficult field.

Literature-based Guided Assessments (LGAs)

Assessment questions are typically given at the end of chapters in traditional biochemistry texts. As they are targeted to specific chapters, they are more likely to address ideas expressed in a chapter and less likely to be summative across chapters. In addition, although specific questions may use data and findings from the research literature, the questions are less likely to intentionally address learning goals tied to students' abilities to read and understand research papers, an arguably difficult task. Finally, the question when tied to specific chapters might not be relevant to larger issues of the day, such as the development of new antivirals targeting pandemic viruses.

To address these issues, we have developed a new approach for assessment questions housed in this book. We have created LGAs with data, graphs, and images derived from 1 or more research papers (with the correct Creative Commons permissions). The LGAs also serve as a way for students to read and explore the research literature in a guided way. We have written traditional

LGAs and those based on the world's greatest challenges (and problems). These include climate change, trauma, information integrity, and addiction. More will be added in the future.

Much of the content covered in biochemistry classes is in service to the premedical curriculum, which focuses almost exclusively on human health as opposed to the emerging notion of “one health". The LGAs linked to global challenges allow us to cover in an organized way the “underserved” areas (based on traditional biochemistry courses) in biochemistry.

What is still needed

We continually update the book. As it has five "volumes, we've added a search bar: **Search Fundamentals of Biochemistry**. Input "**Recent Updates"** (with the quotes) to see relatively new materials and substantive changes.

We still need help in writing, updating, and editing. especially in these areas:

- embedded and end-of-chapter problems (hopefully based on research literature)
- mechanisms of your favorite enzymes or topics you feel are not presented well in other texts
- inclusive and diverse content
- new LGAs

If you are interested in assisting in any of these areas, please contact Henry Jakubowski at hjakubowski@csbsju.edu

Acknowledgments

We greatly appreciate the help of Leslie Loew, Ion Moreau, Jonathan Karr, Bilal Shaikh, and Henry Agnew in integrating Vcell into Fundamentals of Biochemistry. In addition we thank Bartholmew Jardine and Herbert Sauro for the development and use of the Minisidewider software for interactive mathematical simulations in the book. Finally, we greatly appreciate the help of Hui Liu (Anne), Shraddha Nayak, and Janet Iwasa in creating the animations that accompany the VCell and other models.

References

1. Loertscher et al. Identification of Threshold Concepts for Biochemistry. **CBE—Life Sciences EducationVol. 13, No. 3** (2017). https://doi.org/10.1187/cbe.14-04-0066
2. VCell :
    - Schaff, J., C. C. Fink, B. Slepchenko, J. H. Carson, and L. M. Loew. 1997. A general computational framework for modeling cellular structure and function. Biophysical journal 73:1135-1146. PMC1181013
    - Cowan, A. E., Moraru, II, J. C. Schaff, B. M. Slepchenko, and L. M. Loew. 2012. Spatial modeling of cell signaling networks. Methods Cell Biol 110:195-221. PMC3288182

3. https://upload.wikimedia.org/wikiped...nannotated.svg

4. iCn3D: Wang J, Youkharibache P, Zhang D, Lanczycki CJ, Geer RC, Madej T, Phan L, Ward M, Lu S, Marchler GH, Wang Y, Bryant SH, Geer LY, Marchler-Bauer A. iCn3D, a Web-based 3D Viewer for Sharing 1D/2D/3D Representations of Biomolecular Structures. ***Bioinformatics.*** **2020** Jan 1;36(1):131-135. (Epub 2019 June 20.) doi: 10.1093/bioinformatics/btz502. [PubMed PMID: 31218344]

5. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

Licensing

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

Introduction

Fundamentals of Biochemistry is designed to help readers understand many topics across a broad range of area. It covers the the biochemistry foundational concepts recommended by the American Chemical Society (ACS), the American Society for Biochemistry and Molecular Biology (ASBMB) and for the Medical College Admissions Test (MCAT) written by the American Association of Medical Colleges (AAMC). The foundational concepts recommended by these organizations broadly overlap. The ASBMB is the only organization that has also defined specific learning objectives for each of the foundational concepts.

These strongly emphasize scientific inquiry and reasoning skills, which are perhaps best assessed by open-ended questions derived from the literature in which students must employ higher level Bloom skills of application and analysis.

Modified from http://expertbeacon.com/blooms-taxonomy/

The ability to apply, analyze, and evaluate information and concepts are at the heart of scientific inquiry and reasoning skills which are central to the competency standards.

Each chapter and chapter section will (by 2023) contain learning goals and objective and for each chapter, literature-based assessment question and answers that address the foundational concepts and learning goals.

Biochemistry Foundational Concepts and Learning Goals/Objectives

We will present the biochemistry foundational concepts and learning goals/objectives recommended by the following professional organizations.

American Chemical Society (ACS)

The American Chemical Society does not describe learning goals and objectives, but rather suggest concepts in three general areas that should be covered in a biochemistry course. These include:

Biological Structures and Interactions

- Fundamental building blocks (amino acids, carbohydrates, lipids , nucleotides, and prosthetic groups)
- Biopolymers (nucleic acids, peptides/proteins, glycoproteins, and polysaccharides)
- Membranes
- Supramolecular architecture

Biological Reactions

- Kinetics and mechanisms of biological catalysis
- Biosynthetic pathways and strategies/metabolic engineering
- Metabolic cycles, their regulation, and metabolomics
- Organic and inorganic cofactors

Biological Equilibria and Thermodynamics

- Acid-base equilibria
- Thermodynamics of binding and recognition

- Oxidation and reduction processes Electron transport and bioenergetics
- Protein conformation/allostery, folding, oligomerization, and intrinsically disordered proteins (IDPs)

American Society for Biochemistry and Molecular Biology

Of the three professional organization, the ASBMB is the only one that offers both core concepts and learning objectives. Applicable ASBMB learning objectives will be listed at the beginning of each chapter sections (in 2023).

Core Concept 1: Energy is required by and transformed in biological systems.

Learning Objectives:

1.1. Given knowledge of common mechanisms of regulation for biomacromolecules, students should be able to predict the sites and nature of regulation in pathways that transform energy

1.2. Given a knowledge of the basic structure of fatty acids, triglycerides, nucleotides, and carbohydrates; students should be able to compare and contrast the synthesis, storage, and transformation of macromolecules from which living organism harvest derive energy

1.3. Given a macromolecule, students should be able to explain the contribution of entropy, enthalpy and temperature of a macromolecule and water (associated and in bulk solvent) in a folded versus unfolded state

1.4. Given a biological example, students should be able to explain how thermodynamically unfavorable processes can occur.

1.5. Given an enzyme reaction, students should be able to differentiate the effect of an enzyme on the change in free energy of a reaction versus the change in transition state
free energy

1.6. Given entropy, enthalpy and temperature, students should be able to justify why evolutionary selection is constrained by the laws of thermodynamics

Core Concept 2: Macromolecular structure determines function and regulation

Learning Objectives:

2.1. Given the knowledge of bio macromolecules, students should be able to identify, draw(sketch) and know properties (functions) of biomacromolecules

2.2. Given a list of macromolecules, students should be able to devise an experiment on how they interact or interpret results of experiments on their interactions

2.3. Given structural changes of a macromolecule, students should be able to predict the impact of structural substitution would have on macromolecule structure and function

2.4. Given experimental data, students should be able to assess how enzymes facilitate biochemical reactions

2.5. Given that evolutionary forces such as gene duplications and genomic mutations can provide changes in protein structure(s) and function(s) students should be able to explain how protein structures might change while retaining an evolutionarily conserved function

Core Concept 3: Information storage and flow are dynamic and interactive

Learning Objectives:

3.1. Given an understanding of replication, transcription and translation, students should be able to determine how changes in DNA sequence affect the amino acid sequence of the protein it encodes

3.2. Given an understanding of common mechanisms of gene regulation, students will be able to explain or predict changes in transcription in response to biologic variables.

3.3. Given an understanding of genetic information transfer, students should be able to explain the role of RNA in the flow of genetic information

3.4. Given an understanding of evolution and natural selection, students should be able to make predictions on how environmental factors will affect information flow over generations.

3.5. Given an understanding of commonly encountered signal transduction mechanisms, students will be able to predict or design models for information transfer cascades

American Association for Medical Colleges (AAMC) - MCAT

The AAMC has defined the core concepts and content areas covered in their latest version (2015) of the MCAT. These are listed below.

Biological and Biochemical Foundations of Living Systems

Foundational Concept 1: Biomolecules have unique properties that determine how they contribute to the structure and function of cells and how they participate in the processes necessary to maintain life. The content categories for this foundational concept include:

1A. Structure and function of proteins and their constituent amino acids.

1B. Transmission of genetic information from the gene to the protein.

1C. Transmission of heritable information from generation to generation and the processes that increase genetic diversity.

1D. Principles of bioenergetics and fuel molecule metabolism.

Foundational Concept 2: Highly organized assemblies of molecules, cells, and organs interact to carry out the functions of living organis

2A. Assemblies of molecules, cells, and groups of cells within single cellular and multicellular organisms.

2B. The structure, growth, physiology, and genetics of prokaryotes and viruses.

2C. Processes of cell division, differentiation, and specialization.

Foundational Concept 3: Complex systems of tissues and organs sense the internal and external environments of multicellular organisms, and through integrated functioning, maintain a stable internal environment within an ever-changing external environment.

3A. Structure and functions of the nervous and endocrine systems and ways these systems coordinate the organ systems.

3B. Structure and integrative functions of the main organ systems.

Chemical and Physical Foundations of Biological Systems Section

Foundational Concept 4: Complex living organisms transport materials, sense their environment, process signals, and respond to changes using processes understood in terms of physical principles.

4A. Translational motion, forces, work, energy, and equilibrium in living systems.

4B. Importance of fluids for the circulation of blood, gas movement, and gas exchange.

4C. Electrochemistry and electrical circuits and their elements.

4D. How light and sound interact with matter.

4E. Atoms, nuclear decay, electronic structure, and atomic chemical behavior.

Foundational Concept 5: The principles that govern chemical interactions and reactions form the basis for a broader understanding of the molecular dynamics of living systems.

5A. Unique nature of water and its solutions.

5B. Nature of molecules and intermolecular interactions.

5C. Separation and purification methods.

5D. Structure, function, and reactivity of biologically relevant molecules.

5E. Principles of chemical thermodynamics and kinetics.

Embedded Interactive Mathematical Graphs and Models (Vcell and Minisidewider)

All of these chapter sections have embedded graphs (using Calc3D Graphs) and those based on Vcell models imported into Fundamentals of Biochemistry and run interactively in the book using Minisidewinder and other programs. In these graphs, students can move sliders to change concentrations and kinetic and equilibrium/dissociation constants and see immediate effects on the graph.

Fundamentals of Biochemistry

- Preface

Volume I: Structure and Catalysis

Chapter 5: Protein Function

- Chapter 5.1: Binding (non-kinetic graphs)
- Chapter 5.3: Oxygen binding proteins and allosterism (non-kinetic graphs)
- Chapter 5.7: ELISAs (non-kinetic graphs)

Chapter 6: Enzyme Activity

- 6.2: Kinetics without Enzymes
- 6.3: Kinetics with Enzymes
- 6.4: Enzyme Inhibition

Volume II: Bioenergetics and Metabolism

Chapter 14: Principles of Metabolic Regulation

- 14.2: Basic Principles of Metabolic Control Analysis (MCA)
- 14.5: Metabolism and Signaling: The Steady State, Adaptation and Homeostasis

Volume IV: Signaling, Integration of Metabolism, and Special Topics

Chapter 28: Biosignaling - Capstone Volume I

- 28.4: The next step - Downstream intracellular signaling
- 28.7: Calcium Signaling
- 28.13: Regulation of the Cell Cycle by Protein Kinases

CHAPTER OVERVIEW

1: Molecules of Life

1.1: Cellular Foundations of Biochemistry
1.2: Chemical Foundations
1.3: Physical-Chemical Foundations
1.4: Genetic Foundations
1.5: Essential elements for life
1.6: The Building Blocks of Molecules
1.7: Types of Biological Molecules
1.8: Chapter 1 Questions
1.9: Abiotic Origins of Life

1: Molecules of Life is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.1: Cellular Foundations of Biochemistry

Learning Goals

1. **Understand the Role of Water and Ions in Biochemistry**
   - Importance of water's unique properties in biochemical systems.
   - pH and ion gradients influence cellular function and metabolism.
2. **Characterize the Structure and Function of Biomolecules**
   - The four major classes of biomolecules: proteins, nucleic acids, lipids, and carbohydrates.
   - Macromolecular structure relates to function in metabolism and enzyme activity.
3. **Explore Metabolism and Enzyme Function**
   - Catabolic and anabolic reactions in metabolism.
   - Role of enzymes as biological catalysts and how they regulate metabolic pathways.
4. **Examine the Biochemical Environment of the Cell**
   - Crowded nature of the cytoplasm and its effects on protein stability and enzyme function.

By achieving these learning goals, students will develop a solid understanding of the cellular foundations that underlie all biochemical processes, preparing them for more advanced topics in biochemistry and related disciplines.

prompt: Write a series of learning goals for the following web page. The page is designed for junior and senior biochemistry majors.

1.1.1: Introduction

Biochemistry explores the chemical processes that sustain life. This text focuses on four key biomolecule classes: lipids, proteins, nucleic acids, and carbohydrates. It examines their roles in biochemical reactions, metabolism, and cellular function.

Cells function as chemical factories. They import, synthesize, use, and degrade biomolecules to sustain life. **Metabolism** includes both the energy-releasing breakdown of molecules (**catabolism**) and the energy-consuming synthesis of new compounds (**anabolism**). These processes support growth, reproduction, and adaptation to environmental changes.

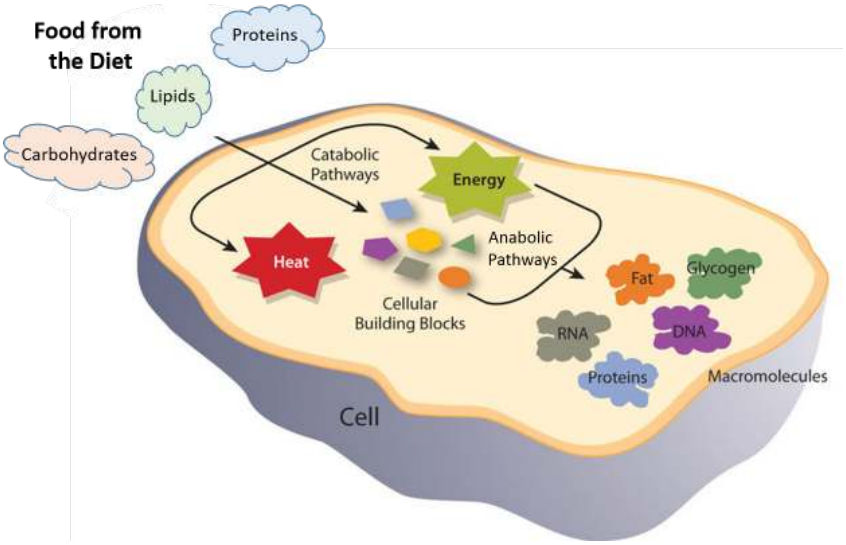


Figure 1.1 Catabolic and Anabolic Reactions. Catabolic reactions involve the breakdown of molecules into smaller components, whereas anabolic reactions build larger molecules from smaller molecules. Catabolic reactions usually release energy, whereas anabolic processes usually require energy. (CC BY-SA-NC 3.0; anonymous)

Enzymes are essential to metabolism. They act as catalysts that speed up chemical reactions. Enzymes also help regulate metabolic pathways by responding to conditions inside the cell and external signals. Their structure determines specificity, either through a **lock-and-key** model or an **induced-fit** model of substrate binding.

Metabolic reactions occur in **pathways**, where one molecule is gradually converted into another. Each step is often controlled by a specific enzyme. This organization allows efficient energy production, biosynthesis, and waste elimination.

Biological membranes play a critical role in biochemistry. The **lipid bilayer** and membrane proteins regulate molecular transport. **Passive and active transport** mechanisms control the movement of nutrients, ions, and waste.

1.1.2: pH and Ion Regulation in Cells

Cellular function depends on the regulation of pH and ion concentrations. The ***cytosol*** typically has a pH between 7.0 and 7.4, but some organelles maintain distinct pH environments. For example, lysosomes have a pH of about 4.8, which is necessary for enzymatic degradation. The mitochondria create a pH gradient across the inner membrane, which drives ATP synthesis.

Ion gradients across membranes are critical for transport, energy production, and cell signaling. The concentration of potassium ions is higher inside the cell, while sodium, chloride, and calcium ions are more concentrated outside. These gradients are maintained by ion transporters and require ATP hydrolysis.

Table 1.1 Average Cellular and Extracellular Ion Concentrations

Ion	Inside (mM)	Outside (mM)
Na^+	140	5
K^+	12	140
Cl-	4	15
Ca^{2+}	1 uM	2

1.1.3: The cell is an amazingly crowded place

In chemistry labs, we typically work with dilute solutions of solutes in a solvent. However, the cellular environment is very different. While the body is about 68% water, the concentration of water varies within cells due to the presence of densely packed biomolecules. Proteins, carbohydrates, and other solutes fill much of the available space, often leaving less room between molecules than the size of a single protein.

This crowded environment has several important effects. Studies show that high intracellular concentrations help stabilize proteins, keeping them in their correctly folded, native state. Crowding also slows the diffusion of molecules, which affects reaction rates and the localization of cellular functions. As a result, biochemical processes are often restricted to specific regions within the cytoplasm, creating specialized microenvironments that enhance efficiency.

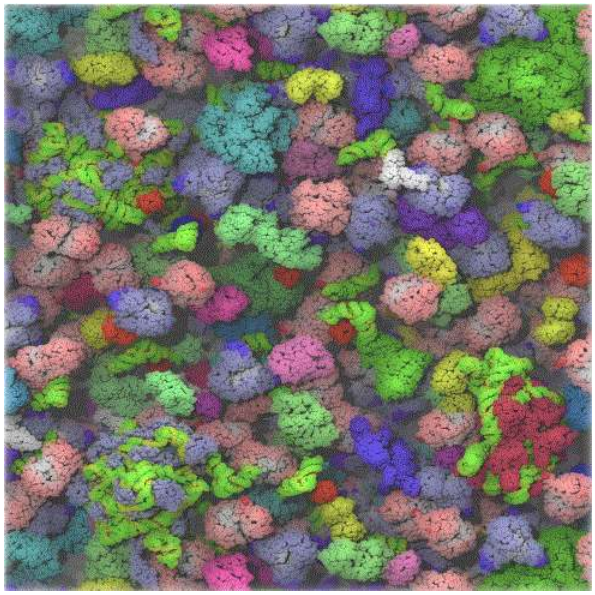


Figure 1.2: The Crowded Cytoplasm of E. Coli. This computer simulation modeled the 50 most abundant macromolecules in the E. coli cytoplasm, with a total of 1008 individual molecules. RNA is shown in green and yellow. Image adapted from McGuffee SR, Elcock AH (2010). PLoS Comput Biol 6(3): e1000694. https://doi.org/10.1371/journal.pcbi.1000694 (open source journal).

Because lab experiments are usually performed in dilute solutions, they may not fully reflect how biomolecules behave inside cells. For example, researchers have shown that adding a neutral copolymer of sucrose and epichlorohydrin to *in vitro* experiments can mimic the crowded cellular environment. This artificially induced crowding caused extracellular molecules to self-organize, leading to structural changes inside the cell. Furthermore, enzyme activity in key metabolic pathways, such as glycolysis, increases significantly under crowded conditions. Crowding may also promote the spatial and temporal association of enzymes, allowing substrates and products to pass efficiently between them in metabolic pathways.

This page titled 1.1: Cellular Foundations of Biochemistry is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **1.1: Cellular Foundations** by Henry Jakubowski and Patricia Flatt is licensed CC BY 4.0.

1.2: Chemical Foundations

Learning Goals (ChatGPT o3-mini)

1. **Fundamental Concepts in Organic Chemistry**
   - Define what constitutes an organic molecule and distinguish it from inorganic compounds.
   - Explain the significance of carbon's versatility in forming millions of unique compounds.
2. **Understanding Functional Groups**
   - Identify and describe common functional groups (e.g., hydroxyl, carbonyl, carboxyl, amine, phosphate, etc.) and their characteristic bonding patterns.
   - Predict the reactivity of a molecule based on the presence and arrangement of its functional groups.
3. **Classification of Hydrocarbons**
   - Differentiate between alkanes, alkenes, alkynes, and aromatics in terms of structure and reactivity.
   - Explain the concepts of saturation versus unsaturation and the implications for chemical behavior.
4. **Reactivity and Structural Diversity**
   - Illustrate how similar functional groups in different molecules lead to predictable reactivity patterns.
   - Analyze examples (such as capsaicin) to identify multiple functional groups within a single compound and discuss their collective impact on molecular behavior.
5. **Specialized Functional Group Classes**
   - Distinguish between closely related groups such as alcohols, phenols, and thiols, as well as ethers versus sulfides.
   - Understand the classification (primary, secondary, tertiary) of both alcohols and amines, noting how definitions differ between these groups.
6. **Organic Phosphates and Carbonyl Compounds**
   - Explain the structural and functional roles of organic phosphates in biomolecules.
   - Compare and contrast aldehydes, ketones, and carboxylic acids along with their derivatives, focusing on their reactivity in biological contexts.
7. **Linking Organic Chemistry to Biochemistry**
   - Demonstrate how functional groups participate in dehydration synthesis reactions that form major macromolecules (proteins, nucleic acids, carbohydrates, and lipids).
   - Recognize the importance of functional group activation in the synthesis of complex biomolecules.
8. **Primary Metabolites and Cellular Function**
   - Define primary metabolites and explain their role in fundamental cellular processes such as energy production, growth, and signaling.
   - Relate the chemical principles of organic reactions to the formation and function of cellular structures (e.g., membranes, cell walls, cytoskeleton).

By achieving these goals, students will deepen their understanding of the structure, reactivity, and biological significance of organic molecules and functional groups, laying a strong foundation for exploring more complex biochemical processes.

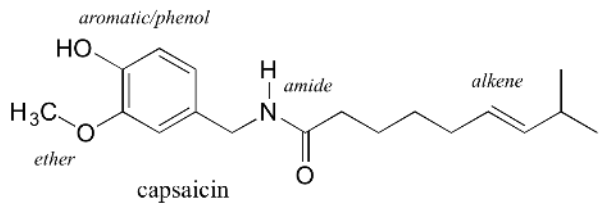
1.2.1: Organic Molecules

On Earth, all carbon-containing molecules have originated from biological, living organisms and are called **organic compounds**. Organic chemicals consist of relatively few similar parts combined in different ways. These structural similarities enable us to predict how a previously unseen compound may react, provided we understand how other molecules containing the same types of parts are known to react.

These parts of organic molecules are called **functional groups** with specific bonding patterns and atoms most commonly found in organic molecules (C, H, O, N, S, and P). Identifying functional groups and predicting reactivity based on their properties is one of the cornerstones of organic chemistry. Functional groups are specific atoms, ions, or groups of atoms having consistent properties. A functional group makes up part of a larger molecule. For example, -OH, the hydroxyl group that characterizes alcohols, contains oxygen with an attached hydrogen. It could be found on any number of different molecules. Just as elements have distinctive properties, functional groups have characteristic chemistries. An -OH functional group on one molecule will tend to react similarly, although perhaps not identically, to an -OH on another molecule.

Organic reactions usually occur in the functional group, so learning about the reactivities of functional groups will prepare you to understand many other aspects of biochemistry.

Functional groups are structural units within organic compounds defined by specific bonding arrangements between specific atoms. The structure of capsaicin, the fiery compound found in hot peppers, has several functional groups, labeled in the figure below and explained throughout this section.

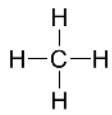


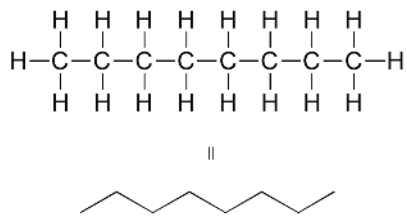
capsaicin

As we progress in biochemistry, it will become increasingly important to recognize the most common functional groups. These *are the key structural elements that define how organic molecules react*. Below is a brief introduction to the major organic functional groups.

1.2.1.1: Alkanes

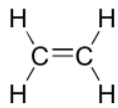
The 'default' in organic chemistry (the *lack* of functional groups) is given the term *alkane*, characterized by single bonds between carbon and carbon or carbon and hydrogen. Methane, CH_4, is the natural gas you may burn in your furnace. Octane (C_8H_{18}) is a component of gasoline.

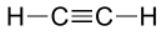

methane


octane

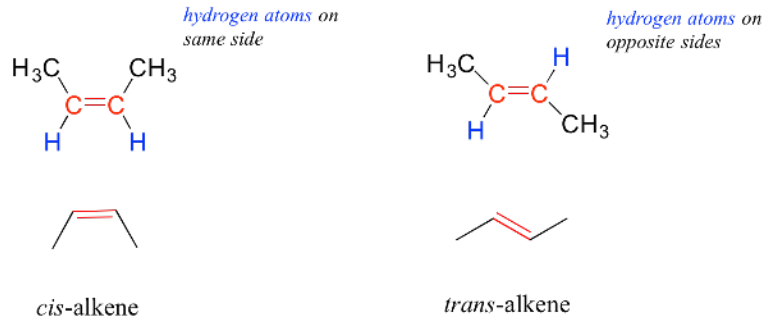
1.2.1.2: Alkenes and Alkynes

Alkenes (sometimes called olefins) have carbon-carbon double bonds, and *alkynes* have carbon-carbon triple bonds. Ethene, the simplest alkene example, is a gas that serves as a cellular signal in fruits to stimulate ripening. (If you want bananas to ripen quickly, put them in a paper bag along with an apple – the apple emits ethane gas (also called ethylene), setting off the bananas' ripening process). Ethyne, commonly known as acetylene, is used as a fuel in welding blowtorches.


ethene
(an alkene)


ethyne
(an alkyne)

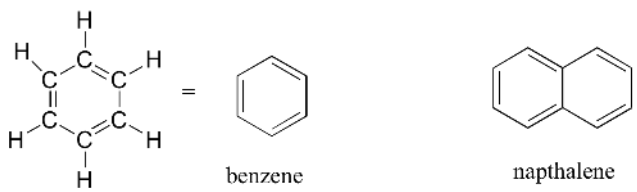
Many alkenes can take two geometric forms: *cis* or *trans*. The *cis* and *trans* forms of a given alkene are different isomers with different physical properties because there is a very high energy barrier to rotation about a double bond. In the example below, the difference between *cis* and *trans* alkenes is readily apparent. Biochemists don't usually use the E (entgegen) and Z (zusammen) labels for groups attached to double bonds (using IUPAC priority numbering).



Alkanes, alkenes, and alkynes are all classified as *hydrocarbons* because they are composed solely of carbon and hydrogen atoms. Alkanes are said to be *saturated hydrocarbons* because the carbons are bonded to the maximum possible number of hydrogens – in other words, they are '*saturated*' with hydrogen atoms. The double and triple-bonded carbons in alkenes and alkynes have fewer hydrogen atoms bonded to them – they are thus referred to as *unsaturated hydrocarbons*.

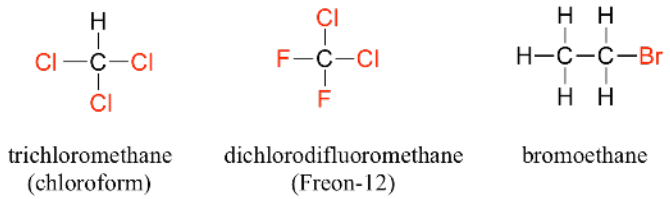
1.2.1.3: Aromatics

The *aromatic group* is exemplified by benzene and naphthalene, a compound with a distinctive 'mothball' smell. Aromatic groups are planar (flat) ring structures with conjugated pi bonding with 4n+2 pi electrons. Given the stability of aromatic groups due to the delocalization of pi electrons, these groups are abundant in nature.



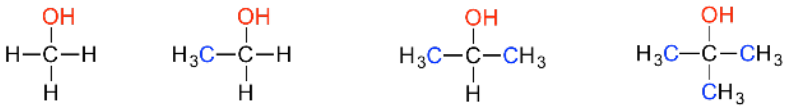
1.2.1.4: Alkyl Halides

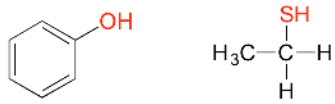
When the carbon atom of an alkane is bonded to one or more halogens, the compound is referred to as an *alkyl halide, or haloalkane*. Chloroform is a valuable solvent in the laboratory and was one of the earliest anesthetic drugs used in surgery. Chlorodifluoromethane was used as a refrigerant and in aerosol sprays until the late twentieth century. Still, its use was discontinued after it was found to have harmful effects on the ozone layer. Bromoethane is a simple alkyl halide often used in organic synthesis. Alkyl halide groups are quite rare in biomolecules.



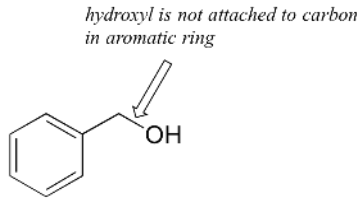
1.2.1.5: Alcohols, Phenols, and Thiols

In the ***alcohol*** functional group, a carbon is single-bonded to an OH group (when part of a larger molecule, the OH group is referred to as a ***hydroxyl group***). All alcohols can be classified as primary, secondary, or tertiary. In primary alcohols, the carbon bonded to the OH group is also bonded to only one other carbon. In secondary and tertiary alcohols, carbon is bonded to two or three other carbons. When the hydroxyl group is *directly* attached to an aromatic ring, the resulting group is called a ***phenol.*** The sulfur analog of an alcohol is called a thiol (from the Greek *thio*, for sulfur).





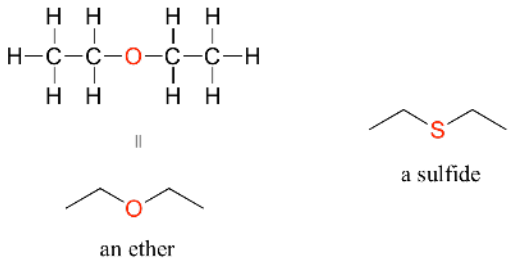
Note that the definition of a phenol states that the hydroxyl oxygen must be *directly* attached to one of the carbons of the aromatic ring. The compound below, therefore, is *not* a phenol – it is a primary alcohol.



The distinction is important because there is a significant difference in the reactivity of alcohols and phenols.

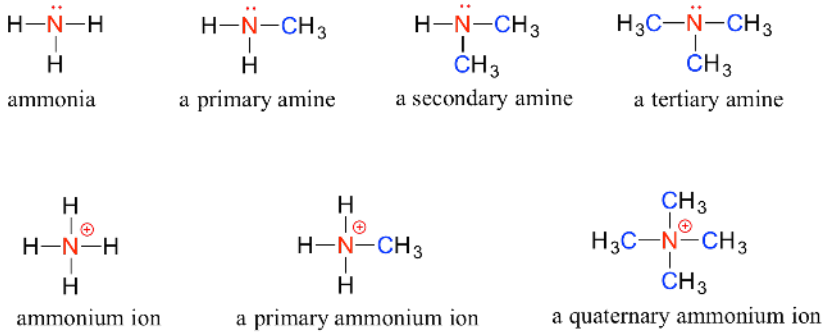
1.2.1.6: Ethers and Sulfides

In an *ether* functional group, oxygen is bonded to two carbons. Below is the structure of diethyl ether, a common laboratory solvent and one of the first compounds to be used as an anesthetic during operations. The sulfur analog of ether is called a *thioether or sulfide*.

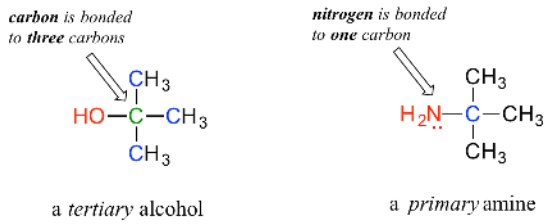


1.2.1.7: Amines

Amines contain nitrogen atoms with single bonds to hydrogen and carbon. Just as there are primary, secondary, and tertiary alcohols, there are primary, secondary, and tertiary amines. Ammonia is a special case with no carbon atoms. One of the most important properties of amines is that they are basic and are readily protonated to form ammonium cations. In the case where nitrogen has four bonds to carbon (which is somewhat unusual in biomolecules), it is referred to as a quaternary ammonium ion.

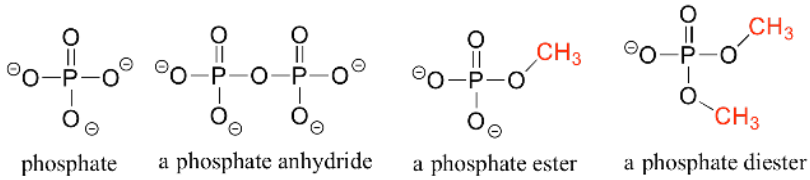


Note: Do not be confused by how the terms ‘primary’, ‘secondary’, and ‘tertiary’ are applied to alcohols and amines – the definitions are different. In alcohols, what matters is how many other carbons the alcohol *carbon* is bonded to, while in amines, what matters is how many carbons the *nitrogen* is bonded to.



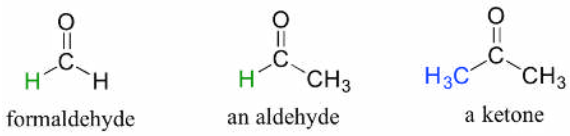
1.2.1.8: Organic Phosphates

Phosphate and its derivative functional groups are ubiquitous in biomolecules. These include *phosphate esters and diesters*. A linkage between two phosphates with a bridging oxygen atom creates a phosphate anhydride.



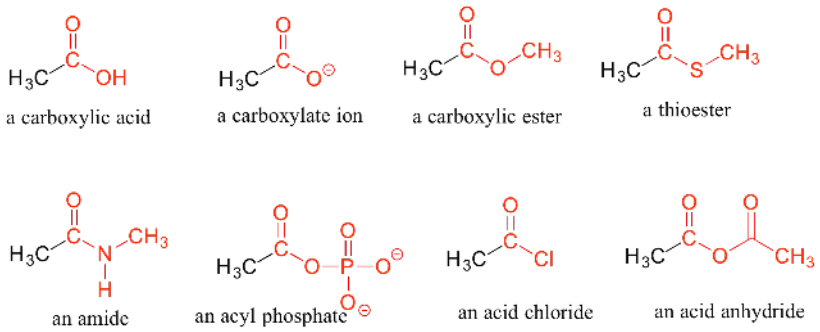
1.2.1.9: Aldehydes and Ketones

Many functional groups contain a carbon-oxygen double bond, commonly called a *carbonyl*. Ketones and aldehydes are two closely related functional groups based on carbonyls that react similarly. In a *ketone*, the carbon atom of a carbonyl is bonded to two other carbons. In an *aldehyde*, the carbonyl carbon is bonded to a hydrogen atom on one side and to a carbon atom on the other. The exception to this definition is formaldehyde, in which the carbonyl carbon has bonds to two hydrogens.



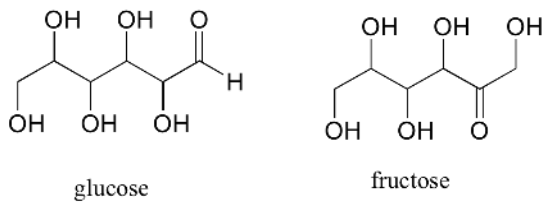
1.2.1.10: Carboxylic Acids and Their Derivatives

When a carbonyl carbon is bonded on one side to a carbon (or hydrogen) and on the other side to an oxygen, nitrogen, or sulfur, the functional group is considered to be one of the carboxylic acid derivatives. The main member of this family is the *carboxylic acid* functional group, in which the carbonyl is bonded to a hydroxyl group. The *carboxylate ion* form has donated the H^+ to the solution. Other derivatives are *carboxylic esters*(usually just called ‘esters’), *thioesters*, *amides*, *acyl phosphates*, *acid chlorides*, and *acid anhydrides*. Except for acid chlorides and acid anhydrides, carboxylic acid derivatives are very common in biological molecules and/or metabolic pathways and will be discussed in further detail in a later chapter.

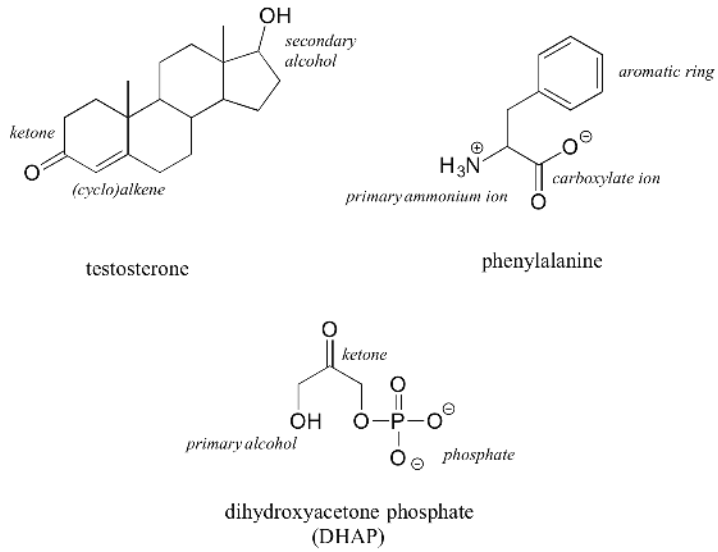


1.2.2: Practice Recognizing Functional Groups in Molecules

A single compound often contains several functional groups, particularly in biological organic chemistry. For example, the six-carbon sugar molecules glucose and fructose contain aldehyde and ketone groups, respectively, containing five alcohol groups. A compound with several alcohol groups is often referred to as a *'polyol'*.

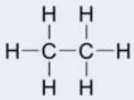
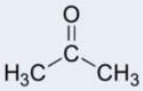
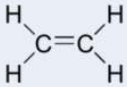
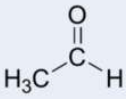
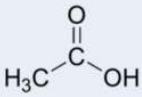
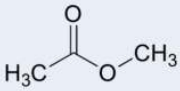
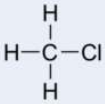
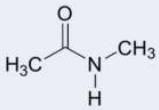


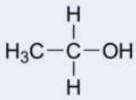
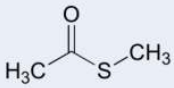
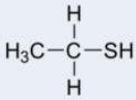
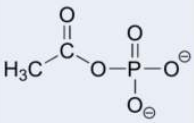
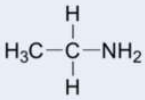
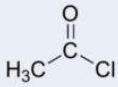
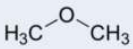
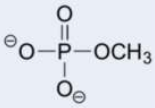
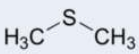
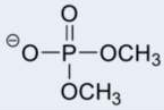
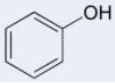
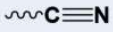
The hormones testosterone and dihydroxyacetone phosphate, as well as the amino acid phenylalanine, all contain multiple functional groups, as labeled below.



Although not exhaustive, this section covers the most important functional groups that we will encounter in biochemistry. Table 1.7 summarizes all the groups listed in this section.

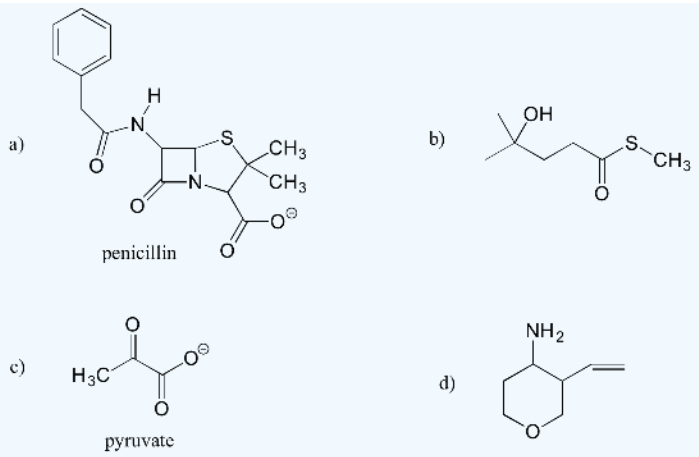
Table 1.7 Common Organic Functional Groups

alkane		ketone	
alkene		aldehyde	
alkyne		carboxylic acid	
aromatic		ester	
alkyl halide		amide	

alcohol		thioester	
thiol		acyl phosphate	
amine		acid chloride	
ether		phosphate monoester	
thioether		phosphate diester	
phenol		nitrile	

Exercise 1.2.1

Identify the functional groups (other than alkanes) in the following organic compounds. State whether alcohols and amines are primary, secondary, or tertiary.



? Exercise 1.2.2

Draw one example of each compound fitting the descriptions below using line structures. Be sure to designate the location of all non-zero formal charges. All atoms should have complete octets (phosphorus may exceed the octet rule). There are many possible correct answers for these, so check your structures with your instructor or tutor.

a. a compound with molecular formula $C_6H_{11}NO$ that includes alkene, secondary amine, and primary alcohol functional groups.
b. an ion with molecular formula $C_3H_5O_6P^{2-}$ that includes aldehyde, secondary alcohol, and phosphate functional groups.
c. A compound with molecular formula C_6H_9NO that has an amide functional group and does *not* have an alkene group.

1.2.3: Primary metabolites

Primary metabolites are components of basic metabolic pathways required for life. They are associated with essential cellular functions such as nutrient assimilation, energy production, and growth/development. Primary metabolites are the building blocks for synthesizing the four major macromolecules in the body: carbohydrates, lipids, proteins, and nucleic acids (DNA and RNA).

Large polymers are made from repeating smaller monomer units (Fig. 1.28). Nucleotides are the monomer units that make up nucleic acids, including DNA and RNA. In contrast, the monomers for proteins are amino acids, carbohydrates are composed of sugar residues, and lipids are composed of fatty acids or acetyl groups.

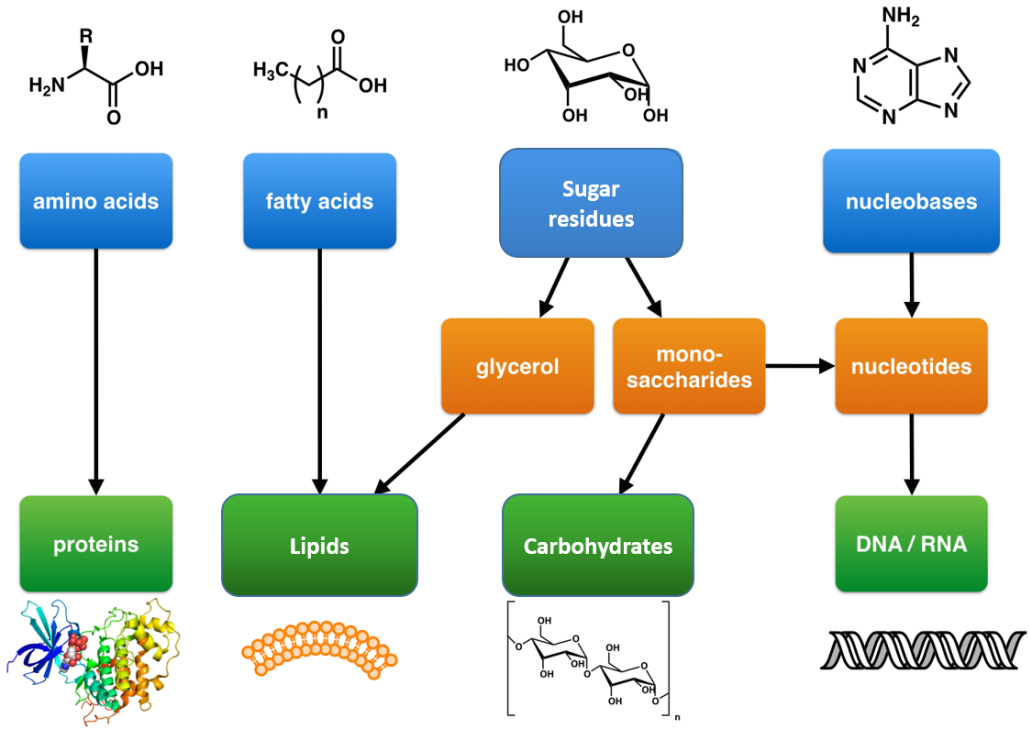


Figure 1.28 The Molecular building blocks of life are made from organic compounds. Modified from: Boghog

1.2.4: Reactions forming the Major Macromolecules

The major macromolecules are constructed by combining repeating monomer subunits through a process known as dehydration synthesis. Interestingly, the organic functional units used in the dehydration synthesis processes for each major macromolecule type have similarities. Thus, it is helpful to look at the reactions together (Figure 1.29)

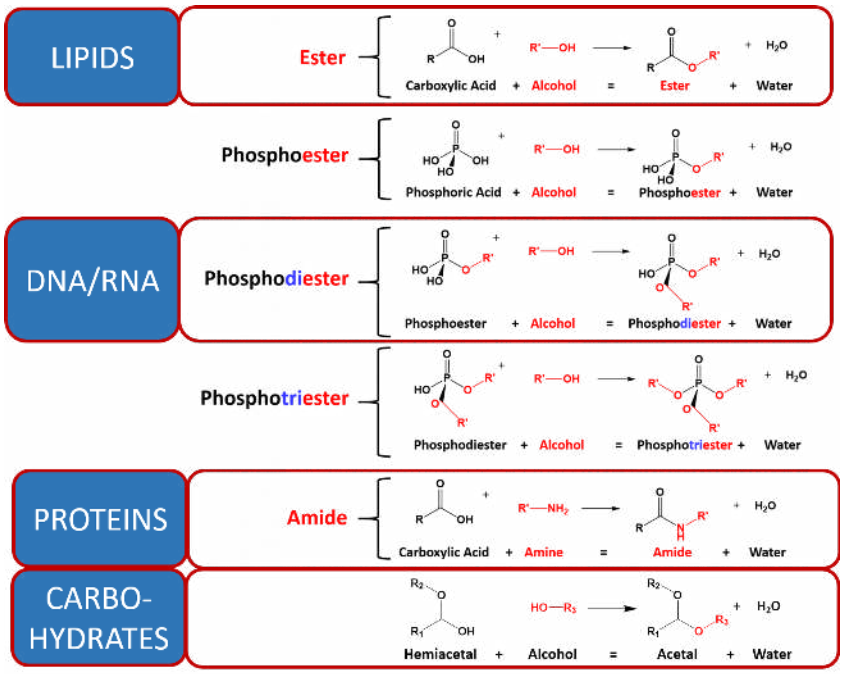


Figure 1.29 Dehydration Synthesis Reactions Involved in Macromolecule Formation. The major organic reactions involved in the biosynthesis of lipids, nucleic acids (DNA/RNA), proteins, and carbohydrates are illustrated. Note that in all of these reactions, there is a functional group that contains two electron-withdrawing groups (the carboxylic acid, phosphoric acid, and the hemiacetal, each having two oxygen atoms attached to a central carbon or phosphorus atom). This forms a reactive, partially positive center atom (carbon in the case of the carboxylic acid and hemiacetal, or phosphorus in the case of the phosphoric acid) that the electronegative oxygen or nitrogen can attack from an alcohol or amine functional group. Within biological systems, many functional groups, such as carboxylic acids, require activation before they can be utilized in synthesis reactions, which will be detailed in later chapters.

Primary metabolites involved with energy production include numerous enzymes that break down food molecules, such as carbohydrates and lipids, and capture the energy released during the hydrolysis of adenosine triphosphate (ATP). *Enzymes* are biological catalysts that speed up chemical reactions. Typically, they are proteins composed of amino acid building blocks. Cells are also composed of primary metabolites. These include cell membranes (e.g., phospholipids), cell walls (e.g., peptidoglycan, chitin), and cytoskeletons (proteins). DNA and RNA, which store and transmit genetic information. Primary metabolites are also involved in cellular signaling, communication, and transport molecules. The structure and function of primary metabolites are key components of this text. These reactions will be detailed in the following chapters.

1.2.5: Summary

This chapter introduces the essential concepts of organic chemistry as they apply to biochemistry, highlighting the central role that carbon-containing compounds play in all life on Earth. It emphasizes that, despite the enormous diversity of organic compounds—numbering in the millions—they are composed of a limited number of recurring structural units known as functional groups. Understanding these groups is fundamental because their predictable reactivity patterns allow us to infer the chemical behavior of unfamiliar compounds.

The chapter begins by discussing the origins of organic molecules and explains how carbon's unique bonding properties enable the formation of complex, diverse molecular architectures. It outlines how functional groups such as hydroxyl, carbonyl, carboxyl, amine, phosphate, and others serve as the "building blocks" that determine the reactivity and interactions of organic molecules. This foundational knowledge is critical for predicting reaction outcomes in both synthetic organic chemistry and biological systems.

A significant portion of the chapter is dedicated to classifying organic compounds based on their structural features. It covers simple hydrocarbons, distinguishing between alkanes (saturated hydrocarbons), alkenes and alkynes (unsaturated hydrocarbons with double and triple bonds, respectively), and aromatic compounds known for their stable, conjugated ring structures. In addition,

the chapter explores various derivatives formed when functional groups are introduced, such as alkyl halides, alcohols, phenols, thiols, ethers, and amines—each with characteristic reactivity and roles in biological systems.

Further, the chapter connects these organic chemistry principles to biochemistry by explaining how dehydration synthesis reactions link monomers into macromolecules. It details the construction of proteins, nucleic acids, carbohydrates, and lipids from their respective monomer units and discusses the importance of functional group activation during these biosynthetic processes. The section on primary metabolites ties together the concept of organic reactivity with cellular function, underscoring the essential roles of these molecules in energy production, cell structure, and signaling.

Overall, the chapter equips students with a framework for understanding the structure and reactivity of organic molecules, laying the groundwork for more advanced studies in biochemistry and molecular biology.

This page titled 1.2: Chemical Foundations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **1.2: Chemical Foundations** by Henry Jakubowski and Patricia Flatt is licensed CC BY 4.0.
- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

1.3: Physical-Chemical Foundations

Learning Goals (ChatGPT o3-mini)

1. **Understand the Driving Forces of Reactions**
   - Explain how energy differences (ΔG) between reactants and products dictate the spontaneity and extent of chemical reactions.
   - Use the "ball on a hill" analogy to relate potential energy differences to reaction spontaneity.
2. **Distinguish Reversible and Irreversible Reactions**
   - Describe the characteristics of irreversible reactions versus reversible reactions and illustrate how reaction progress and equilibrium are reflected in concentration-time profiles.
   - Interpret different scenarios (favoring products, favoring reactants, or balanced) based on changes in reactant and product concentrations.
3. **Apply Equilibrium Concepts**
   - Define the equilibrium constant (K_{eq}) and explain how it quantitatively reflects the relative favorability of reactants versus products.
   - Relate the value of K_{eq} (greater than, less than, or equal to 1) to the extent of reaction and spontaneous direction.
4. **Analyze the Role of Concentration in Reaction Energetics**
   - Differentiate between the intrinsic stability of reactants/products (ΔG^o) and the influence of concentration on reaction spontaneity via the reaction quotient (Q_{rx}).
   - Explain how changing concentrations (as per Le Chatelier's Principle) can shift equilibrium positions.
5. **Calculate and Interpret Gibbs Free Energy Changes**
   - Use the relationship $\Delta G = \Delta G^o + RT \ln Q$ to calculate the free energy change under non-standard conditions.
   - Discuss how the sign and magnitude of ΔG determine whether a reaction proceeds toward products, remains at equilibrium, or reverses.
6. **Connect Thermodynamic Parameters: Enthalpy and Entropy**
   - Define enthalpy (ΔH) and entropy (ΔS), and explain their roles in determining the spontaneity of both exothermic and endothermic reactions.
   - Compare and contrast exothermic ($\Delta H < 0$) and endothermic ($\Delta H > 0$) processes in the context of biological reactions.
7. **Integrate the Second Law of Thermodynamics in Biochemical Contexts**
   - Explain the significance of the second law of thermodynamics ($\Delta S_total > 0$) and how it governs spontaneous processes.
   - Relate macroscopic thermodynamic changes (ΔG_sys) to microscopic rearrangements and energy dispersion (ΔS_{sys} and ΔS_{surr}) in biochemical systems.
8. **Apply Thermodynamic Principles to Real-World Biochemical Reactions**
   - Analyze case studies, such as the reactions of strong versus weak acids with water, to illustrate how intrinsic molecular stability and concentration effects combine to determine reaction direction and extent.
   - Evaluate the thermodynamic rationale behind phenomena like water evaporation and molecular binding, and discuss how these principles underpin cellular metabolism.

By achieving these learning goals, students will be equipped to analyze and predict the behavior of biochemical reactions, understand the interplay between energy, equilibrium, and entropy, and apply these concepts to complex biological systems.

The types and numbers of chemical reactions in biological cells are staggering. Compared to physical and chemical reactions in a controlled and closed environment, biological reactions occur in open systems with input and output of energy and chemical

"feedstocks." Yet, they are governed by the same physical principles that control all reactions. We can gain insight into biological reactions and how they are controlled by considering the principles you have used in many preceding classes: energy changes, equilibria, and thermodynamics. Let's review them!

1.3.1: Reactions and Energy Changes

Why do reactions vary in extent from completely irreversible in the forward reaction to reversible reactions favoring the reactants? It may be helpful to understand a simple physical reaction before attempting more complex chemical reactions. Let's start with a simple ball on a hill. Does a ball at the top of a hill roll downhill spontaneously, or does the opposite happen? No one has ever seen a ball roll spontaneously uphill unless a lot of energy was added to the ball. This physical reaction appears irreversible and is driven to the lower potential energy at the bottom of the hill. The potential energy gap is related to the "extent" and spontaneity of this reaction. As we have undoubtedly observed, natural processes tend to go to a lower energy state. By analogy, we will consider the driving force for a chemical reaction to be the free energy difference, ΔG, between reactants and products. ΔG determines the extent and spontaneity of the reaction.

1.3.1.1: Reversible/Irreversible Reactions, Extent of Reactions, Equilibria

Consider a hypothetical reversible reaction in which you start with some reactants, $\mathbf{A}$ and $\mathbf{B}$, each at a 1 M concentration (1 mol of each/L solution). but no products, $\mathbf{P}$ and $\mathbf{Q}$. For ease, assume that the total volume of solution is 1 L, so that we start with 1 mol each of $\mathbf{A}$ and $\mathbf{B}$. At time $t = 0$, the concentration of products is 0. The reaction can be written as:

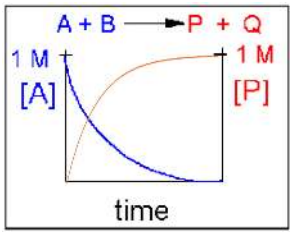
$$\mathrm{A} + \mathrm{B} \rightleftharpoons \mathrm{P} + \mathrm{Q}.$$

As time progresses, the amounts or concentrations of $\mathbf{A}$ and $\mathbf{B}$ decrease as the amounts or concentrations of products $\mathbf{P}$ and $\mathbf{Q}$ increase. At some point, no further changes occur in the amounts or concentrations of the remaining reactants or products. At this point, the reaction is in equilibrium, a term often used in our common vocabulary to denote a system undergoing no net change.

Most of the reactions we will study occur in solution, so we will deal with concentrations (mol/L or mmol/mL = M). Let's consider how the concentrations of reactants and products change over time. Depending on the extent to which a reaction is reversible, four different scenarios can be imagined:

Scenario 1: Irreversible reaction in which the reverse reaction occurs to a negligible extent.

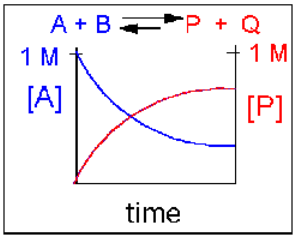
In this reaction, the reverse reaction occurs to such a small extent that it can be neglected. The only reaction that occurs is the conversion of reactants to products. Hence, all the reactants are converted into the product. At equilibrium [A] = 0. Since 1 mol of A reacted, it must form 1 mol of P and 1 mol of Q - i.e., the concentration of products at equilibrium is 1 M. At an earlier time of the reaction (let's pick a time when [A] = 0.8 M), only part of the reactants have reacted (in this case 0.2 M), producing an equal amount of products, P and Q. Graphs of [A] and [P] as a function of time are shown below. [A] decreases nonlinearly to 0 M while [P] increases reciprocally to 1 M concentration. This is illustrated in the graph below.



Examples of irreversible reactions include those of strong acids (such as nitric, sulfuric, and hydrochloric) with bases (OH- and water), as well as the much more complex combustion reactions, such as the burning of sugars (like those found in trees) and hydrocarbons (like octane) to form CO_2 and H_2O.

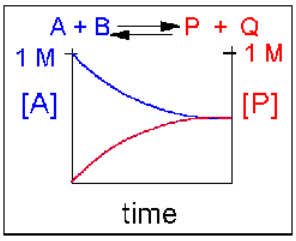
Scenario 2: Reversible reaction in which the forward reaction is favored.

Again, [A] decreases and [P] increases, but in this case, some [A] remains because the reaction is reversible. As [A]and [B] decrease, [P] and [Q] increase, which increases the chance that they will collide and form the product. Since P and Q can react to form reactants, the [A] at equilibrium is not zero, as is shown below.



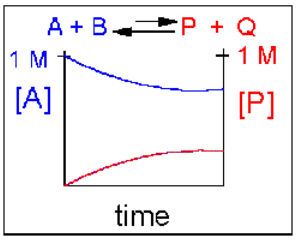
Scenario 3: Reversible Reaction in which forward and reverse reactions are equally favored.

Again, [A] decreases and [P] increases, but in this case, some [A] remains because the reaction is reversible. As [A] and [B] decrease, [P] and [Q] increase, thereby increasing the likelihood that they will collide and form the product. Since P and Q can react to form reactants, the [A] at equilibrium is not zero, as is shown below. Because the reactants and products are equally favored, their concentrations will be equal at equilibrium.



Scenario 4: Reversible Reaction in which the reverse reaction is favored.

Again, [A] decreases, and [P] increases, but in this case, some A remains since the reaction is reversible. As [A] and [B] decrease, [P] and [Q] increase, thereby increasing the likelihood that they will collide and form the product. Since P and Q can react to form reactants, the [A] at equilibrium is not zero, as shown below. Because the reaction favors the reactants, their concentration will be higher at equilibrium than that of the products.



An example of this kind of reaction, which favors the reactants, is the reaction of acetic acid (a weak acid) with water.

$$\mathrm{CH_3CO_2H\,(aq) + H_2O\,(l) \rightleftharpoons H_3O^+\,(aq) + CH_3CO_2{}^-\,(aq)}$$

1.3.2: Equilibrium Constants

Without much experience in chemistry, it is challenging to examine the structures of the reactants and products and determine whether the reaction is irreversible or reversible, favoring either the reactants or the products (except for obvious irreversible reactions described above). However, this data can be found in tables of equilibrium constants. The equilibrium constant, as its name implies, is constant, independent of the concentration of the reactants and products. A $K_{eq} > 1$ implies that the products are favored. A $K_{eq} < 1$ implies that reactants are favored. Reactants and products are equally favored when $K_{eq} = 1$. For the more general reaction,

$$a\,\mathrm{A} + b\,\mathrm{B} \rightleftharpoons p\,\mathrm{P} + q\,\mathrm{Q}$$

where a, b, p, and q are the stoichiometric coefficients,

$$\mathrm{K_{eq}} = \frac{[\mathrm{P}]^{\mathrm{p}}_{\mathrm{eq}}[\mathrm{Q}]^{\mathrm{q}}_{\mathrm{eq}}}{[\mathrm{A}]^{\mathrm{a}}_{\mathrm{eq}}[\mathrm{B}]^{\mathrm{b}}_{\mathrm{eq}}} \tag{1.3.1}$$

where all the concentrations are at their equilibrium values. For a simple reaction where a, b, p, and q are all 1, then $K_{eq} = ([P][Q])/([A][B])$.

(Note: Equilibrium constants are truly constant only at a given temperature, pressure, and solvent condition. Likewise, they depend on concentration to the extent that their activities change with concentration.)

For an irreversible reaction, such as the reaction of a 0.1 M HCl (aq) in water, $[HCl]_{eq}$ = 0, you cannot easily measure a K_{eq}. However, if we assume the reaction goes in reverse to an almost imperceptible degree, $[HCl]_{eq}$ might equal 10-10 M. Hence K_{eq} >> 1.

In summary, the extent of different reactions can range from completely irreversible (favoring only the products) to reversible (favoring the reactants).

Our next goal is to understand what controls the extent of a reaction. That is, of course, the change in the Gibbs free energy. Two different pairs of factors influence the ΔG. One pair is the reactants' concentration and their inherent reactivity relative to the products (as reflected in the K_{eq}). The other pair is enthalpy/entropy changes. We will now consider the first pair.

Contributions of Molecule Stability (K_{eq}) and Concentration to ΔG

Consider the reactions of hydrochloric acid and acetic acid with water.

$$\mathrm{HCl\,(aq) + H_2O\,(l) \longrightarrow H_3O^+\,(aq) + Cl^-\,(aq)}$$
$$\mathrm{CH_3CO_2H\,(aq) + H_2O\,(l) \longrightarrow H_3O^+\,(aq) + CH_3CO_2{-}\,(aq)}$$

Assume that at t = 0, each acid is placed into water at a concentration of 0.1 M. When equilibrium is reached, essentially no HCl remains in the solution. In contrast, 99% of the acetic acid remains. Why are they so different? We rationalized that HCl (aq) is a much stronger acid than H_3O^+(aq), which itself is a much stronger acid than CH_3CO_2H (aq). Why? Something about the structure of these acids (and the bases) makes HCl much more intrinsically unstable, much higher in energy, and hence much more reactive than the acid it forms, H_3O^+(aq). Likewise, H_3O^+(aq) is much more intrinsically unstable, much higher in energy, and hence more reactive than CH_3CO_2H (aq). This has nothing to do with concentration, since the initial concentration of both HCl (aq) and CH_3CO_2H (aq) was identical. This observation is reflected in the K_{eq} for these acids (>>1 for HCl and <<1 for acetic acid). This difference in the intrinsic stability of reactants relative to products (which is independent of concentration) is one factor contributing to ΔG.

The other factor is concentration. A 0.25 M (0.25 mol/L or 0.25 mmol/ml) solution of acetic acid does not conduct electricity, implying that very few ions of H_3O^+(aq) + $CH_3CO_2^-$ (aq) exist in the solution. However, a dim light becomes evident if more concentrated acetic acid is added. Adding more reactants appeared to drive the reaction toward more products, even though the reverse reaction is favored based on the intrinsic stability of reactants and products alone. Before the concentrated acid was added, the system was at equilibrium. Adding concentrated acid perturbed the equilibrium, which drove the reaction to form additional products. This is an example of Le Châtelier's Principle, which states that if a reaction at equilibrium is perturbed, the reaction will shift in the direction that counteracts the perturbation. Hence, if

- more reactant is added, the reaction shifts to form more products
- more product is added, the reaction shifts to form more reactants
- products are selectively removed (by distillation, crystallization, or further reaction to produce another species), the reaction shifts to form more products.
- reactants are removed (as above), the reaction shifts to form more reactants.
- heat is added to an exothermic reaction, the reaction shifts to remove the excess heat by shifting to form more reactants. (opposite for an endothermic reaction).

1.3.3: Change in Free Energy G

Without doing a complicated derivation, these simple examples suggest that the total ΔG can be expressed as the sum of the two contributions, showing the effects of the intrinsic stability (K_{eq}) and concentration:

$$\Delta G_{\text{total}} = \Delta G_{\text{stability}} + \Delta G_{\text{concentration}} \tag{1.3.2}$$

which becomes for the simple reaction $\mathrm{A + B \rightleftharpoons P + Q}$ (after a rigorous derivation):

$$\Delta G = \Delta G^0 + RT \ln \frac{[P][Q]}{[A][B]} \\ = \Delta G^0 + RT \ln Q_{rx} \tag{1.3.3}$$

where ΔG^o reflects the contribution from the relative intrinsic stability of reactants and products, and the second term reflects the contribution from the relative concentrations of reactants and products (which has nothing to do with stability). Q_{rx} is the **reaction quotient**, which for the reaction A + B ↔ P + Q is given by:

$$Q_{rx} = ([P][Q])/([A][B]) \tag{1.3.4}$$

at any point in the reaction.

Meaning of ΔG

Remember that ΔG is the "driving" force for a reaction, analogous to the difference in potential energy for a ball on a hill. Go back to that analogy. If the ball starts at the top of the hill, does it roll downhill? Of course. It transitions from high to low potential energy. The reaction can be written as:

$$\text{Ball}_{\text{top}} \rightarrow \text{Ball}_{\text{bottom}}$$

for which the change in potential energy, $\Delta PE = PE_{bottom} - PE_{top} < 0$. If the ball starts at the bottom, will it go to the top? Obviously not. For that reaction,

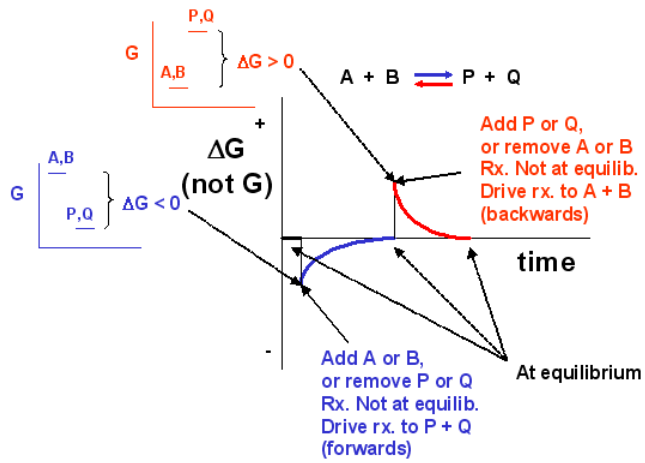
$$\text{Ball}_{\text{bottom}} \rightarrow \text{Ball}_{\text{top}}, \Delta PE > 0.$$

If the top of the hill were at the same height as the bottom of the hill (obviously an absurd situation), the ball would not move. It would effectively be at equilibrium, a state of no change. For this reaction, Ball $_{top}$ --> Ball $_{bottom}$, the ΔPE = 0. As the ball starts rolling down the hill, its potential energy approaches the potential it would have at the bottom. Hence, the ΔPE changes from negative to increasingly positive until it reaches its minimum value, at which point ΔPE = 0, and movement ceases. If the ΔPE is not 0, the ball will move until the ΔPE = 0.

Likewise, for a chemical reaction that favors products, ΔG < 0. The system is not at equilibrium, and the reaction will go toward products. As the reaction proceeds, products build up, and there is less of a driving force for reactants to go to products (Le Chatelier's Principle), so the ΔG becomes more positive until the ΔG = 0 and the reaction is at equilibrium. A reaction with a ΔG > 0 is likewise not at equilibrium, so it will go in the appropriate direction until equilibrium is reached. Hence, for the reaction A + B <==> P + Q,

- if ΔG < 0, the reaction goes toward products P and Q
- if ΔG = 0, the reaction is at equilibrium, and no further change occurs in the concentration of reactants and products.
- if ΔG > 0, the reaction goes toward reactants A and B.

We cannot easily measure the actual free energy (G) of reactants or products, but we can readily determine ΔG. These points are illustrated in the graph below of ΔG vs time for the hypothetical reaction A + B ↔ P + Q. (Also notice the two insert graphs - in blue and red - which show, in analogy to the ball on the hill graphs, the values of ΔG at the two points where the perturbation to the equilibrium was made.)



Notice that ΔG changes continuously until the system reaches equilibrium. Initially, the equilibrium is perturbed, causing the system to deviate from equilibrium (shown in blue). The perturbation favored the products. After reaching equilibrium, the system was perturbed again to favor the reverse reaction. Notice, in this case, the ΔG for the reaction as written: A + B ↔ P + Q is positive - i.e., it is not in equilibrium. Therefore, the reaction (as written) goes backward to reactants. It is essential to recognize that the reported ΔG represents the reaction as written.

Now let's apply $\Delta G = \Delta G^o + RT\ln Q = \Delta G^o + RT\ln ([P][Q])/([A][B])$ to two reactions we discussed above:

- $HCl(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + Cl^-(aq)$
- $CH_3CO_2H(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + CH_3CO_2^-(aq)$

At time t = 0, 0.1 mol of HCl and CH_3CO_2H were added to two different beakers. The forward reaction is favored at this point, but obviously to different extents. The RTln Q would be identical for both acids, since each reactant is at 0.1 M, and no products exist. However, the $\Delta\mathbf{G}^o$ is negative for HCl and positive for acetic acid since HCl is a strong acid. Hence, at t=0, **ΔG** for the HCl reaction is much more negative than for acetic acid. This is summarized in the table below. The direction of the arrow shows if products (-->) or reactants (<---) are favored. The size of the arrow indicates approximately the extent to which the ΔG term is favored

Reaction at t=0	$\Delta\mathbf{G}^o$	RTln Q	ΔG
$HCl(aq) + H_2O(l)$	---------->	---------->	---------->
$CH_3CO_2H(aq) + H_2O(l)$	<----------	---------->	->

Now, when equilibrium is reached, no net change occurs in the concentration of reactants and products, and ΔG = 0. In the case of HCl, there is just an infinitesimal amount of HCl left and 0.1 M of each product, so concentration favors HCl formation. However, the intrinsic relative stability of reactants and products still favors products. In the case of acetic acid, most acetic acid remains (0.099 M) with little product (0.001 M), so concentration favors products. However, the intrinsic relative stability of reactants and products still favors reactants. This is summarized in the table below.

Reaction at equilibrium.	ΔG^ostab	RTln Q	ΔG
$HCl(aq) + H_2O(l)$	---------->	<----------	favors neither, = 0
$CH_3CO_2H(aq) + H_2O(l)$	<----------	---------->	favors neither, = 0

Compare the two tables above (one at time t=0 and the other at equilibrium). Notice:

- ΔG^o does not change in a given set of conditions since it has nothing to do with concentration.
- Only RTln Q changes during a reaction until equilibrium is achieved.

Meaning of ΔG^o

To get a better understanding of the significance of ΔG^o, let's consider the following equations under two different conditions:

$$\Delta G = \Delta G^0 + RT \ln \frac{[P][Q]}{[A][B]} = \Delta G^0 + RT \ln Q_{rx} \tag{1.3.5}$$

Condition I: Reaction at equilibrium, ΔG = 0

The equation reduces to:

$$\Delta G^0 = -RT \ln \frac{[P]_{eq}[Q]_{eq}}{[A]_{eq}[B]_{eq}} = -2.303RT \log K_{eq} \tag{1.3.6}$$

This supports our idea that ΔG^o is independent of concentration since K_{eq} should also be independent of concentration.

Condition II: The concentration of all reactants and products is 1 M (standard state, assuming solution reaction)

The equation reduces to:

$$\Delta G = \Delta G^0 + RT \ln \frac{[1][1]}{[1][1]} = \Delta G^0 + 2.303RT \log 1 = \Delta G^0$$

$$\begin{aligned} \Delta G &= \Delta G^o + RT \ln \left(\frac{[1][1]}{[1][1]} \right) \\ &= \Delta G^o + 2.303RT \log 1 \\ &= \Delta G^o \end{aligned} \tag{1.3.7}$$

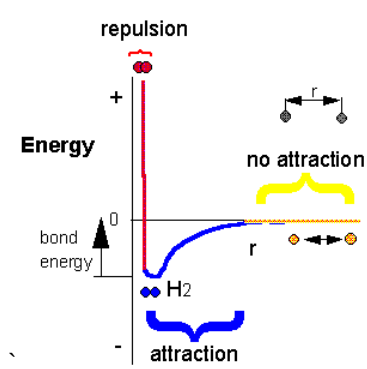
This implies that when all reactants are at this concentration, defined as the standard state (1 M for solutes), the ΔG at that particular moment is the ΔG^o for the reaction. If one of the reactants or products is H_3O^+, it would make little biological sense to calculate ΔG^o for the reaction using the standard state of $[H_3O^+]$ = 1 M or a pH of zero. Instead, it is assumed that the pH = 7, $[H_3O^+] = 10^{-7}$ M. A new symbol is used for ΔG^o under these conditions, $\Delta G^{o'}$.

1.3.4: Heat, Enthalpy, and Entropy

Consider the association reaction of hydrogen atoms into molecular hydrogen

$$H + H \longrightarrow H_2.$$

Does this reaction occur spontaneously? It does. You should remember that individual **H** atoms are unstable since they don't have a completed outer shell of electrons - in this case, a duet. As they approach, they can interact to form a covalent bond, releasing energy in the process. The bonded state is lower in energy than two separated H atoms. This should be clear since energy has to be added to a molecule of H_2 to break the bond. We refer to this as the **bond dissociation energy**.



Now consider a more complicated reaction, the burning of octane.

$$2\,C_8H_{18}\,(l) + 25\,O_2\,(g) \longrightarrow 16\,CO_2\,(g) + 18\,H_2O\,(g)$$

To carry out this reaction, every C-C, C-H, and O-O bond in the reactants must be broken (each of which requires an input of energy), but energy is released when the products' C-O and H-O covalent bonds are formed. Is more energy needed to break the

bonds in the reactants, or is more energy released when forming bonds in the product? The answer should be clear. The products must have a lower energy than the reactants, as large amounts of heat and light energy are released during the combustion of gasoline and other hydrocarbons.

These reactions suggest that energy must be released for a reaction to proceed to any extent in a given direction.

Now consider, however, the following reaction:

$$\mathrm{Ba(OH)_2 \cdot 8\,H_2O\,(s) + 2\,NH_4SCN\,(s) \longrightarrow 10\,H_2O\,(l) + 2\,NH_3\,(g) + Ba(SCN)_2\,(aq)}$$

When these two solids are added to a beaker and stirred, a reaction occurs, as evidenced by the formation of a liquid (water) and the smell of ammonia. Surprisingly, heat is not released into the surroundings in this reaction. Instead, heat was absorbed from the surroundings, causing the beaker to become so cold that it froze to a piece of wood (with a layer of water added to the wood) on which it was placed. This reaction appears to contradict our understanding that a reaction proceeds in a direction where heat is liberated. Reactions that liberate heat and raise the temperature of the surroundings are called **exothermic** reactions. Reactions that absorb heat from the surroundings and, hence, lower the surroundings' temperature are **endothermic** reactions. To answer the question, we need to consider entropy.

1.3.5: A review of thermodynamics

You may remember from General Chemistry that the change in the internal energy of a system, ΔE, is given by:

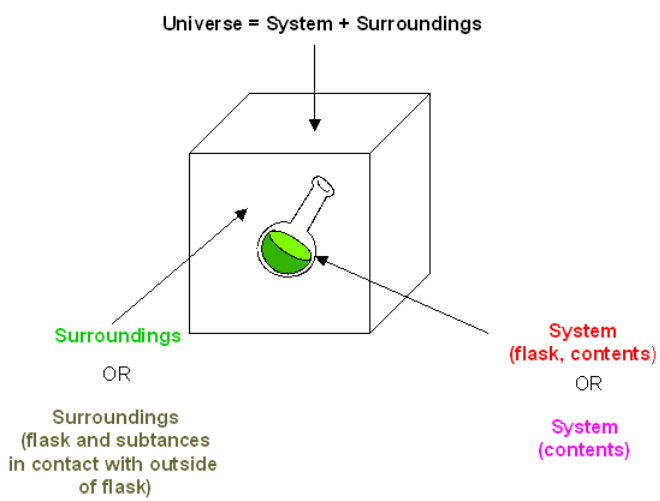
$$\begin{aligned} \Delta E_{sys} &= q + w \\ &= q - P_{ext}\Delta V \end{aligned} \tag{1.3.8}$$

where q is the heat (thermal energy) transferred to (+) or from the system (-), w is the work done on (+) or by (-) the system. This is one expression for the 1st Law of Thermodynamics

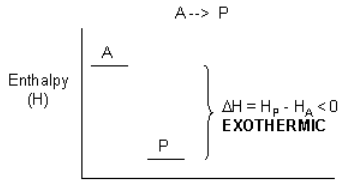
If only pressure/volume (PV) work is done (and not electrical work, for example), $w = -P_{ext}\Delta V$, where P_{ext} is the external pressure resisting a volume change in the system, ΔV. Under these conditions, the heat transfer at constant P, q_P is given by:

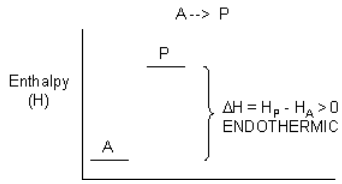
$$\begin{aligned} \mathrm{\Delta E_{sys} - w = \Delta E_{sys} + P_{ext}\Delta V} \\ &= \mathrm{q_P} \\ &= \mathrm{\Delta H_{sys}} \end{aligned} \tag{1.3.9}$$

q_p, which can easily be measured in a coffee cup calorimeter, is equal to the **change in enthalpy**, ΔH, of the system.



In exothermic reactions, the reactants have more thermal energy than the products, and the heat released (measured in kilocalories or kilojoules) is the difference between the energy of the products and the reactants. When heat energy is used to measure the difference in energy, we refer to the energy as enthalpy (H) and the heat released as the change in enthalpy (ΔH), as illustrated below.





For exothermic reactions, $\Delta H < 0$. For endothermic reactions, $\Delta H > 0$.

The equation $\Delta E_{sys} = q + w = q - P_{ext}\Delta V$ is one expression of the First Law of Thermodynamics. Another statement of energy conservation is:

$$\Delta E_{tot} = \Delta E_{universe} = \Delta E_{sys} + \Delta E_{surrounding} = 0 \tag{1.3.10}$$

Something other than heat being released from the system decides whether a reaction proceeds to a significant extent. That is, the spontaneity of a reaction must depend on more than just ΔH_{sys}. Another example of a spontaneous natural reaction is the evaporation of water (a physical, not chemical, process).

$$H_2O(l) \longrightarrow H_2O(g)$$

Heat is absorbed from the surroundings to break the intermolecular forces (H-bonds) among the water molecules (the system), thereby turning the liquid into a gas. If the surroundings are the skin, evaporating sweat cools the body. Why are these reactions spontaneous and essentially irreversible even though they are endothermic? Notice that in both of these endothermic reactions (the reactions of $Ba(OH)_2.8H_2O(s)$ and $2NH_4SCN(s)$ and the evaporation of water), the products are more disorganized (more disordered) than the reactants. A solid is more ordered than a liquid or a gas, and a liquid is more ordered than a gas. In nature, ordered things tend to become more disordered over time. Entropy (S), the other factor (in addition to enthalpy changes), is often considered to measure a system's disorder. The greater the entropy, the greater the disorder. For reactions that go from order (low S) to disorder (high S), the change in S, $\Delta S > 0$. For the reaction that proceeds from low order to high order, $\Delta S < 0$.

Caution

However, this common description of entropy is quite misleading. Macroscopic examples describing ordered/disordered states (such as the cleanliness of your room or the shuffling of a deck of cards) are inappropriate since entropy deals with microscopic states.

The driving force for spontaneous reactions is the dispersion of energy and matter. Increases in entropy for reactions involving matter occur when gases or solutes in solution are dispersed, thereby increasing positional entropy. For reactions involving energy changes, entropy increases when energy is dispersed as random, undirected thermal motion, increasing thermal entropy. In this sense, entropy, S (a measure of "spreadedness") is a measure of the number of different ways (microstates) that particles or energy can be arranged (W), not a measure of disorder! W is an abbreviation for the German word ***Wahrscheinlichkeit***, which means probability. It can be shown that for a solute dissolving in a solvent, $W_{sys} = W_{solute} \times W_{solvent}$. Note that this is a multiplicative function. Entropy is a logarithmic function of W, allowing the additivity of solute and solvent W values —a feature also found in other thermodynamic state functions, such as ΔE, ΔH, and ΔS. Hence

$$\ln W \text{ sys } = \ln W_{solute} + \ln W_{solvent} \tag{1.3.11}$$

Boltzmann showed that for molecules,

$$S = k \ln W \tag{1.3.12}$$

where k is the Boltzmann constant (1.68 x 10^{-23} J/K), S units: J/K

or

$$S = kN_A \ln W = R \ln W \tag{1.3.13}$$

Boltzmann realized the connection between the macroscopic entropy of a system and the microscopic order/disorder of a system through the equation $S = k \ln W$. Increasing S (macroscopic property) occurs with an increasing number of possible microscopic states for the atoms and molecules in a system.

The dissolution of a solute in water and the expansion of a gas into a vacuum, both of which proceed spontaneously toward an increase in matter dispersal, are examples of familiar processes characterized by $\Delta S_{sys} > 0$. We will see in future chapters that entropy changes in the solvent, solutes, and protein are critical determinants of protein folding.

The spontaneity of exothermic and endothermic processes will depend on the

$$\Delta S_{tot} = \Delta S_{\text{surr}} + \Delta S_{\text{sys}} \tag{1.3.14}$$

ΔS_{sys} often depends on matter dispersal (positional entropy). ΔS_{surr} depends on energy changes in the surroundings, $\Delta H_{surr} = -\Delta H_{sys}$ (thermal entropy).

It is more convenient to express thermodynamic properties in terms of the system being studied rather than its surroundings. This can be readily done for the $\Delta S_{surr,}$ which depends on ΔH_{sys} and the temperature. First, consider the dependency on ΔH_{sys}. Thermal energy flows into or out of the system, and since $\Delta H_{sys} = - \Delta H_{surr}$,

ΔS_{surr} is proportional to $-\Delta H_{sys}$

- For an exothermic reaction, $\Delta S_{surr} > 0$ (since $\Delta H_{sys} < 0$) and the reaction is favored;
- For an endothermic reaction, $\Delta S_{surr} < 0$, (since $\Delta H_{sys} > 0$), and the reaction is disfavored;

ΔS_{surr} also depends on the temperature T of the surroundings:

ΔS_{surr} is proportional to 1/T

If the T_{surr} is high, a given heat transfer to or from the surroundings will have a smaller effect on the ΔS_{surr}. Conversely, if the T_{surr} is low, the effect on ΔS_{surr} will be greater. (Atkins uses the analogy of the effect of a sneeze in a library compared to a crowded street; An American Chemistry General Chemistry text uses the analogy of giving $5 to a friend with $1000 compared to one with just $10.) Hence,

$$\Delta S_{\text{surr}} = \frac{-\Delta \text{H}_{\text{sys}}}{\text{T}} \tag{1.3.15}$$

(Note: from a more rigorous thermodynamic approach, entropy can be determined from

$$\text{dS} = \frac{\text{dq}_{\text{rev}}}{\text{T}} \tag{1.3.16}$$

Once again,

$$\Delta S_{\text{tot}} = \Delta S_{\text{surr}} + \Delta S_{\text{sys}} \tag{1.3.17}$$

ΔS_{tot} depends on both enthalpy changes in the system and entropy changes in the surroundings. Hence,

$$\Delta S_{\text{tot}} = \frac{-\Delta \text{H}_{\text{sys}}}{\text{T}} + \Delta \text{S}_{\text{sys}} \tag{1.3.18}$$

Multiplying both sides by $-T$ gives

$$-\text{T}\Delta \text{S}_{\text{tot}} = \Delta \text{H}_{\text{sys}} + \text{T}\Delta \text{S}_{\text{sys}} \tag{1.3.19}$$

The thermodynamic function Gibb's Free Energy, G, can be defined as:

$$G = H - TS \tag{1.3.20}$$

At constant T and P,

$$\Delta G = \Delta H - T\Delta S \tag{1.3.21}$$

Hence

$$\Delta G_{sys} = \Delta H_{sys} - T\Delta S_{sys} = -T\Delta S_{tot} \tag{1.3.22}$$

Spontaneity is determined by ΔS_{tot} OR ΔG_{sys} since $\Delta S_{tot} = -\Delta G_{sys}/T$. ΔG_{sys} is widely used in discussing spontaneity since it is a state function, depends only on the enthalpy and entropy changes in the system, and is negative (as is the potential energy change for a falling object) for all spontaneous processes.

The second law of thermodynamics can be succinctly stated: For any spontaneous process, the $\Delta S_{tot} > 0$. Unlike energy (from the First Law), entropy is not conserved.

1.3.6: Summary

Chapter Summary: Reaction Energetics and Thermodynamics in Biological Systems

This chapter delves into the fundamental principles governing the vast array of chemical reactions that occur within biological cells. It begins by framing these reactions in the context of open systems, where both energy and chemical feedstocks flow in and out. Yet, the underlying behavior is dictated by the same thermodynamic laws that control all chemical processes.

Key Concepts and Analogies

- The chapter introduces the concept of free energy (ΔG) as the driving force of reactions, using the simple "ball on a hill" analogy to illustrate how reactions move from higher to lower energy states.
- It explains how the extent and spontaneity of reactions are determined by the free energy difference between reactants and products, much like a ball rolling downhill due to a drop in potential energy.

Reversible and Irreversible Reactions

- The discussion distinguishes between irreversible reactions (in which the reverse reaction is negligible) and reversible reactions, which can reach equilibrium.
- Four scenarios are described, ranging from reactions that go to completion (irreversible) to those in which the equilibrium favors either the products or the reactants, or is balanced equally.

Equilibrium and the Reaction Quotient

- The equilibrium constant (K_eq) is introduced as a measure of the relative favorability of products versus reactants, with K_eq values greater than, less than, or equal to 1 indicating different extents of reaction.
- The reaction quotient (Q_rx) is presented as a snapshot of the reaction state at any point in time, with its relationship to ΔG given by the equation:
  $\Delta G=\Delta G^0+RT\ln Q_{rx}$

Intrinsic Stability Versus Concentration Effects

- The chapter distinguishes between ΔG^0, which reflects the intrinsic energy differences (or stability) of reactants and products, and the contribution of reactant/product concentrations.
- Examples such as the reactions of HCl and acetic acid with water show that, even when starting with equal concentrations, the inherent stability differences drive the reactions to different extents.

Thermodynamics: Enthalpy and Entropy

- The text reviews the first and second laws of thermodynamics, establishing that energy conservation (ΔE = q + w) and the concept of entropy (S) are crucial for understanding reaction spontaneity.
- Enthalpy (ΔH) is discussed in the context of exothermic (ΔH < 0) and endothermic (ΔH > 0) reactions, while entropy changes (ΔS) capture the degree of dispersal or "spread" of energy and matter.
- The Gibbs free energy equation, $\Delta G=\Delta H-T\Delta S$, is derived and emphasized as a critical tool for predicting whether a reaction will proceed spontaneously.

Integration and Application

- The chapter culminates in the integration of these concepts by showing how both the energetic (enthalpy) and dispersive (entropy) factors combine to determine the system's overall free energy change.
- Examples such as the evaporation of water and the reaction between $Ba(OH)_2 \cdot 8H_2O$ and NH_4SCN illustrate that even endothermic processes can be spontaneous when accompanied by a significant increase in entropy.

In summary, this chapter provides a comprehensive framework for understanding the energetics of biochemical reactions. It equips students with the tools to analyze how free energy, equilibrium, and the interplay between enthalpy and entropy govern the direction and extent of reactions—knowledge that is essential for exploring metabolism and other complex cellular processes.

This page titled 1.3: Physical-Chemical Foundations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **1.3: Physical-Chemical Foundations** by Henry Jakubowski and Patricia Flatt is licensed CC BY 4.0.
- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

1.4: Genetic Foundations

Learning Goals (ChatGPT o3-mini)

1. **Evolution and Natural Selection**
   - Describe the theory of biological evolution by natural selection and explain how heritable genetic variation drives adaptation over generations.
   - Differentiate between genotype and phenotype, and discuss how allelic variation determines phenotypic traits.
2. **Molecular Basis of Heredity**
   - Explain how the chemical structure and mutability of DNA underpin natural selection and evolutionary change.
   - Relate the roles of DNA, RNA, and proteins in transmitting and expressing genetic information.
3. **The Genetic Code and Its Decipherment**
   - Describe the triplet nature of codons and explain how codon specificity determines the amino acid sequence in proteins.
   - Summarize key experiments (e.g., those by Nirenberg, Matthaei, and Khorana) that led to the elucidation of the genetic code.
4. **Central Dogma of Molecular Biology**
   - Outline the processes of DNA replication, transcription, and translation, and explain how they interconnect to express genetic information.
   - Compare the processes in prokaryotes and eukaryotes, with emphasis on the roles of introns and splicing in eukaryotic gene expression.
5. **Mutations and Protein Evolution**
   - Identify different types of mutations (e.g., missense, nonsense, frameshift) and discuss their potential effects on protein function and phenotype.
   - Explain the concepts of homologs (orthologs and paralogs) and analogs, and discuss how protein sequence comparisons can shed light on evolutionary relationships.
6. **Gene Expression and Regulation**
   - Discuss how gene expression is regulated at multiple levels, including transcriptional control and epigenetic modifications, and explain the impact on phenotype and evolution.
   - Evaluate examples of how environmental factors can induce heritable changes in gene expression without altering the DNA sequence.
7. **Modern Molecular Biology Techniques**
   - Explain the principles behind key DNA manipulation techniques such as restriction enzyme digestion, DNA cloning, PCR, and site-specific mutagenesis.
   - Describe the methodologies used in DNA sequencing (e.g., Sanger sequencing and real-time methods) and DNA fingerprinting, and discuss how these techniques have revolutionized biological research and forensic science.
8. **Bioinformatics and Comparative Genomics**
   - Understand how bioinformatic tools are used to compare DNA and protein sequences, infer phylogenetic relationships, and predict gene function.
   - Discuss the challenges and limitations of sequence-based gene annotation and the importance of experimental validation.

By achieving these goals, students will develop a comprehensive understanding of the molecular principles underlying evolution, gene expression, and modern genetic technologies, thereby equipping them to engage with advanced topics in biochemistry and molecular biology.

1.4.1: Introduction

The development of complex biological organisms on our planet has arisen through the evolutionary mechanism of ***natural selection***. The British naturalist Charles Darwin proposed the theory of biological evolution by natural selection in his book *'On the Origin of Species'*, published in 1859. Darwin defined ***evolution*** as "descent with modification," the idea that species change over time, give rise to new species, and share a common ancestor. The mechanism that Darwin proposed for evolution is ***natural selection***. Because resources are limited, organisms with heritable traits that favor survival and reproduction will tend to leave more offspring than their peers, increasing the frequency of those traits within a population over generations. Thus, natural selection causes populations to ***adapt***, or become increasingly well-suited, to their environments over time. Natural selection depends on the environment and requires existing heritable variation in a group.

Natural selection acts on an organism's ***phenotype*** or physical characteristics. ***Phenotype*** is determined by an organism's genetic makeup (***genotype***) and the environment in which the organism lives. When different organisms in a population possess different versions of a gene for a certain trait, each version is known as an ***allele***. It is primarily this genetic variation that underlies differences in phenotype. Only a single gene governs some traits, but the interactions of many genes influence most traits. A variation in one of the many genes contributing to a trait may have only a small effect on the phenotype; together, these genes can produce a continuum of possible phenotypic values.

For example, interactions between equine coat color genes determine a horse's color. Many colors are possible, but changes in only a few genes produce all variations. *Extension* and *agouti* are particularly well-known genes with dramatic effects. For example, differences in the *agouti gene* can help determine whether a horse is bay or black. A change in the *extension gene* can, in turn, result in a horse being chestnut-colored instead (Figure 1.30). Yet other gene variants are responsible for many other coat color possibilities, including palomino, buckskin, and cremello horses.

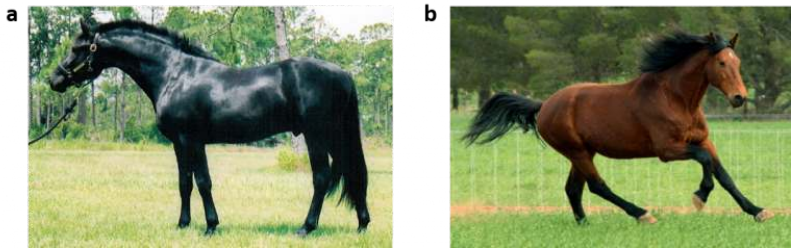


Figure 1.30 Genotype Variations as Determinants of Horse Coat Color. Horses capable of producing the black pigment eumelanin have at least one copy of the dominant *extension gene* (*E/E or E/e*). Interestingly, the *agouti gene* controls the restriction of true black pigment (eumelanin) in the coat. Horses expressing an *extension dominant gene and* recessive at the *agouti gene locus (a/a*) will be black, as shown in *(a).* Meanwhile, horses that are dominant for *extension (E/E or E/a)* and dominant for the *agouti genotype* (*A/A* or *A/a)* will never be fully black. Depending on other gene loci, they will instead show coloration patterns such as bay, *as shown in (b). Image (a)* provided by: Serendipityblue; *Image (b)* provided by: CMSporthorses

Thus, the primary molecular mechanism driving natural selection is determined by the heritability and mutability of genetic traits encoded in deoxyribonucleic acid (DNA). In Chapter 4, you will learn about the structural characteristics of DNA. In contrast, Chapter 9 focuses on the biochemical mechanisms of DNA replication, detailing the importance of DNA repair and the molecular mechanisms of genetic evolution.

1.4.2: Genetic Code

Notably, the phenotypic traits determined by an organism's genetic makeup are not directly controlled by the genetic material, DNA, but rather by the proteins encoded by the genes. In 1945, geneticist George Beadle proposed the one gene-one enzyme hypothesis, suggesting that genes are highly specific for encoding a protein sequence. However, it would take 16 more years before the biochemical nature of this process was deduced. Efforts to understand how proteins are encoded began after DNA's structure was discovered in 1953. George Gamow postulated that sets of three bases must be employed to encode the 20 standard amino acids living cells use to build proteins, allowing a maximum of $4^3 = 64$ amino acids.

The Crick, Brenner, Barnett, and Watts-Tobin experiment first demonstrated that codons have three DNA bases (Figure 1.31). Marshall Nirenberg and Heinrich J. Matthaei were the first to reveal the nature of a codon in 1961.

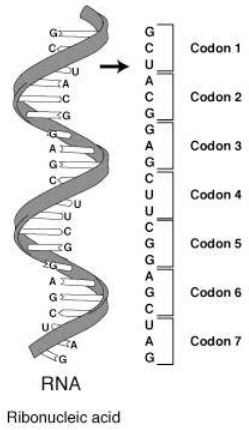


Figure 1.31 Codons Consist of Sets of Three Bases. A series of codons in part of a messenger RNA (mRNA) molecule. Each codon consists of three nucleotides, usually corresponding to a single amino acid. The nucleotides are abbreviated with the letters A, U, G, and C. This is mRNA, which uses U (uracil). DNA uses T (thymine) instead. This mRNA molecule will instruct a ribosome to synthesize a protein according to this code. Image by Sverdrup

They used a cell-free system to translate a poly-uracil RNA sequence (i.e., UUUUU…) and discovered that the polypeptide they had synthesized consisted only of the amino acid phenylalanine. Thus, they deduced that the codon UUU specified this amino acid.

This was followed by experiments in Severo Ochoa's laboratory that demonstrated that the poly-adenine RNA sequence (AAAAA…) coded for the polypeptide poly-lysine and that the poly-cytosine RNA sequence (CCCCC…) coded for the polypeptide poly-proline. Therefore, the codon AAA specified the amino acid lysine, and the codon CCC specified the amino acid proline. Using various copolymers, most of the remaining codons were determined.

Subsequent work by Har Gobind Khorana identified the rest of the genetic code. Shortly thereafter, Robert W. Holley determined the structure of transfer RNA (tRNA), the adapter molecule that facilitates the translation of RNA into protein. This work was based upon Ochoa's earlier studies, yielding the latter the Nobel Prize in Physiology or Medicine in 1959 for work on the enzymology of RNA synthesis.

Extending this work, Nirenberg and Philip Leder revealed the triplet nature of the code and deciphered its codons (Figure 1.32). In these experiments, various combinations of mRNA were passed through a filter containing ribosomes, the cellular components that translate RNA into protein. Unique triplets promoted the binding of specific tRNAs to the ribosome. Leder and Nirenberg determined the sequences of 54 out of 64 codons. Khorana, Holley, and Nirenberg received the 1968 Nobel Prize for their work.

The three stop codons were named by discoverers Richard Epstein and Charles Steinberg. "Amber" was named after their friend Harris Bernstein, whose last name means "amber" in German. The other two stop codons were named "ochre" and "opal" to keep the "color names" theme.

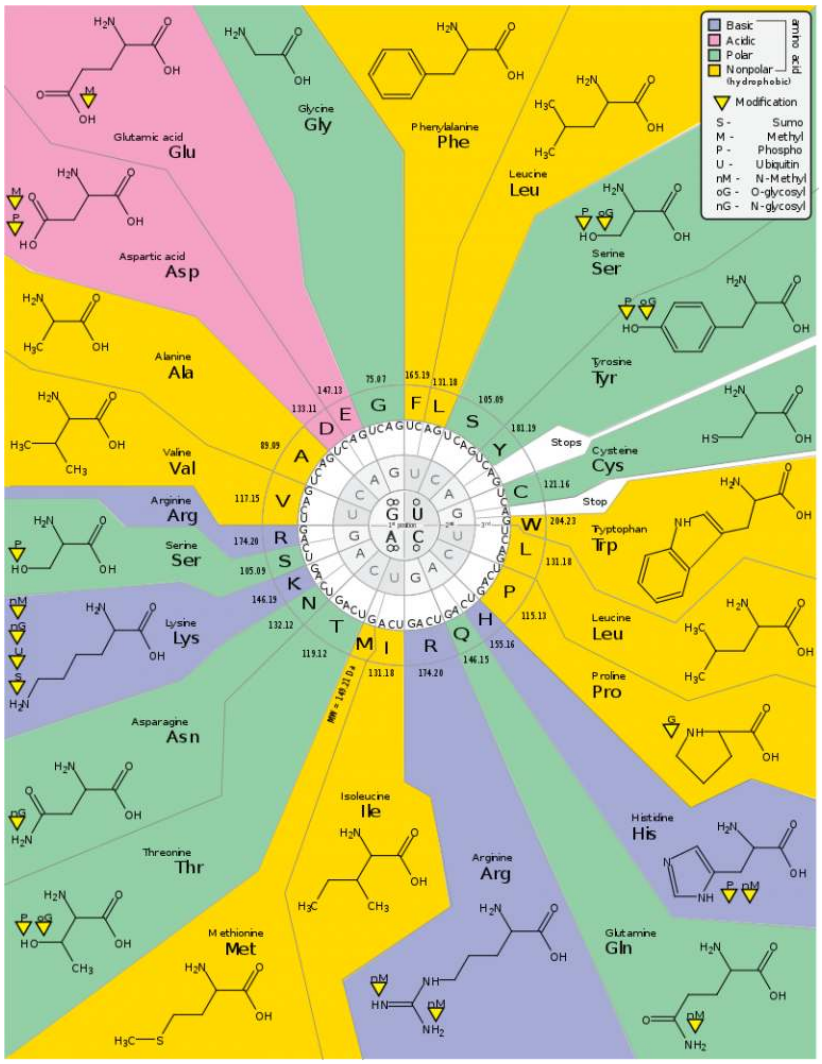


Figure 1.32 The Genetic Code. *Image edited by Seth Miller, Original file designed and produced by: Kosi Gramatikoff courtesy of Abgent*

Each gene contains a reading frame defined by the initial triplet of nucleotides from which translation starts. It sets the frame for a run of successive, non-overlapping codons, known as an ***open reading frame (ORF)***. For example, the string 5′-AAATGAACG-3′, if read from the first position, contains the codons AAA, TGA, and ACG; if read from the second position, it contains the codons AAT and GAA; and if read from the third position, it contains the codons ATG and AAC. Every sequence can, thus, be read in its 5′ → 3′ direction in three reading frames, each producing a possibly distinct amino acid sequence: in the given example, Lys (K)-Trp (W)-Thr (T), Asn (N)-Glu (E), or Met (M)-Asn (N), respectively. When DNA is double-stranded, six possible reading frames are defined: three in the forward orientation on one strand and three in reverse on the opposite strand. A start codon defines the protein-coding frame, typically the first AUG (ATG) codon in the RNA (or DNA) sequence.

Three stop codons terminate translation: UAG is *amber*, UGA is *opal* (sometimes also called *ochre*), and UAA is is *ochre*. Stop codons are also called “termination” or “nonsense” codons. They signal the release of the nascent polypeptide from the ribosome.

Mutations

During DNA replication, errors occasionally occur during the polymerization of the second strand. These errors, called mutations, can affect an organism’s phenotype, especially if they occur within a gene's protein-coding sequence. Error rates are typically one error in every 10–100 million bases due to the “proofreading” ability of DNA polymerases.

Missense mutations and ***nonsense mutations*** are types of point mutations that can cause genetic diseases, such as sickle-cell disease and thalassemia, respectively. Clinically important ***missense mutations*** generally change the properties of the coded

amino acid residue among basic, acidic, polar, or non-polar states, whereas ***nonsense mutations*** result in a stop codon.

Mutations that disrupt the reading frame by **indels** (insertions or deletions) of non-multiple-of-3 nucleotide bases are known as ***frameshift mutations***. These mutations typically result in a completely different translation from the original RNA and likely cause a stop codon to be read, thereby truncating the protein. These mutations may impair the protein's function and are thus rare in *in vivo* protein-coding sequences. One reason inheritance of frameshift mutations is rare is that if the protein being translated is essential for growth under the organism's selective pressures, the absence of a functional protein may cause death before the organism becomes viable. Frameshift mutations can result in severe genetic diseases, such as Tay–Sachs disease.

Although most mutations that alter protein sequences are either harmful or neutral, some are beneficial. These mutations may enable the mutant organism to withstand particular environmental stresses better than wild-type organisms or reproduce more quickly. In these cases, a mutation will become more common in a population through natural selection. Different sequence variations of the same gene or protein within a single organism, within a population, or between different species are known as ***sequence polymorphisms***. Larger-scale gene duplication events can also drive evolutionary change.

1.4.3: Similar Proteins

The evolution of proteins is studied by comparing the sequences and structures of proteins from many organisms representing distinct evolutionary clades. If the sequences/structures of two proteins are similar, indicating that they diverged from a common origin, they are called ***homologous proteins***. More specifically, homologous proteins in two distinct species are called ***orthologs***. In contrast, homologous proteins encoded by a single species's genome are referred to as ***paralogs***. Unrelated genes with separate evolutionary origins that each encode proteins with similar functions are termed ***analogs*** (Figure 1.33).

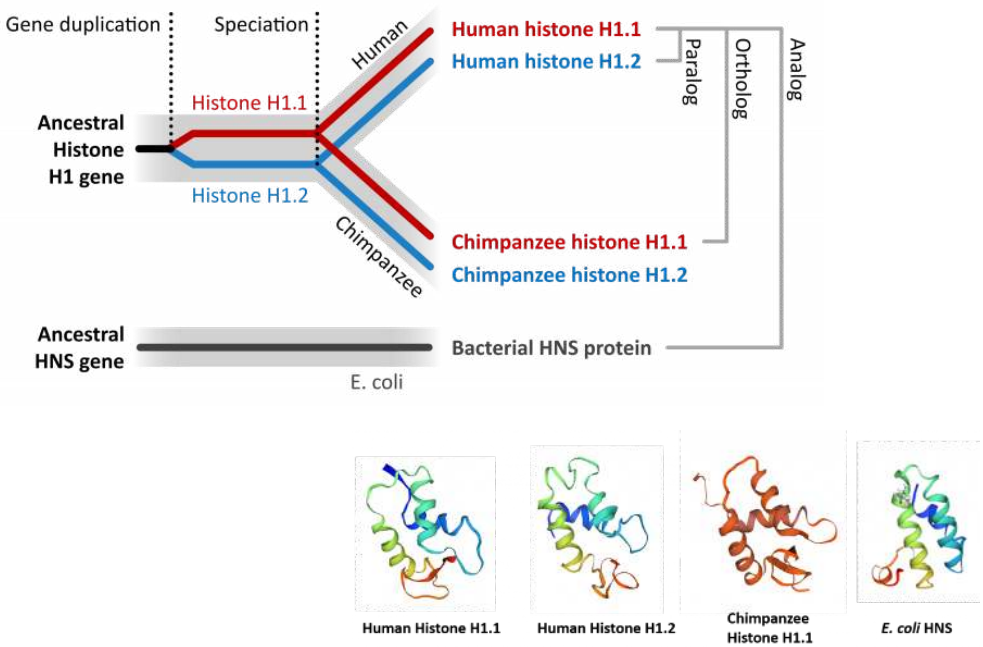


Figure 1.33 Genetic Evolution of Protein Sequences. *(Upper Panel)* An ancestral gene duplicates to produce two paralogs (Gene A and B). A speciation event produces orthologs in the two daughter species. In a separate species, an unrelated gene (Gene C) has a similar function but a distinct evolutionary origin, and thus it is an analog. *(Lower Panel)* 3-D protein models were retrieved or modeled using SWISS-MODEL: Human Histone H1.1 (Q02539), Human Histone H1.2 (P16403), E. coli HNS (P0ACF8). Histone H1.1 from the chimpanzee (*Pan troglodytes* XP_016810512.1) was modeled using Human Histone H1.1 as a template. Note that the *E. coli* HNS protein is typically modeled as a dimer. Only a single monomer is shown here. Upper Image by Thomas Shafee

DNA sequencing techniques have rapidly improved over the last 15 to 20 years, making it possible to sequence the entire genomes of organisms and, thus, predict the entire proteome of an organism based on the translation of the sequenced genome, followed by the annotation of predicted ORFs using phylogenetic comparison of similar genes/proteins from other known organisms. This has given rise to the field of ***Bioinformatics,*** which utilizes computer science, mathematics, and statistical analysis to analyze the vast

quantities of biological data generated in genome sequencing projects. The phylogenetic relationships, and hence ancestral relationships, of various genes, proteins, and, ultimately, organisms can be established through the statistical analysis of sequence alignments. Such phylogenetic trees have shown that protein sequence similarities closely reflect evolutionary relationships among organisms.

Protein evolution refers to the changes in protein shape, function, and composition that occur over time. Through quantitative analysis and experimentation, scientists have strived to understand the rate and causes of protein evolution. Scientists could estimate protein evolution rates by using the amino acid sequences of hemoglobin and cytochrome c from multiple species. They found that the rates varied among proteins. Each protein has its own rate, which remains constant across phylogenies (i.e., hemoglobin evolves at a different rate than cytochrome c, but hemoglobins from humans, mice, and other species have comparable rates of evolution). Not all regions within a protein mutate at the same rate; functionally important areas mutate more slowly, and amino acid substitutions involving similar amino acids occur more often than dissimilar substitutions. Overall, the level of polymorphisms in proteins seems to be fairly constant. Several species (including humans, fruit flies, and mice) have similar levels of protein polymorphism.

Gene duplication events, followed by mutation, can also give rise to paralogs with distinct functions within an organism. This can make it difficult to annotate genomes based solely on sequence, as homologous protein sequences may not have similar functions *in vivo*. Approximately 10-25% of annotations made on sequence homology are estimated to be incorrect and require experimental validation. For example, human pancreatic ribonuclease is a digestive enzyme that breaks down nucleic acids. The angiogenin protein is a paralog of pancreatic ribonuclease and shares high sequence homology and a similar 3D shape (Figure 1.34). However, the functions of these proteins are quite different. Angiogenin induces vascularization by activating transcriptional processes in endothelial cells. However, if the function of only one of these homologs were known, it would be easy to mistakenly hypothesize that the homologous protein would be similar in function. Thus, care must be taken when using bioinformatic tools to avoid overestimating the predictive ability of sequence alignments.

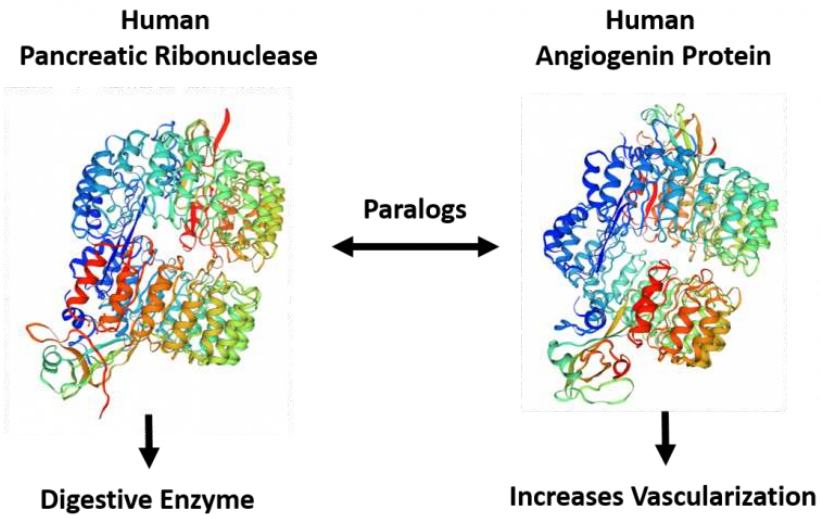


Figure 1.34 Homologous Proteins Do Not Always Have Homologous Functions. In the example above, the digestive enzyme, pancreatic ribonuclease is a paralog of the angiogenin protein and shares an ancestral origin. However, the functions of each of these proteins are quite divergent and have evolved such that they do not share homologous functions. 3-D protein models were retrieved using SWISS-MODEL: Human Pancreatic Ribonuclease (P07998) and Human Angiogenin (P03950)

The control of gene expression is crucial to all life processes, enabling the differentiation of cells to form various body structures and organs, as well as smaller, more reversible changes that allow an organism to respond to different environmental stimuli. In Chapter 12, you will explore the major biochemical mechanisms used to control gene expression within cells. This will include the discussion of a fairly new and exciting field of study known as ***epigenetics***. In addition to the heritability of traits through the passage of genetic information, it is fast becoming clear that the environmental factors that an organism is exposed to throughout its life can affect gene expression without physically altering the DNA sequence, and that these changes in expression patterns can be long-lasting and can even be inherited in the following generations. The term ***epigenetics*** means ‘on top of’ or ‘in addition to’

genetics and focuses on the heritable gene expression patterns induced by an organism's exposure or experience within its environment.

For example, in human populations, stressful events such as starvation can have lasting imprints on children who are born under these conditions. These children have higher risks of obesity and metabolic disorders as adults, including the development of type 2 diabetes. These predispositions can be passed to children born during starvation and their future offspring, indicating that environmental events can influence gene expression patterns across multiple generations. In more controlled laboratory experiments using rats, it has been demonstrated that the more a mother rat licks and nurtures her offspring, the calmer and more relaxed the offspring will be as adults. Mother rats that are less nurturing and ignore their young have offspring that will grow up displaying higher levels of anxiety. These changes are not caused by genetic differences between the offspring but rather by differences in gene expression patterns. Calm and relaxed mice can be altered to show high anxiety by exposing them to agents that alter gene expression patterns. A future chapter will cover the mechanisms that control such heritable alterations in gene expression patterns.

1.4.4: Central Dogma of Biology

DNA encodes the genetic material. It must be replicated during cell division. In transcription, the information is decoded into RNA, which is then translated into a protein sequence. Collectively, these processes are referred to as the Central Dogma of Biology. A variant occurs when RNA is decoded into DNA, a process called reverse transcription. These processes are briefly described below and further detailed in subsequent chapters.

Replication

DNA must be duplicated, a process called **replication,** before a cell divides. DNA replication allows each daughter cell to inherit a full complement of chromosomes.

Animation of Replication

Transcription and Splicing

For a given gene, only one strand of the DNA serves as the template for **transcription**. An example is shown below. In this example, the bottom (blue) strand serves as the template strand, also known as the minus (-) strand or the sense strand. This strand serves as a template for mRNA synthesis. The enzyme RNA polymerase synthesizes an mRNA in the 5' to 3' direction complementary to this template strand. The opposite DNA strand (red) is called the coding strand, the nontemplate strand, the plus (+) strand, or the antisense strand.

The easiest way to find the corresponding mRNA sequence (shown in green below) is to read the coding, nontemplate, plus (+), or antisense strand directly in the 5' to 3' direction, substituting U for T.

DNA: coding, nontemplate, plus (+), or antisense strand

```
5′ T G A C C T T C G A A C G G G A T G G A A A G G 3′
3′ A C T G G A A G C T T G C C C T A C C T T T C C 5′
```

DNA: noncoding, template, minus (-), or sense strand

```
5′ U G A C C U U C G A A C G G G A U G G A A A G G 3′
```

RNA strand

As we've learned more about the structure of DNA, RNA, and proteins, it has become clear that transcription and translation differ in eukaryotes and prokaryotes. Specifically, eukaryotes have intervening sequences of DNA (introns) within a given gene that separate coding fragments of DNA (exons). A primary transcript is made from the DNA, and the introns are sliced out, and the exons join in a contiguous stretch to form messenger RNA, which leaves the nucleus. Translation occurs in the cytoplasm. Remember, prokaryotes do not have a nucleus.

- Animation of Transcription
- Animation of mRNA Splicing

Translation

Information in an mRNA sequence is decoded to form a protein. In this process, a triplet of nucleotides (a codon) in RNA encodes a single amino acid. Translation occurs on a large RNA-protein complex called the ribosome. An intermediary transfer RNA (tRNA) molecule becomes covalently linked to a single amino acid by the enzyme tRNA-acyl synthetase. This "charged" tRNA binds via its complementary anticodon to the triplet codon in the mRNA. The ribosome/tRNA complex ratchets down the mRNA, allowing a new "charged" tRNA complex to bind at an adjacent site. The two adjacent amino acids form a peptide bond driven by ATP cleavage. This process repeats until a "stop" codon appears in the mRNA sequence. The genetic code reveals the relationship between the triplet mRNA codon and the corresponding amino acid in the growing peptide chain.

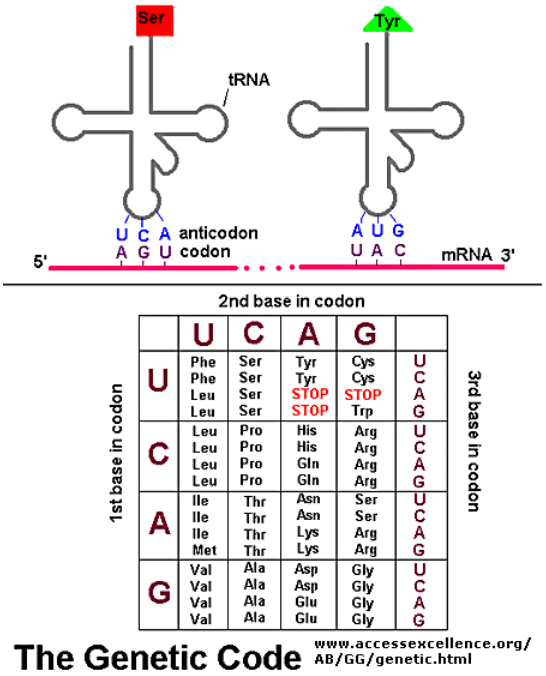


Figure 1.4.1: Codon:Anticodon interactions between mRNA and tRNA. (Copyright; author via source)

As mentioned in the Protein Chapter (amino acid section), two other amino acids occasionally appear in proteins (excluding those altered by post-translational modification). One is **selenocysteine (Sec)**, found in Archaea, eubacteria, and animals. The other is **pyrrolysine (Pyl)**, found in Archaea.

- Selenocysteine (Sec): First, the amino acid serine is attached to a tRNA specific for Sec to form Ser-tRNA^{Sec}. In bacteria, an enzyme (**selenocysteine synthase, or SelA) uses** selenophosphate to donate selenium. In Archae and Eukaryotes, an enzyme (**O-phosphoseryl-tRNA kinase or PSTK)** phosphorylated the -OH on Ser. Then the enzyme **O-phosphoseryl-tRNA:selenocysteinyl-tRNA synthase (SepSecS)** (or selenocysteine synthase, SecS) converts the O-phosphoseryl-tRNA to form the Sec. Selenocysteine-tRNA recognizes UGA, another stop codon. This usual tRNA complex would recognize only a small fraction of stop codons in the mRNA sequence group.
- **Pyrrolysine (Pyl): First,** pyrrolysine is synthesized using three enzymes (**PylB, PylC, and PylD**) using two molecules of L-lysine. Then, pyrrolsine is attached directly to its tRNA by the enzyme **pyrrolysyl-tRNA synthetase (PylRS). The tRNA^{Pyl}** has a CUA anticodon, which recognizes the UAG (amber) stop codon in mRNA.

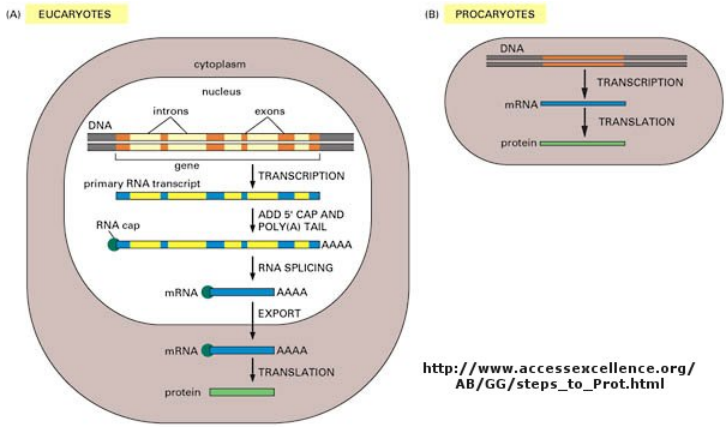


Figure: Central Dogma Differences in Eukaryotes and Prokaryotes

Animation of Translation

1.4.5: What is a gene?

The definition of a gene can differ depending on whom you ask. The word "gene" has become a cultural icon of our time. Can our genes explain what it is to be human? The definition of a gene has changed with time. Eukaryotic genes contain exons (coding regions) and introns (intervening sequences) that are transcribed to produce a primary transcript. In a post-transcriptional process, introns are spliced out by the spliceosome to produce a messenger RNA (mRNA), which is translated into a protein sequence. (See diagram above.)

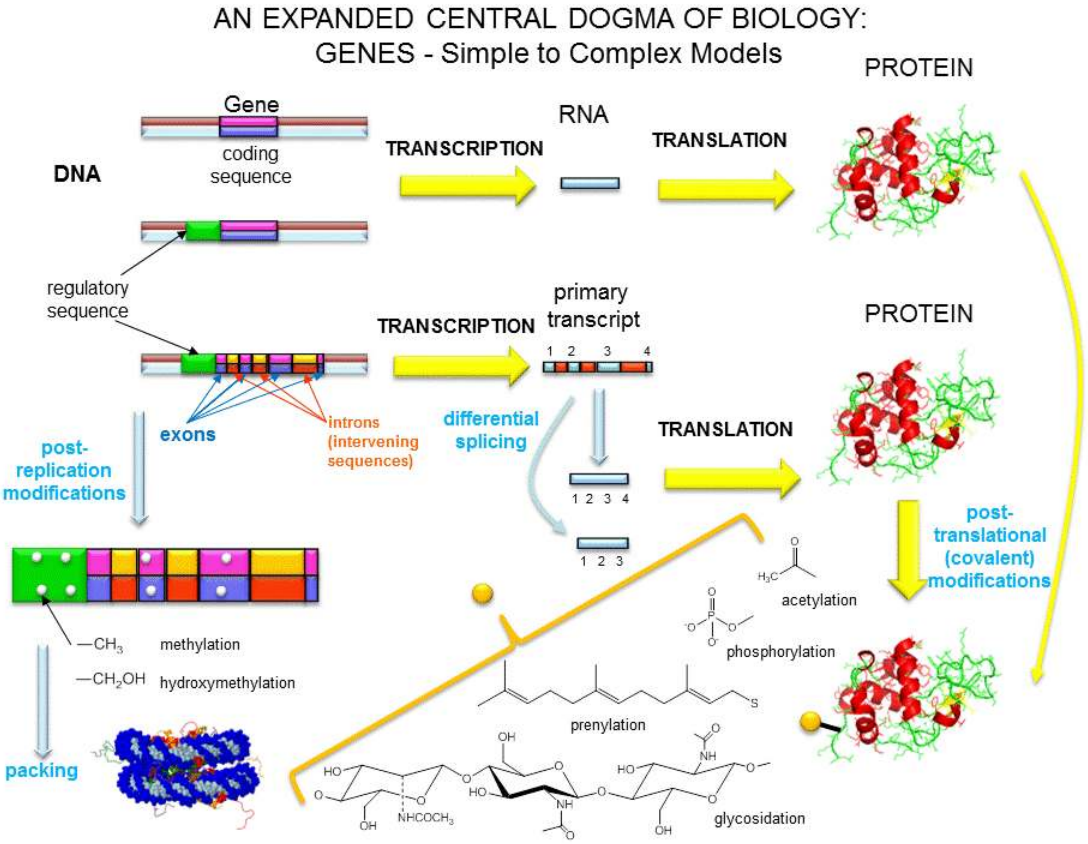


Figure 1.4.1: Figure: A view of genes and their products: Simplicity to Complexity

Over the last 100 years, as our understanding of biochemistry has increased, the definition of a gene has evolved from

- the basis of inheritable traits

- certain regions of chromosomes
- a segment of a chromosome that produces one enzyme
- a segment of a chromosome that produces one protein
- a segment of a chromosome that produces a functional product

The last definition was necessary because some gene products with functional roles (structural and catalytic) are RNA molecules. The last definition also includes regulatory regions of the chromosome involved in transcriptional control. Snyder and Gerstein have developed five criteria for gene identification, which is important as organisms' complete genomes are analyzed for genes.

1. identification of an open reading frame (ORF) - this would include a series of codons bounded by start and stop codons. This becomes progressively more challenging if the gene has many exons embedded in long introns.
2. specific DNA features within genes - these would include a bias towards certain codons found in genes or splice sites (to remove intron RNA)
3. comparing putative gene sequences for homology with known genes from different organisms, but avoiding sequences that might be conserved regulatory regions.
4. identification of RNA transcripts or expressed protein (which does not require DNA sequence analysi,s as the top three steps do) -
5. inactivating (chemically or through specific mutagenesis) a gene product (RNA or protein).

New findings make it even more complicated to define a gene, especially if the transcripts of a "gene region" are studied. Cheng et al. studied all transcripts from 10 human chromosomes and 8 cell lines. They identified a substantial number of distinct transcripts, many of which overlapped. Splicing often occurs between nonadjacent introns. Transcripts were found from both strands and were from regions containing introns and exons. Other studies found up to 5% of transcripts continued through the end of the "gene" into other genes. 63% of the entire mouse genome, comprised of only 2% exons, is transcribed.

1.4.6: The Language of DNA

In this short chapter, you will briefly learn how modern molecular biologists manipulate DNA, the blueprint for all of life. The details will be found in subsequent chapters. The four-letter alphabet (A, G, C, and T) that makes up DNA represents a language that, when transcribed and translated, leads to the myriad of proteins that make us who we are as a species and as individuals. Let's continue with the metaphor that DNA is a language. To master that language, as with any other language, we need to be able to read, write, copy, and edit that language. If you were using a word processor to find one line in a hundred-page document or one article from one book out of the Library of Congress, you would also need a way to search the large print base available. You might want to compare two copies of a file to see if they differ. Scientists can read, write, copy, edit, search, and compare the language of the genome. These abilities, acquired over the last twenty years, have revolutionized our understanding of life and given us the potential to alter life for good or evil.

DNA in human chromosomes exists as one long double-stranded molecule. It is too long to study and manipulate physically in the lab. Using a battery of enzymes, the DNA of chromosomes can be chemically cleaved into smaller fragments, which are more readily manipulable. (Similar techniques are used to sequence proteins, which require overlapping polypeptide fragments to be made.) After the fragments are created, they must be separated to be studied. DNA fragments can be separated based on a structural feature that distinguishes them. Polarity cannot be used, since all DNA fragments have negatively charged phosphate groups in the sugar-phosphate backbone. Although each fragment would have a unique sequence, it would be hard to separate all the different fragments by, for instance, attaching some molecule that binds to a unique sequence in the major groove of a given fragment to a big bead and using that bead to separate that one unique fragment. You would need a different bead for each unique fragment! The best way to separate fragments is to base separation on fragment size using electrophoresis on an agarose or polyacrylamide gel.

A carbohydrate extract called agarose is made from algae. Water is added to the extract, which is then heated. The carbohydrate extract dissolves in the water to form a viscous solution. The agarose solution is poured into a mold (similar to warm Jell-O) and allowed to solidify. A plastic comb with wide teeth was placed in the agarose when it was still liquid. When the agarose is solid, the comb can be removed, leaving little wells in its place. A solution of DNA fragments can be placed in the wells. The agarose slab with the sample is covered with a buffer solution, and electrodes are placed at each end of the slab. The negative electrode is placed near the well-end of the agarose slab, while the positive electrode is placed at the other end. If a voltage is applied across the agarose slab, the negatively charged DNA fragments will move through the agarose gel toward the positive electrode. This

migration of charged molecules in solution toward an oppositely charged electrode is called electrophoresis. Pretend you are one of the fragments.

To you, the gel looks like a tangled cobweb. You sneak through the openings in the web as you move straight forward to the positive electrode. The larger the fragment, the slower you move because getting through the tangled web is hard. Conversely, the shorter the fragment, the faster you move. Using this technique and its many variants, oligonucleotides differing by just one nucleotide can be separated. In electrophoresis of DNA fragments, a fluorescent, uncharged dye, ethidium bromide, is added to the buffer solution. This dye intercalates between DNA base pairs, imparting a fluorescent yellow-green color to the DNA when exposed to UV light on an agarose gel.

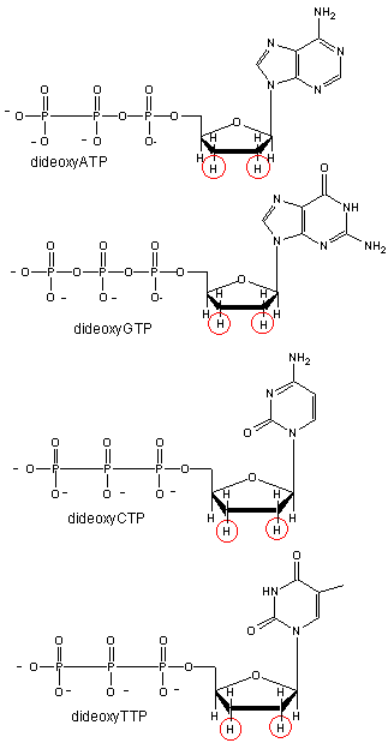
Reading DNA

We will discuss one method for reading the DNA sequence. This method, developed by Sanger, won him a second Nobel Prize. To sequence a single-stranded piece of DNA, the complementary strand is synthesized. Four different reaction mixtures are set up. Each contains all four radioactive deoxynucleotides (dATP, dCTP, dGTP, and dTTP) required for the reaction, as well as DNA polymerase.

Additionally, dideoxyATP (ddATP) is added to one of the reaction tubes. The dATP and ddATP attach randomly to the growing 3' end of the complementary strand. If ddATP is added, no further nucleotides can be added, since its 3' end has an H rather than an OH. That's why it's called dideoxy. The new chain is terminated. If dATP is added, the chain will continue growing until another A nucleotide is needed. Hence, a series of discrete DNA fragments will be generated, all terminated upon addition of ddATP. The same scenario occurs for the other three tubes, which contain dCTP and ddCTP, dTTP and ddTTP, and dGTP and ddGTP, respectively. All the fragments generated in each tube will be placed in separate lanes for electrophoresis, where they will separate by size.

Didexoynucleotides

Figure: Didexoynucleotides



PROBLEM: As shown below, you will pretend to sequence a single-stranded piece of DNA. The enzyme DNA polymerase adds the new nucleotides to the primer, GACT, in the 5' to 3' direction. You will set up four reaction tubes. Each tube contains all the dXTPs. In addition, add ddATP to tube 1, ddTTP to tube 2, ddCTP to tube 3, and ddGTP to tube 4. For each separate reaction mixture, determine all the possible sequences made by writing the possible sequences on one of the unfinished complementary

sequences below. Cut the completed sequences from the page, choose the size of the polynucleotide sequences made, and place them as they would migrate (based on size) in the appropriate lane of an imaginary gel, which you have drawn on a piece of paper. Lane 1 will contain the nucleotides produced in tube 1, and so on. Then draw lines under the positions of the cutout nucleotides to represent DNA bands in the gel. Read the sequence of the complementary DNA synthesized. Then write the sequence of the ssDNA that was to be sequenced.

5' T C A A C G A T C T G A 3' (STAND TO SEQUENCE)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

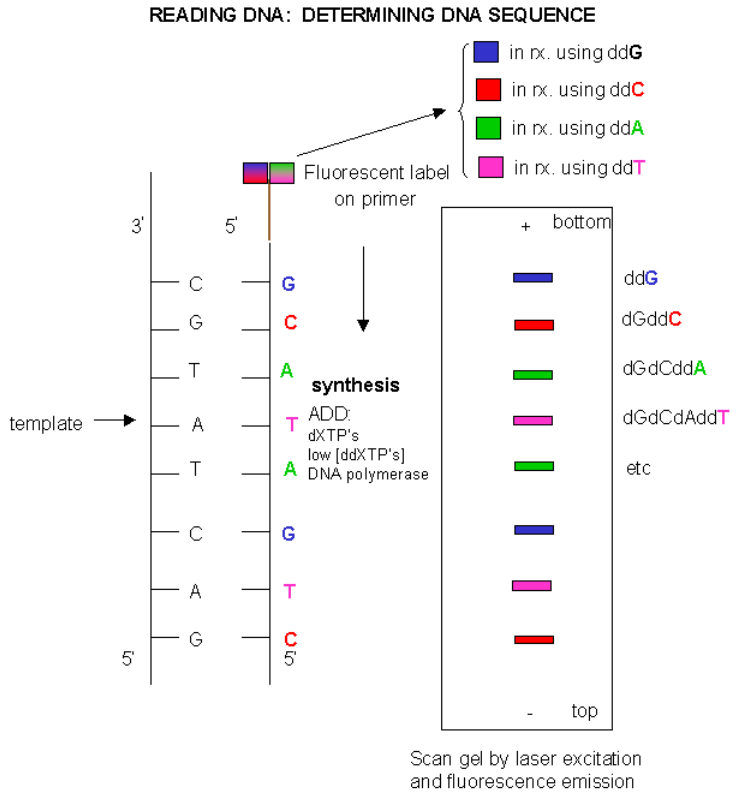
3' G A C T 5' (primer)

3' G A C T 5' (primer)

3' G A C T 5' (primer)

Since the DNA fragments have no detectable color, they can not be directly visualized in the gel. Alternative methods are used. In the one described above, radiolabeled ddXTPs were used. Once the sequencing gel is run, it can be dried, and the bands can be visualized by radioautography (also called autoradiography). An X-ray film is placed over the dried gel in a dark environment. The radiolabeled bands will emit radiation, exposing the X-ray film directly above them. The film can be developed to detect the bands. The primer can be labeled with a fluorescent dye using a newer technique. If a different dye is used for each reaction mixture, all the mixtures can be run in a single gel lane. (Only one reaction mix containing all the ddXTPs together is performed.) The gel can then be scanned by a laser, which detects fluorescence from the dyes at different wavelengths.

Figure: DNA sequencing using different fluorescent primers for each ddXTP reaction



A recent advance in sequencing enables the real-time determination of a sequence. The four deoxynucleotides are each labeled with a different fluorophore on the 5' phosphate (not the base as above). A tethered DNA polymerase elongates the DNA on a template, releasing the fluorophore into solution (i.e., the fluorophore is not incorporated into the DNA chain). The reaction occurs in a visualization chamber called a zero-mode waveguide, a cylindrical metallic chamber with a width of 70 nm and a volume of 20 zeptoliters (20 x 10^{-21} L). It sits on a glass support, through which the sample is illuminated with a laser. Given the small volume, non-incorporated fluorescently tagged deoxynucleotides diffuse in and out on the microsecond timescale. When a deoxynucleotide is incorporated into DNA, its residence time is on the millisecond timescale. This enables prolonged fluorescence detection, resulting in a high signal-to-noise ratio. Newer technologies that sequence DNA by moving it through pores in membranes could bring sequencing down to $1000/genome or less.

- Animation of Sanger Sequencing
- Nanopore sequencing

Writing DNA

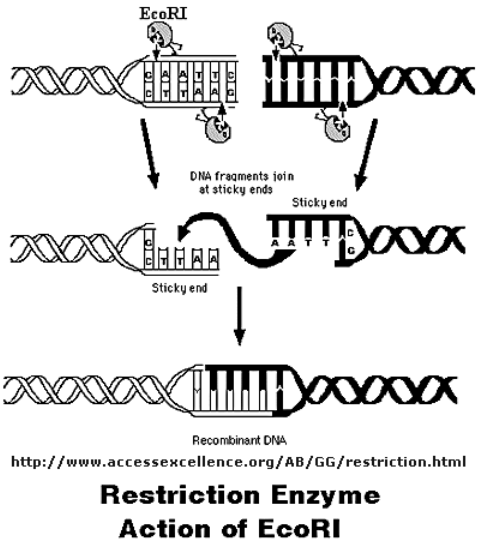
Oligonucleotides can be synthesized on a solid bead. Adding one nucleotide at a time allows control over the oligonucleotide's sequence and length.

Copying DNA

Several methods exist for copying a DNA sequence millions of times. Most methods use plasmids (found in bacteria) and viruses (which can infect any cell). The DNA of the plasmid or virus is engineered to contain a copy of a specific DNA sequence of interest. The plasmid or virus is then reintroduced into the cell, where it undergoes amplification.

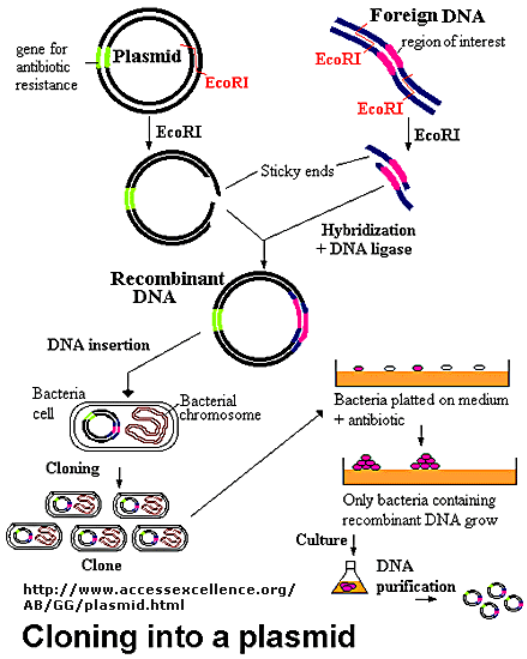
Initially, DNA containing a gene or regulatory sequence of interest is cut at specific places with an enzyme called a restriction endonuclease, or restriction enzyme for short. The enzyme doesn't cleave DNA anywhere but at "restricted" places in the sequence, much as an endoprotease cleaves a protein after a given amino acid within a protein chain. The restriction endonuclease must cleave both strands of double-stranded DNA (dsDNA). It can cut the strands cleanly to leave blunt ends or, in a staggered fashion, to leave small ssDNA tails. Multiple such sites are scattered randomly throughout the genome. The gene of interest must be flanked on either side by such a sequence. The same enzyme cleaves the DNA of plasmids or viruses.

Figure: Cleaving DNA with the Restriction Enzyme EcoR1



The foreign DNA fragment can then be added to the plasmid or viral DNA to form a recombinant DNA molecule. This DNA cloning technique is the basis for the entire field of recombinant DNA technology.

Figure: Cloning a Restriction Fragment into a Plasmid



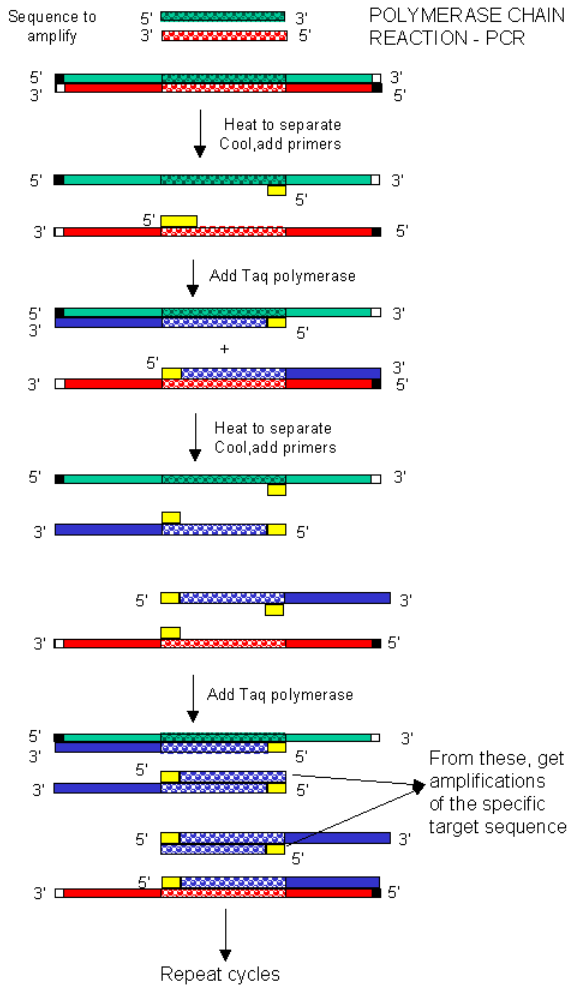
Animation of Gene Splicing

The plasmid can be introduced into bacteria, which then take it up through a process called **transformation**. The plasmid can be replicated in the bacteria, which will copy the DNA fragment of interest. Typically, the plasmid carries a gene that confers antibiotic resistance in the bacteria. Only bacteria that carry the plasmid (presumably the insert) will be able to grow. To isolate the desired fragment, the plasmids are isolated from bacteria and cleaved with the same restriction enzyme, after which the fragment can be purified. Additionally, the bacteria can be induced to express the protein encoded by the foreign gene. In lab 4, we will transform bacteria with a plasmid containing the gene for human adipocyte acid phosphatase beta, HAAP-B, and induce gene expression.

A similar method can be used to copy DNA, in which the foreign fragment is recombined with the DNA of a bacteriophage, a virus that infects bacteria, such as E. coli. The recombinant DNA can then be packaged into viruses, as shown below. When the virus infects the bacteria, it instructs the cells to make millions of new viruses, copying the foreign fragment of interest.

Sometimes, "cloning" or copying a fragment of DNA is not what an investigator wants. For instance, if the genomic DNA comes from a human cell, the gene will contain introns. If you put this DNA into a plasmid or bacteriophage, the introns go with it. Bacteria can replicate this DNA, but often, one wants not just to copy (amplify) the DNA but also to transcribe it into RNA and then translate it into protein. Bacteria, however, cannot splice out the intron, so mature mRNA cannot be produced. If one could clone into the bacteria's DNA without the introns, this problem would not exist. One such possible method involves starting with the actual mRNA for a protein of interest. In this technique, a dsDNA copy is made from an ss-mRNA molecule. Such dsDNA is called cDNA, for complementary or copy DNA. This can then be cloned into a plasmid or bacteriophage vector and amplified as described above.

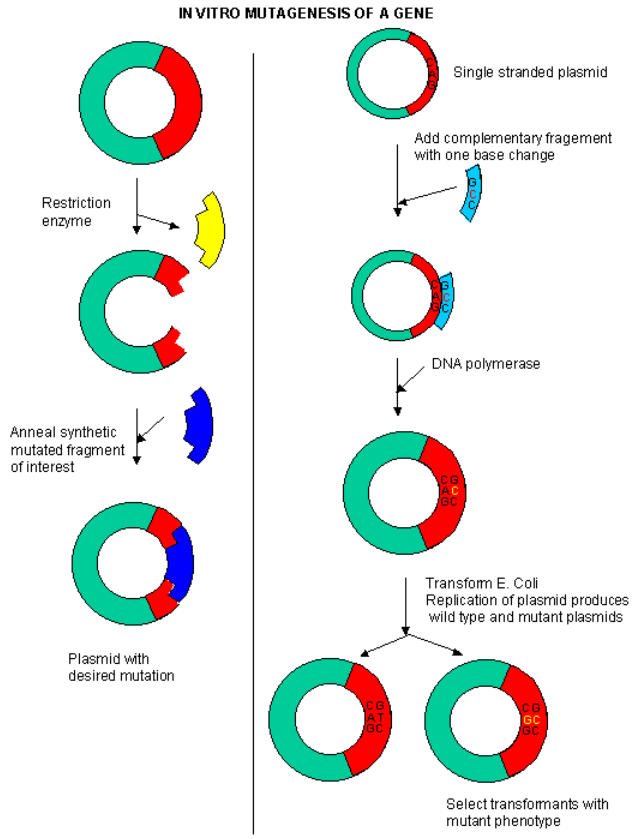
In the mid-'80s, a new method was developed to copy (amplify) DNA in a test tube. It doesn't require a plasmid or a virus. It requires only a DNA fragment and some primers (small oligonucleotides complementary to sections of DNA on each strand that straddle the section to be amplified). Add dATP, dCTP, dGTP, dTTP, and a heat-stable DNA polymerase from the organism Thermus aquaticus (which inhabits hot springs), and you're ready to proceed. The mixture is first heated to a temperature that separates the dsDNA strands. The temperature is lowered, allowing a large stoichiometric excess of the primers to anneal to the ssDNA. Heat-stable Taq polymerase (from Thermus aquaticus) extends primers. The temperature is raised again, allowing the dsDNA strand to separate. Upon cooling, the primers anneal to both the original and the newly synthesized DNA from the previous cycle, and DNA synthesis resumes. This cycle is repeated, as shown in the diagram. This chain reaction is called the polymerase chain reaction (PCR). The target DNA synthesized is amplified a million times in 20 cycles or a billion times in 30 cycles, which can be done in a few hours.



- Animation of PCR

Editing DNA

We will spend considerable time discussing how specific amino acids can be covalently modified to identify the presence of a particular amino acid or to modify the protein's activity. It is routine to use recombinant DNA technology to alter one or more nucleotides, thereby changing the amino acid or adding or deleting one or more amino acids. This technique, known as site-specific mutagenesis, is widely employed by protein chemists to assess the significance of a specific amino acid in a protein's folding, structure, and function. The techniques are described in the diagram below;



Searching DNA

Where on a chromosome is the gene that codes for a given protein? One way to find the gene is to synthesize a small oligonucleotide "probe" complementary to part of the gene's actual DNA sequence (determined from previous experiments). Attach a fluorescent molecule to the DNA probe. Then, use a cell preparation in which the chromosomes can be seen under the microscope. The base is added, which unwinds the double-stranded DNA helix. A fluorescent probe is added that binds to the chromosome at the site of the gene complementary to the DNA. Hybridization is the process whereby a single-stranded nucleotide sequence (the target) binds through H-bonds to another complementary nucleotide sequence (the probe).

What if you don't know the nucleotide sequence of the gene but know the protein's amino acid sequence, as in the example shown below? From the genetic code table, you could predict the sequence of all possible RNA molecules complementary to the DNA in the gene. Since some amino acids have more than one codon, many possible sequences of DNA could code for the protein fragment. The link below lists all corresponding mRNA sequences that could encode a short amino acid sequence. The 20-mer sequences of minimal nucleotide degeneracy should be used as genomic probes.

Comparing DNA

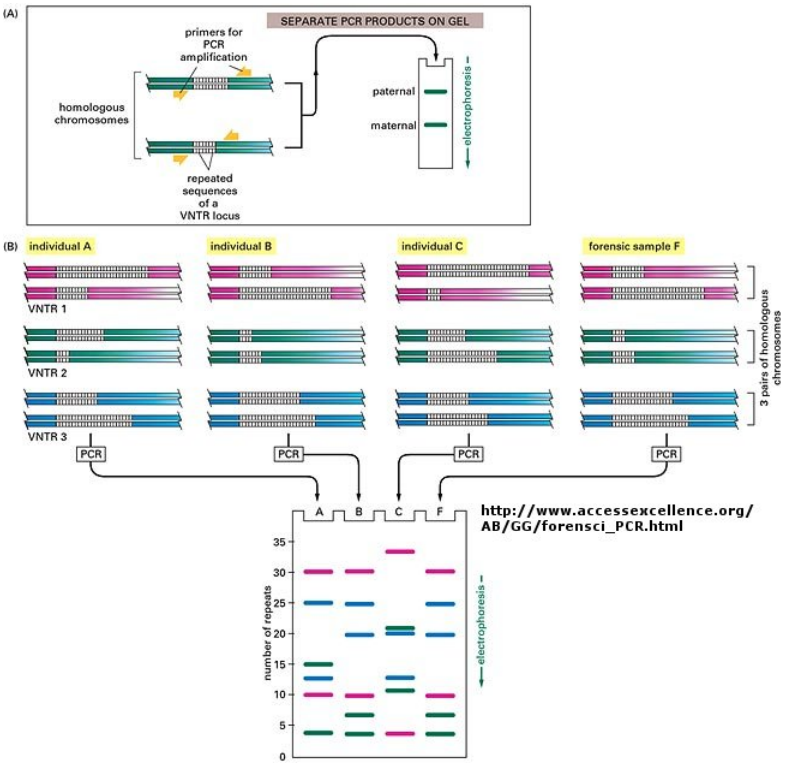
The DNA sequence of each individual must be unique to that individual, differing from every other individual in the world (except identical twins). The difference must be less than the differences between a human and a chimp, which are 98.5 % identical. Let us say that each of us has DNA sequences that are 99.9% identical to those of some "normal humans". Given that we have about 4 billion base pairs of DNA, we are all different by about 0.001 x 4,000,000,000, or about 4 million base pairs. This means that, on average, we have one nucleotide difference for each 1000 DNA base pairs. Some of these are encoded in genes, but most are likely located between DNA, and many are clustered in areas of highly repetitive DNA at the ends of chromosomes (called telomeres) and in the middle (called centromeres).

Remember that restriction enzyme sites are also interspersed randomly along the DNA. If differences in DNA among individuals occur within sequences cleaved by restriction enzymes, then in some individuals a particular enzyme may not cleave at the usual site but instead at a more distal site. Hence, the sizes of the restriction enzyme fragments should differ among individuals. When

cut with a battery of restriction enzymes, each person's DNA should yield a unique set of DNA fragments, each with a size specific to that individual. Each person's DNA has a unique **Restriction Fragment Length Polymorphism (RFLP).** How could you detect such a polymorphism?

You know how to cut sample DNA with restriction enzymes and separate the fragments on an agarose gel. However, an additional step is required, as thousands of fragments could appear on the gel, resulting in a single large continuous smear. If, however, each fragment could be reacted with a set of small, radioactive DNA probes that are complementary to certain highly polymorphic sections of DNA (like telomeric DNA) and then visualized, only a few sets of discrete bands would be observed in the agarose gel. These discrete bands would differ from the DNA bands seen in another individual's gene, which would be treated similarly. This technique, known as Southern Blotting, works as shown below. DNA fragments are separated by electrophoresis in an agarose gel. The dsDNA fragments are unwound by heating, then a piece of nitrocellulose filter paper is placed on top of the gel, and the DNA from the gel transfers to the filter paper. Then, a small radioactive oligonucleotide probe complementary to a polymorphic site on the DNA is added to the paper. It binds only to the fragment containing DNA complementary to the probe. The filter paper is dried, and a piece of X-ray film is placed over the sheet. A set of radioactive fragments (not complementary to the probe) is also run. They serve as markers to ensure that gel electrophoresis and transfer to filter paper were performed correctly.

When this technique is applied in forensic cases or paternity cases, it is referred to as DNA fingerprinting. With current techniques, investigators can unequivocally state that the odds of a particular pattern not belonging to a suspect are 1 in 1,000,000. The X-ray film shown below is a copy of real forensic evidence obtained from a rape case. The Southern blot results from suspect 1, suspect 2, the victim, and the forensic evidence are shown. Analyze the data.



1.4.7: Summary

This chapter weaves together fundamental concepts in molecular evolution and modern genetic technologies to provide a comprehensive view of how biological information is stored, transmitted, and modified. It begins by revisiting Darwin's theory of evolution by natural selection, emphasizing that genetic variation—arising from mutations in DNA—is the engine driving the adaptation and diversification of species. The relationship between genotype and phenotype is explored through examples such as variations in equine coat color, illustrating how specific alleles combine to produce a continuum of observable traits.

The chapter then delves into the molecular basis of heredity, detailing the discovery and decipherment of the genetic code. It highlights key experiments that established the triplet nature of codons and their role in specifying the 20 standard amino acids,

thereby forming the basis for protein synthesis. This discovery not only solidified the concept of the Central Dogma—where DNA is replicated, transcribed into RNA, and translated into protein—but also underscored the importance of non-protein-coding sequences and regulatory regions in complex organisms.

Moving forward, the text examines the mechanisms and consequences of mutations, including point mutations (missense and nonsense) and frameshift mutations. It discusses how these genetic alterations can impact protein function and lead to disease. The chapter also distinguishes between homologous proteins (orthologs and paralogs) and analogs, providing insight into protein evolution and the use of sequence comparisons to reconstruct phylogenetic relationships.

A significant portion of the chapter is devoted to modern molecular biology techniques that have revolutionized our ability to read, copy, edit, and compare DNA. Detailed explanations are provided for methods such as restriction enzyme digestion, DNA cloning, and polymerase chain reaction (PCR), as well as advanced DNA sequencing techniques—including Sanger sequencing and real-time nanopore sequencing. Additionally, the chapter covers DNA fingerprinting and Southern blotting as powerful tools for genetic analysis and forensic applications.

In summary, this chapter integrates evolutionary theory with molecular mechanisms and technological advances, equipping students with a deep understanding of how genetic information is encoded, maintained, and manipulated—knowledge that is crucial for exploring the frontiers of biochemistry and molecular biology.

This page titled 1.4: Genetic Foundations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **1.4: Genetic Foundations** by Henry Jakubowski and Patricia Flatt is licensed CC BY 4.0.
- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

1.5: Essential elements for life

The diverse environments and ecosystems on Earth have a wide range of conditions in terms of temperature, available nutrients, acidity, salinity, and energy sources. Prokaryotes are very well equipped to make their living out of a vast array of nutrients and conditions. To live, prokaryotes need a source of energy, a source of carbon, and some additional nutrients.

Macronutrients

Cells are essentially a well-organized assemblage of macromolecules and water. Recall that macromolecules are produced by the polymerization of smaller units called monomers. For cells to build all of the molecules required to sustain life, they need certain substances, collectively called **nutrients**. When prokaryotes grow in nature, they obtain their nutrients from the environment. Nutrients that are required in large amounts are called macronutrients, whereas those required in smaller or trace amounts are called micronutrients. Just a handful of elements are considered macronutrients—carbon, hydrogen, oxygen, nitrogen, phosphorus, and sulfur. (A mnemonic for remembering these elements is the acronym *CHONPS*.)

Why are these macronutrients needed in large amounts? They are the components of organic compounds in cells, including water. Carbon is the major element in all macromolecules: carbohydrates, proteins, nucleic acids, lipids, and many other compounds. Carbon accounts for about 50 percent of the composition of the cell. Nitrogen represents 12 percent of the total dry weight of a typical cell and is a component of proteins, nucleic acids, and other cell constituents. Most of the nitrogen available in nature is either atmospheric nitrogen (N_2) or another inorganic form. Diatomic (N_2) nitrogen, however, can be converted into an organic form only by certain organisms, called nitrogen-fixing organisms. Both hydrogen and oxygen are part of many organic compounds and of water. Phosphorus is required by all organisms for the synthesis of nucleotides and phospholipids. Sulfur is part of the structure of some amino acids such as cysteine and methionine, and is also present in several vitamins and coenzymes. Other important macronutrients are potassium (K), magnesium (Mg), calcium (Ca), and sodium (Na). Although these elements are required in smaller amounts, they are very important for the structure and function of the prokaryotic cell.

Micronutrients

In addition to these macronutrients, prokaryotes require various metallic elements in small amounts. These are referred to as micronutrients or trace elements. For example, iron is necessary for the function of the cytochromes involved in electron-transport reactions. Some prokaryotes require other elements—such as boron (B), chromium (Cr), and manganese (Mn)—primarily as enzyme cofactors.

Contributors and Attributions

CC licensed content, Shared previously

- Biology. **Provided by**: OpenStax CNX. **Located at**: http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@10.8. **License**: *CC BY: Attribution*. **License Terms**: Download for free at http://cnx.org/contents/185cbf87-c72...f21b5eabd@10.8

1.5: Essential elements for life is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.5: Essential elements for life** has no license indicated.
- **6.10: Needs of Prokaryotes** has no license indicated.

1.6: The Building Blocks of Molecules

At its most fundamental level, life is made up of matter. Matter occupies space and has mass. All matter is composed of elements, substances that cannot be broken down or transformed chemically into other substances. Each element is made of atoms, each with a constant number of protons and unique properties. A total of 118 elements have been defined; however, only 92 occur naturally, and fewer than 30 are found in living cells. 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.

Atoms

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 subatomic particles, it would no longer have the properties of hydrogen.

At the most basic level, all organisms are made of a combination of elements. They contain atoms that combine together to form molecules. In multicellular organisms, such as animals, molecules can interact to form cells that combine to form tissues, which make up organs. These combinations continue until entire multicellular organisms are formed.

All atoms contain protons, electrons, and neutrons (Figure 1.6.1). The only exception is hydrogen (H), which is 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.

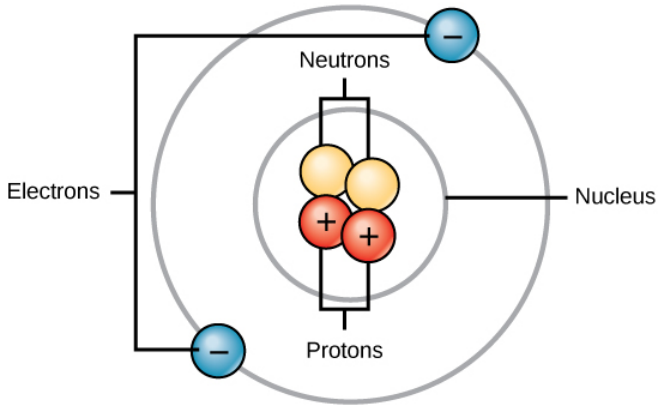


***Figure** 1.6.1**:** Atoms are made up of protons and neutrons located within the nucleus, and electrons surrounding the nucleus.*

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.

Because protons and neutrons each have a mass of 1, the mass of an atom is equal to the number of protons and neutrons of that atom. The number of electrons does not factor into the overall mass, because their mass is so small.

As stated earlier, each element has its own unique properties. Each 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, or atomic mass, 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.

These numbers provide information about the elements and how they will react when combined. Different elements have different melting and boiling points, and are in different states (liquid, solid, or gas) at room temperature. They also combine in different

ways. Some form specific types of bonds, whereas others do not. How they combine is based on the number of electrons present. Because of these characteristics, the elements are arranged into the periodic table of elements, a chart of the elements that includes the atomic number and relative atomic mass of each element. The periodic table also provides key information about the properties of elements (*Figure* **1.6.2**)—often indicated by color-coding. The arrangement of the table also shows how the electrons in each element are organized and provides important details about how atoms will react with each other to form molecules.

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 will lose protons, other subatomic particles, or energy to form more stable elements. These are called radioactive isotopes or radioisotopes.

ART CONNECTION

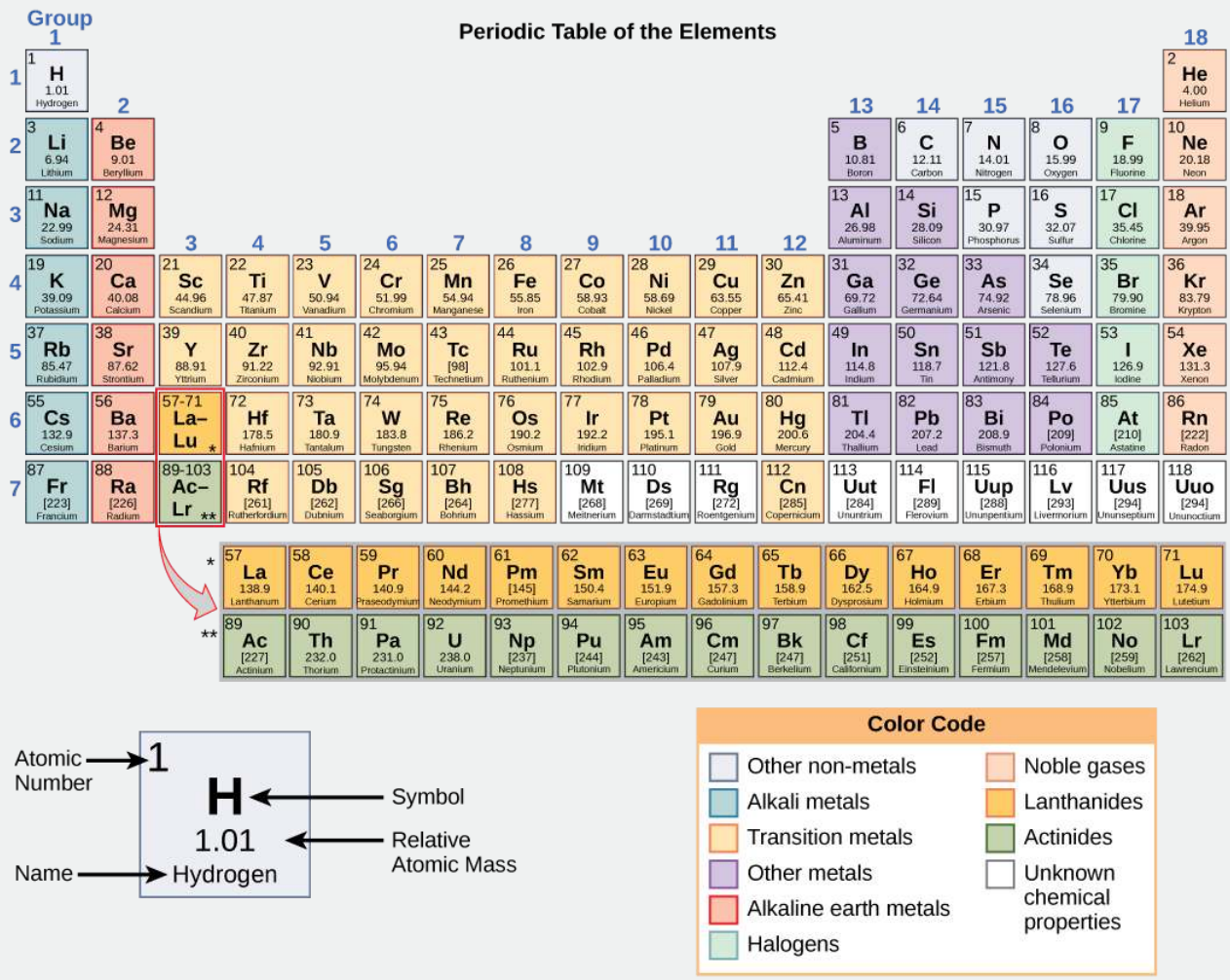


Figure **1.6.2**: Arranged in columns and rows based on the characteristics of the elements, the periodic table provides key information about the elements and how they might interact with each other to form molecules. Most periodic tables provide a key or legend to the information they contain.

Chemical Bonds

How elements interact with one another depends on how their electrons are arranged and how many openings for electrons exist at the outermost region where electrons are present in an atom. Electrons exist at energy levels that form shells around the nucleus.

The closest shell can hold up to two electrons. The closest shell to the nucleus is always filled first, before any other shell can be filled. Hydrogen has one electron; therefore, it has only one spot occupied within the lowest shell. Helium has two electrons; therefore, it can completely fill the lowest shell with its two electrons. If you look at the periodic table, you will see that hydrogen and helium are the only two elements in the first row. This is because they only have electrons in their first shell. Hydrogen and helium are the only two elements that have the lowest shell and no other shells.

The second and third energy levels can hold up to eight electrons. The eight electrons are arranged in four pairs and one position in each pair is filled with an electron before any pairs are completed.

Looking at the periodic table again (Figure **1.6.2**), you will notice that there are seven rows. These rows correspond to the number of shells that the elements within that row have. The elements within a particular row have increasing numbers of electrons as the columns proceed from left to right. Although each element has the same number of shells, not all of the shells are completely filled with electrons. If you look at the second row of the periodic table, you will find lithium (Li), beryllium (Be), boron (B), carbon (C), nitrogen (N), oxygen (O), fluorine (F), and neon (Ne). These all have electrons that occupy only the first and second shells. Lithium has only one electron in its outermost shell, beryllium has two electrons, boron has three, and so on, until the entire shell is filled with eight electrons, as is the case with neon.

Not all elements have enough electrons to fill their outermost shells, but an atom is at its most stable when all of the electron positions in the outermost shell are filled. Because of these vacancies in the outermost shells, we see the formation of chemical bonds, or interactions between two or more of the same or different elements that result in the formation of molecules. To achieve greater stability, atoms will tend to completely fill their outer shells and will bond with other elements to accomplish this goal by sharing electrons, accepting electrons from another atom, or donating electrons to another atom. Because the outermost shells of the elements with low atomic numbers (up to calcium, with atomic number 20) can hold eight electrons, this is referred to as the octet rule. An element can donate, accept, or share electrons with other elements to fill its outer shell and satisfy the octet rule.

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. Positive ions are formed by losing electrons and are called cations. Negative ions are formed by gaining electrons and are called anions.

For example, sodium only has one electron in its outermost shell. It takes less energy for sodium to donate that one electron than it does to accept seven more electrons to fill the outer shell. If sodium loses an electron, it now has 11 protons and only 10 electrons, leaving it with an overall charge of +1. It is now called a sodium ion.

The chlorine atom has seven electrons in its outer shell. Again, it is more energy-efficient for chlorine to gain one electron than to lose seven. Therefore, it tends to gain an electron to create an ion with 17 protons and 18 electrons, giving it a net negative (–1) charge. It is now called a chloride ion. This movement of electrons from one element to another is referred to as electron transfer. As Figure **1.6.4** illustrates, a sodium atom (Na) only has one electron in its outermost shell, whereas a chlorine atom (Cl) has seven electrons in its outermost shell. A sodium atom will donate its one electron to empty its shell, and a chlorine atom will accept that electron to fill its shell, becoming chloride. Both ions now satisfy the octet rule and have complete outermost shells. Because the number of electrons is no longer equal to the number of protons, each is now an ion and has a +1 (sodium) or –1 (chloride) charge.

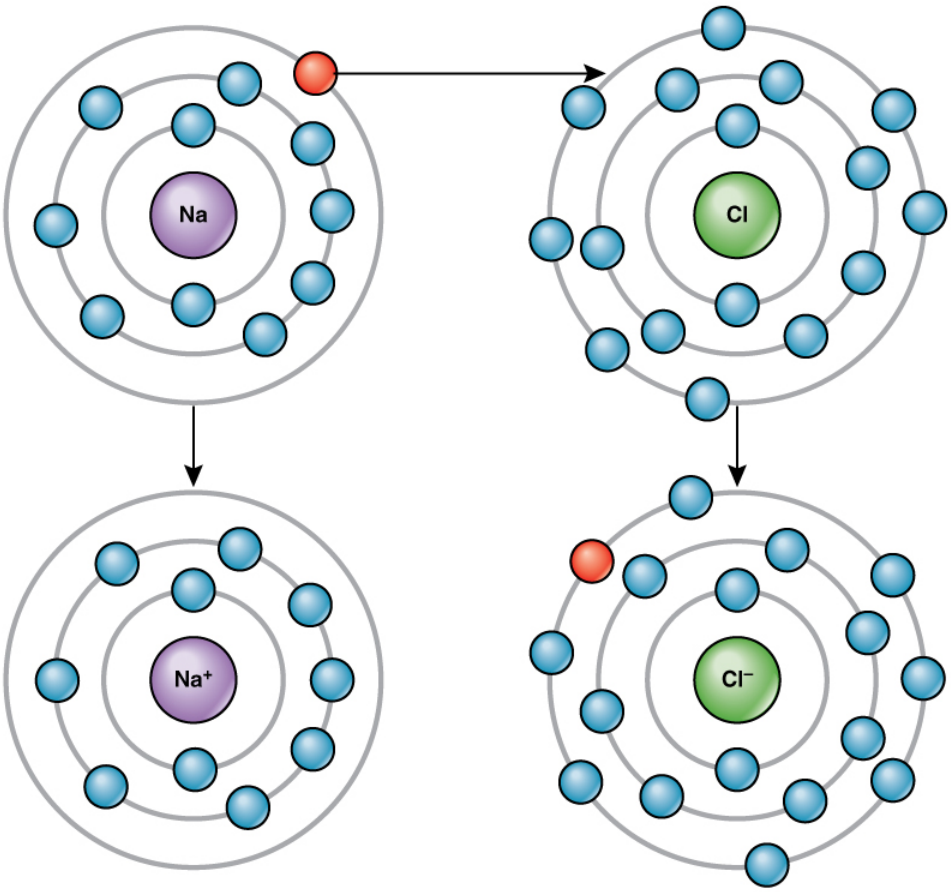


Figure 1.6.4: Elements tend to fill their outermost shells with electrons. To do this, they can either donate or accept electrons from other elements.

Ionic Bonds

There are four types of bonds or interactions: ionic, covalent, hydrogen bonds, and van der Waals interactions. Ionic and covalent bonds are strong interactions that require a larger energy input to break apart. When an element donates an electron from its outer shell, as in the sodium atom example above, a positive ion is formed. The element accepting the electron is now negatively charged. Because positive and negative charges attract, these ions stay together and form an ionic bond, or a bond between ions. The elements bond together with the electron from one element staying predominantly with the other element. When Na^+ and Cl^- ions combine to produce NaCl, an electron from a sodium atom stays with the other seven from the chlorine atom, and the sodium and chloride ions attract each other in a lattice of ions with a net zero charge.

Covalent Bonds

Another type of strong chemical bond between two or more atoms is a covalent bond. These bonds form when an electron is shared between two elements and are the strongest and most common form of chemical bond in living organisms. Covalent bonds form between the elements that make up the biological molecules in our cells. Unlike ionic bonds, covalent bonds do not dissociate in water.

The hydrogen and oxygen atoms that combine to form water molecules are bound together by covalent bonds. The electron from the hydrogen atom divides its time between the outer shell of the hydrogen atom and the incomplete outer shell of the oxygen atom. To completely fill the outer shell of an oxygen atom, two electrons from two hydrogen atoms are needed, hence the subscript “2” in H_2O. The electrons are shared between the atoms, dividing their time between them to “fill” the outer shell of each. This sharing is a lower energy state for all of the atoms involved than if they existed without their outer shells filled.

There are two types of covalent bonds: polar and nonpolar. Nonpolar covalent bonds form between two atoms of the same element or between different elements that share the electrons equally. For example, an oxygen atom can bond with another oxygen atom to fill their outer shells. This association is nonpolar because the electrons will be equally distributed between each oxygen atom. Two

covalent bonds form between the two oxygen atoms because oxygen requires two shared electrons to fill its outermost shell. Nitrogen atoms will form three covalent bonds (also called triple covalent) between two atoms of nitrogen because each nitrogen atom needs three electrons to fill its outermost shell. Another example of a nonpolar covalent bond is found in the methane (CH_4) molecule. The carbon atom has four electrons in its outermost shell and needs four more to fill it. It gets these four from four hydrogen atoms, each atom providing one. These elements all share the electrons equally, creating four nonpolar covalent bonds (Figure 1.6.5).

In a polar covalent bond, the electrons shared by the atoms spend more time closer to one nucleus than to the other nucleus. Because of the unequal distribution of electrons between the different nuclei, a slightly positive (δ+) or slightly negative (δ–) charge develops. The covalent bonds between hydrogen and oxygen atoms in water are polar covalent bonds. The shared electrons spend more time near the oxygen nucleus, giving it a small negative charge, than they spend near the hydrogen nuclei, giving these molecules a small positive charge.

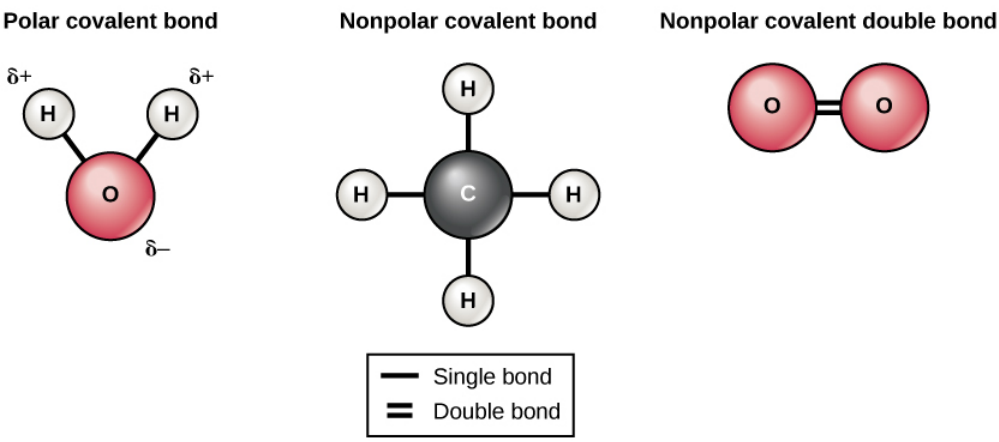


Figure 1.6.5: *The water molecule (left) depicts a polar bond with a slightly positive charge on the hydrogen atoms and a slightly negative charge on the oxygen. Examples of nonpolar bonds include methane (middle) and oxygen (right).*

Hydrogen Bonds

Ionic and covalent bonds are strong bonds that require considerable energy to break. However, not all bonds between elements are ionic or covalent bonds. Weaker bonds can also form. These are attractions that occur between positive and negative charges that do not require much energy to break. Two weak bonds that occur frequently are hydrogen bonds and van der Waals interactions. These bonds give rise to the unique properties of water and the unique structures of DNA and proteins.

When polar covalent bonds containing a hydrogen atom form, the hydrogen atom in that bond has a slightly positive charge. This is because the shared electron is pulled more strongly toward the other element and away from the hydrogen nucleus. Because the hydrogen atom is slightly positive (δ+), it will be attracted to neighboring negative partial charges (δ–). When this happens, a weak interaction occurs between the δ+ charge of the hydrogen atom of one molecule and the δ– charge of the other molecule. This interaction is called a hydrogen bond. This type of bond is common; for example, the liquid nature of water is caused by the hydrogen bonds between water molecules (Figure 1.6.6). Hydrogen bonds give water the unique properties that sustain life. If it were not for hydrogen bonding, water would be a gas rather than a liquid at room temperature.

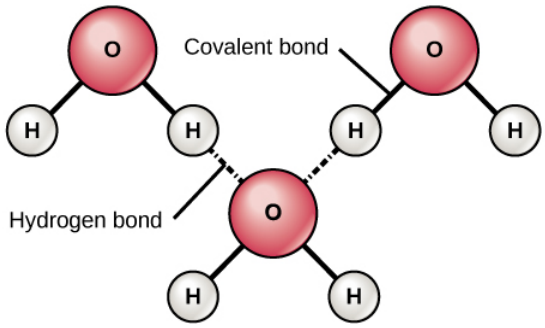


Figure 1.6.6: Hydrogen bonds form between slightly positive (δ+) and slightly negative (δ–) charges of polar covalent molecules, such as water.

Hydrogen bonds can form between different molecules and they do not always have to include a water molecule. Hydrogen atoms in polar bonds within any molecule can form bonds with other adjacent molecules. For example, hydrogen bonds hold together two long strands of DNA to give the DNA molecule its characteristic double-stranded structure. Hydrogen bonds are also responsible for some of the three-dimensional structure of proteins.

van der Waals Interactions

Like hydrogen bonds, van der Waals interactions are weak attractions or interactions between molecules. They occur between polar, covalently bound, atoms in different molecules. Some of these weak attractions are caused by temporary partial charges formed when electrons move around a nucleus. These weak interactions between molecules are important in biological systems.

Matter is anything that occupies space and has mass. It is made up of atoms of different elements. All of the 92 elements that occur naturally have unique qualities that allow them to combine in various ways to create compounds or molecules. Atoms, which consist of protons, neutrons, and electrons, are the smallest units of an element that retain all of the properties of that element. Electrons can be donated or shared between atoms to create bonds, including ionic, covalent, and hydrogen bonds, as well as van der Waals interactions.

Glossary

anion

a negative ion formed by gaining electrons

atomic number

the number of protons in an atom

cation

a positive ion formed by losing electrons

chemical bond

an interaction between two or more of the same or different elements that results in the formation of molecules

covalent bond

a type of strong bond between two or more of the same or different elements; forms when electrons are shared between elements

electron

a negatively charged particle that resides outside of the nucleus in the electron orbital; lacks functional mass and has a charge of –1

electron transfer

the movement of electrons from one element to another

element

one of 118 unique substances that cannot be broken down into smaller substances and retain the characteristic of that substance; each element has a specified number of protons and unique properties

hydrogen bond

a weak bond between partially positively charged hydrogen atoms and partially negatively charged elements or molecules

ion

an atom or compound that does not contain equal numbers of protons and electrons, and therefore has a net charge

ionic bond

a chemical bond that forms between ions of opposite charges

isotope

one or more forms of an element that have different numbers of neutrons

mass number

the number of protons plus neutrons in an atom

matter

anything that has mass and occupies space

neutron

a particle with no charge that resides in the nucleus of an atom; has a mass of 1

nonpolar covalent bond

a type of covalent bond that forms between atoms when electrons are shared equally between atoms, resulting in no regions with partial charges as in polar covalent bonds

nucleus

(chemistry) the dense center of an atom made up of protons and (except in the case of a hydrogen atom) neutrons

octet rule

states that the outermost shell of an element with a low atomic number can hold eight electrons

periodic table of elements

an organizational chart of elements, indicating the atomic number and mass number of each element; also provides key information about the properties of elements

polar covalent bond

a type of covalent bond in which electrons are pulled toward one atom and away from another, resulting in slightly positive and slightly negative charged regions of the molecule

proton

a positively charged particle that resides in the nucleus of an atom; has a mass of 1 and a charge of +1

radioactive isotope

an isotope that spontaneously emits particles or energy to form a more stable element

van der Waals interaction

a weak attraction or interaction between molecules caused by slightly positively charged or slightly negatively charged atoms

Contributors and Attributions

- Samantha Fowler (Clayton State University), Rebecca Roush (Sandhills Community College), James Wise (Hampton University). Original content by OpenStax (CC BY 4.0; Access for free at https://cnx.org/contents/b3c1e1d2-83...4-e119a8aafbdd).

This page titled 1.6: The Building Blocks of Molecules is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by OpenStax.

- **2.1: The Building Blocks of Molecules** by OpenStax is licensed CC BY 4.0.
- **1.6: The Building Blocks of Molecules** by OpenStax is licensed CC BY 4.0.

1.7: Types of Biological Molecules

The large molecules necessary for life that are built from smaller organic molecules are called biological macromolecules. There are four major classes of biological macromolecules (carbohydrates, lipids, proteins, and nucleic acids), and each is an important component of the cell and performs a wide array of functions. Combined, these molecules make up the majority of a cell's mass. Biological macromolecules are organic, meaning that they contain carbon (with some exceptions, like carbon dioxide). In addition, they may contain hydrogen, oxygen, nitrogen, phosphorus, sulfur, and additional minor elements.

Carbon

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, if not most, 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 Bonding

Carbon contains four electrons in its outer shell. Therefore, it 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 **1.7.1**).

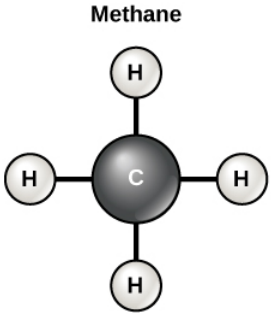


Figure **1.7.1**: Carbon can form four covalent bonds to create an organic molecule. The simplest carbon molecule is methane (CH_4), depicted here.

However, structures that are more complex are made using carbon. Any of the hydrogen atoms can be replaced with another carbon atom covalently bonded to the first carbon atom. In this way, long and branching chains of carbon compounds can be made (Figure **1.7.2a**). The carbon atoms may bond with atoms of other elements, such as nitrogen, oxygen, and phosphorus (Figure **1.7.2b**). The molecules may also form rings, which themselves can link with other rings (Figure **1.7.2c**). This diversity of molecular forms accounts for the diversity of functions of the biological macromolecules and is based to a large degree on the ability of carbon to form multiple bonds with itself and other atoms.

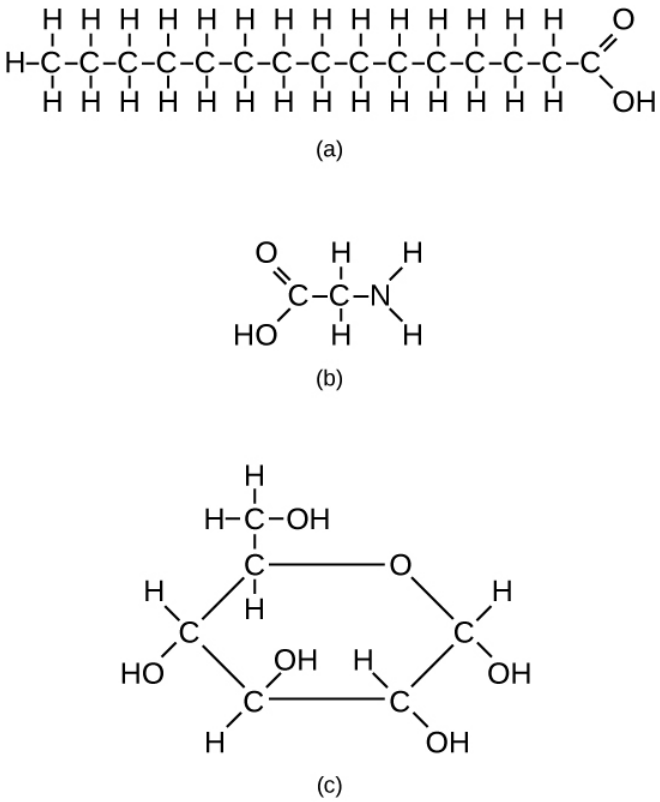


Figure 1.7.2: *These examples show three molecules (found in living organisms) that contain carbon atoms bonded in various ways to other carbon atoms and the atoms of other elements. (a) This molecule of stearic acid has a long chain of carbon atoms. (b) Glycine, a component of proteins, contains carbon, nitrogen, oxygen, and hydrogen atoms. (c) Glucose, a sugar, has a ring of carbon atoms and one oxygen atom.*

Carbohydrates

Carbohydrates are macromolecules with which most consumers are somewhat familiar. To lose weight, some individuals adhere to "low-carb" diets. Athletes, in contrast, often "carb-load" before important competitions to ensure that they have sufficient energy to compete at a high level. Carbohydrates are, in fact, an essential part of our diet; grains, fruits, and vegetables are all natural sources of carbohydrates. Carbohydrates provide energy to the body, particularly through glucose, a simple sugar. Carbohydrates also have other important functions in humans, animals, and plants.

Carbohydrates can be represented by the formula $(CH_2O)_n$, where n is the number of carbon atoms in the molecule. In other words, the ratio of carbon to hydrogen to oxygen is 1:2:1 in carbohydrate molecules. Carbohydrates are classified into three subtypes: monosaccharides, disaccharides, and polysaccharides.

Monosaccharides (mono- = "one"; sacchar- = "sweet") are simple sugars, the most common of which is glucose. In monosaccharides, the number of carbon atoms usually ranges from three to six. Most monosaccharide names end with the suffix -ose. Depending on the number of carbon atoms in the sugar, they may be known as trioses (three carbon atoms), pentoses (five carbon atoms), and hexoses (six carbon atoms).

Monosaccharides may exist as a linear chain or as ring-shaped molecules; in aqueous solutions, they are usually found in the ring form.

The chemical formula for glucose is $C_6H_{12}O_6$. In most living species, glucose is an important source of energy. During cellular respiration, energy is released from glucose, and that energy is used to help make adenosine triphosphate (ATP). Plants synthesize glucose using carbon dioxide and water by the process of photosynthesis, and the glucose, in turn, is used for the energy requirements of the plant. The excess synthesized glucose is often stored as starch that is broken down by other organisms that feed on plants.

Galactose (part of lactose, or milk sugar) and fructose (found in fruit) are other common monosaccharides. Although glucose, galactose, and fructose all have the same chemical formula ($C_6H_{12}O_6$), they differ structurally and chemically (and are known as isomers) because of differing arrangements of atoms in the carbon chain (Figure **1.7.3**).

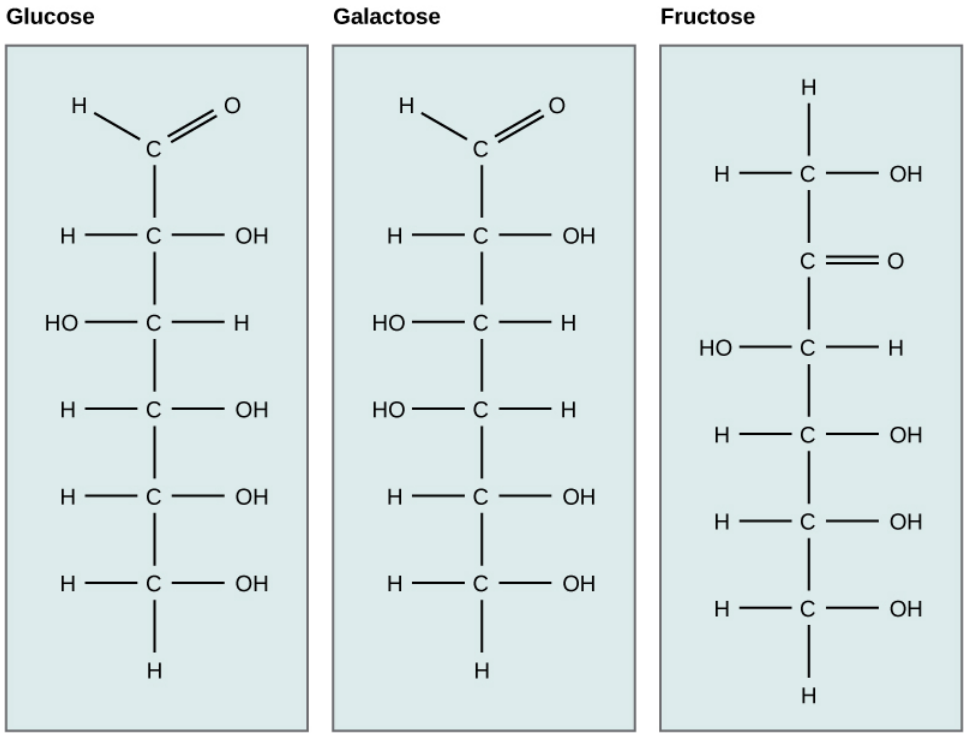


Figure **1.7.3**: Glucose, galactose, and fructose are isomeric monosaccharides, meaning that they have the same chemical formula but slightly different structures.

Disaccharides (di- = “two”) form when two monosaccharides undergo a dehydration reaction (a reaction in which the removal of a water molecule occurs). During this process, the hydroxyl group (–OH) of one monosaccharide combines with a hydrogen atom of another monosaccharide, releasing a molecule of water (H_2O) and forming a covalent bond between atoms in the two sugar molecules.

Common disaccharides include lactose, maltose, and sucrose. Lactose is a disaccharide consisting of the monomers glucose and galactose. It is found naturally in milk. Maltose, or malt sugar, is a disaccharide formed from a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is composed of the monomers glucose and fructose.

A long chain of monosaccharides linked by covalent bonds is known as a polysaccharide (poly- = “many”). The chain may be branched or unbranched, and it may contain different types of monosaccharides. Polysaccharides may be very large molecules. Starch, glycogen, cellulose, and chitin are examples of polysaccharides.

Starch is the stored form of sugars in plants and is made up of amylose and amylopectin (both polymers of glucose). Plants are able to synthesize glucose, and the excess glucose is stored as starch in different plant parts, including roots and seeds. The starch that is consumed by animals is broken down into smaller molecules, such as glucose. The cells can then absorb the glucose.

Glycogen is the storage form of glucose in humans and other vertebrates, and is made up of monomers of glucose. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever glucose levels decrease, glycogen is broken down to release glucose.

Cellulose is one of the most abundant natural biopolymers. The cell walls of plants are mostly made of cellulose, which provides structural support to the cell. Wood and paper are mostly cellulosic in nature. Cellulose is made up of glucose monomers that are linked by bonds between particular carbon atoms in the glucose molecule.

Every other glucose monomer in cellulose is flipped over and packed tightly as extended long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. Cellulose passing through our digestive system is called

dietary fiber. While the glucose-glucose bonds in cellulose cannot be broken down by human digestive enzymes, herbivores such as cows, buffalos, and horses are able to digest grass that is rich in cellulose and use it as a food source. In these animals, certain species of bacteria reside in the rumen (part of the digestive system of herbivores) and secrete the enzyme cellulase. The appendix also contains bacteria that break down cellulose, giving it an important role in the digestive systems of ruminants. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal.

Carbohydrates serve other functions in different animals. Arthropods, such as insects, spiders, and crabs, have an outer skeleton, called the exoskeleton, which protects their internal body parts. This exoskeleton is made of the biological macromolecule chitin, which is a nitrogenous carbohydrate. It is made of repeating units of a modified sugar containing nitrogen.

Thus, through differences in molecular structure, carbohydrates are able to serve the very different functions of energy storage (starch and glycogen) and structural support and protection (cellulose and chitin) (Figure 1.7.4).

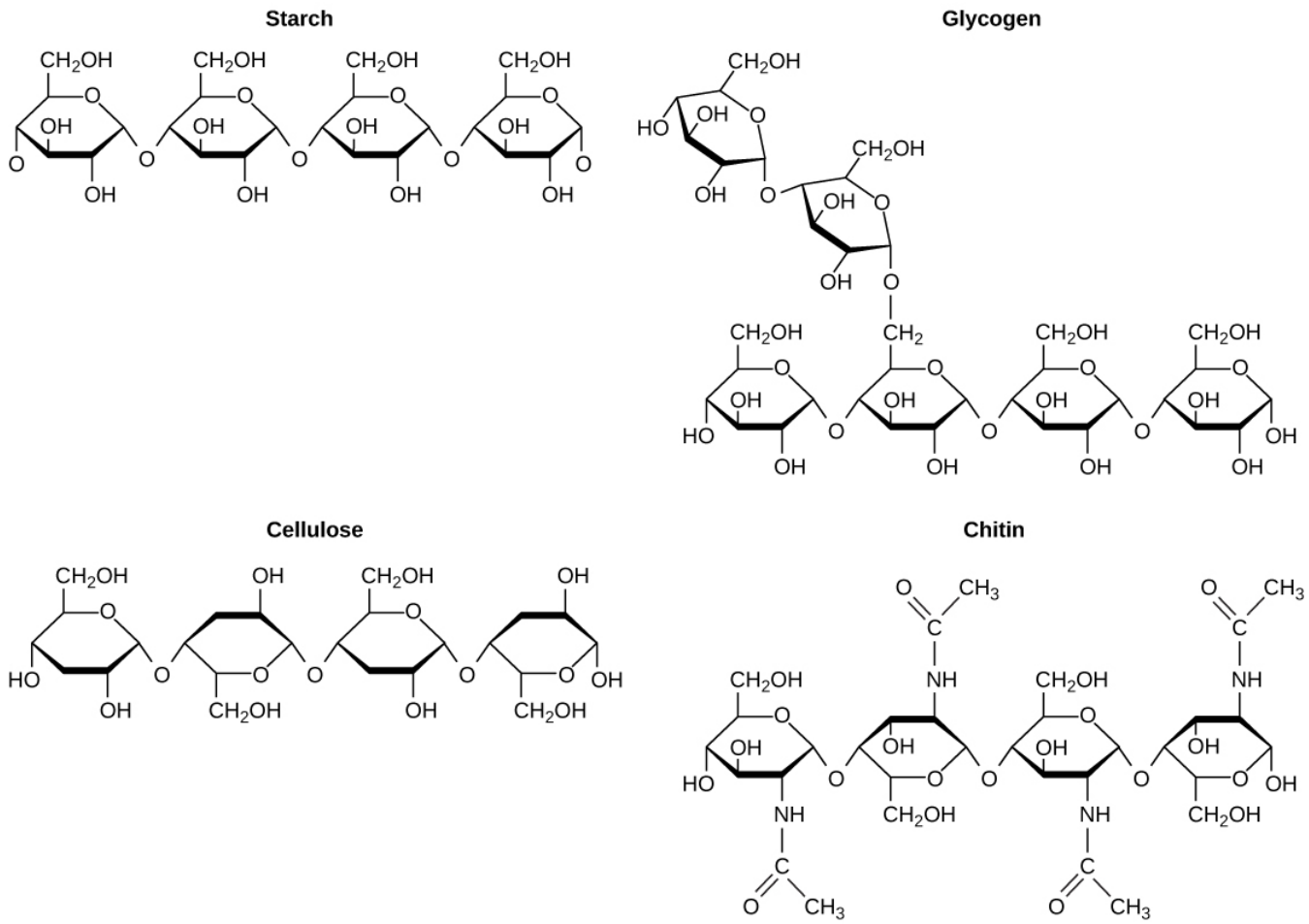


Figure 1.7.4: *Although their structures and functions differ, all polysaccharide carbohydrates are made up of monosaccharides and have the chemical formula $(CH_2O)n$.*

CAREERS IN ACTION: Registered Dietitian

Obesity is a worldwide health concern, and many diseases, such as diabetes and heart disease, are becoming more prevalent because of obesity. This is one of the reasons why registered dietitians are increasingly sought after for advice. Registered dietitians help plan food and nutrition programs for individuals in various settings. They often work with patients in health-care facilities, designing nutrition plans to prevent and treat diseases. For example, dietitians may teach a patient with diabetes how to manage blood-sugar levels by eating the correct types and amounts of carbohydrates. Dietitians may also work in nursing homes, schools, and private practices.

To become a registered dietitian, one needs to earn at least a bachelor's degree in dietetics, nutrition, food technology, or a related field. In addition, registered dietitians must complete a supervised internship program and pass a national exam. Those

who pursue careers in dietetics take courses in nutrition, chemistry, biochemistry, biology, microbiology, and human physiology. Dietitians must become experts in the chemistry and functions of food (proteins, carbohydrates, and fats).

Lipids

Lipids include a diverse group of compounds that are united by a common feature. Lipids are hydrophobic ("water-fearing"), or insoluble in water, because they are nonpolar molecules. This is because they are hydrocarbons that include only nonpolar carbon-carbon or carbon-hydrogen bonds. 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 (Figure 1.7.5). 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 plasma membrane. Lipids include fats, oils, waxes, phospholipids, and steroids.

Figure 1.7.5: Hydrophobic lipids in the fur of aquatic mammals, such as this river otter, protect them from the elements. (credit: Ken Bosma)

A fat molecule, such as a triglyceride, consists of two main components—glycerol and fatty acids. Glycerol is an organic compound with three carbon atoms, five hydrogen atoms, and three hydroxyl (–OH) groups. Fatty acids have a long chain of hydrocarbons to which an acidic carboxyl group is attached, hence the name "fatty acid." The number of carbons in the fatty acid may range from 4 to 36; most common are those containing 12–18 carbons. In a fat molecule, a fatty acid is attached to each of the three oxygen atoms in the –OH groups of the glycerol molecule with a covalent bond (Figure 1.7.6).

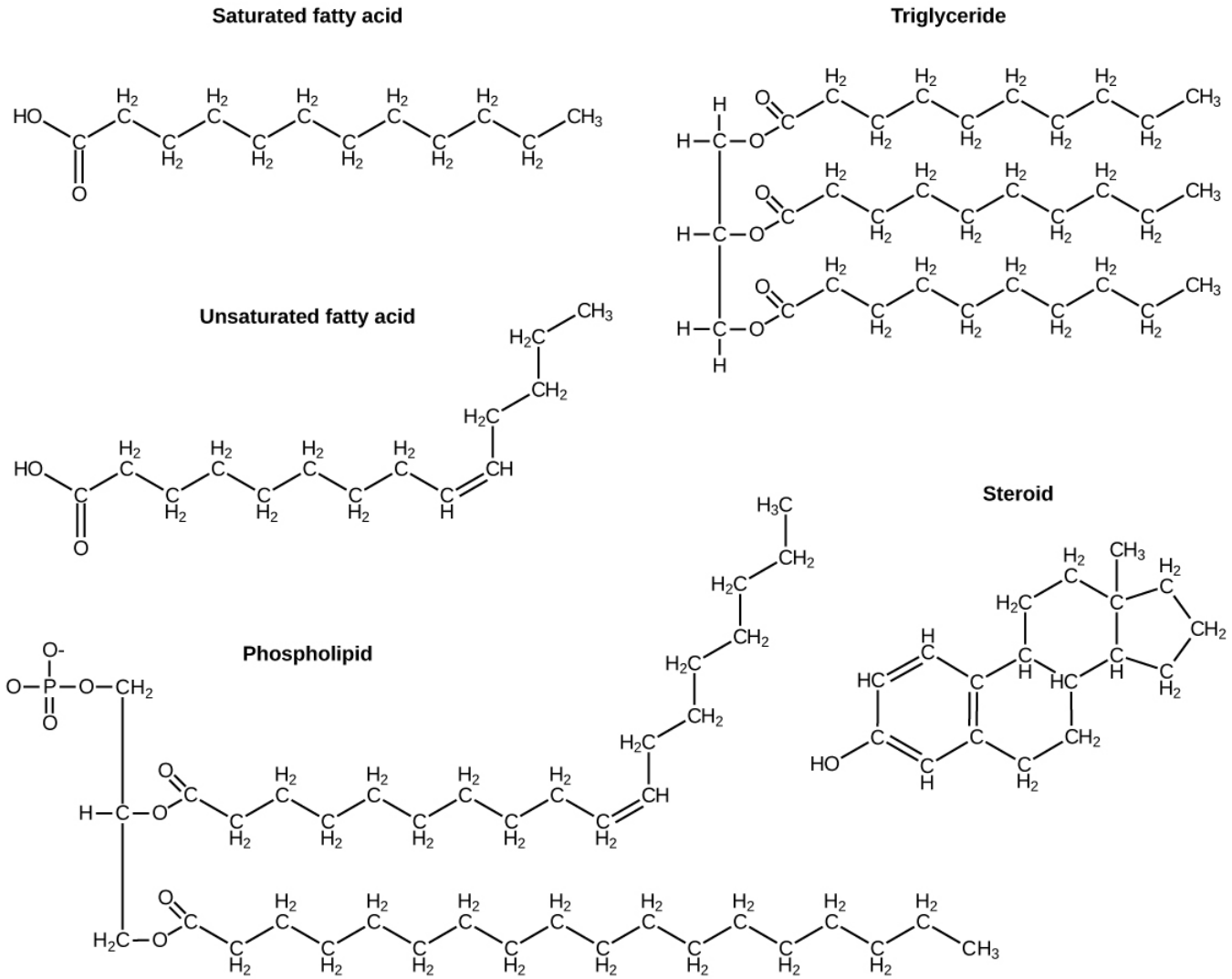


Figure 1.7.6: *Lipids include fats, such as triglycerides, which are made up of fatty acids and glycerol, phospholipids, and steroids.*

During this covalent bond formation, three water molecules are released. The three fatty acids in the fat may be similar or dissimilar. These fats are also called triglycerides because they have three fatty acids. Some fatty acids have common names that specify their origin. For example, palmitic acid, a saturated fatty acid, is derived from the palm tree. Arachidic acid is derived from *Arachis hypogaea*, the scientific name for peanuts.

Fatty acids may be saturated or unsaturated. In a fatty acid chain, if there are only single bonds between neighboring carbons in the hydrocarbon chain, the fatty acid is saturated. Saturated fatty acids are saturated with hydrogen; in other words, the number of hydrogen atoms attached to the carbon skeleton is maximized.

When the hydrocarbon chain contains a double bond, the fatty acid is an unsaturated fatty acid.

Most unsaturated fats are liquid at room temperature and are called oils. If there is one double bond in the molecule, then it is known as a monounsaturated fat (e.g., olive oil), and if there is more than one double bond, then it is known as a polyunsaturated fat (e.g., canola oil).

Saturated fats tend to get packed tightly and are solid at room temperature. Animal fats with stearic acid and palmitic acid contained in meat, and the fat with butyric acid contained in butter, are examples of saturated fats. Mammals store fats in specialized cells called adipocytes, where globules of fat occupy most of the cell. In plants, fat or oil is stored in seeds and is used as a source of energy during embryonic development.

Unsaturated fats or oils are usually of plant origin and contain unsaturated fatty acids. The double bond causes a bend or a “kink” that prevents the fatty acids from packing tightly, keeping them liquid at room temperature. Olive oil, corn oil, canola oil, and cod liver oil are examples of unsaturated fats. Unsaturated fats help to improve blood cholesterol levels, whereas saturated fats contribute to plaque formation in the arteries, which increases the risk of a heart attack.

In the food industry, oils are artificially hydrogenated to make them semi-solid, leading to less spoilage and increased shelf life. Simply speaking, hydrogen gas is bubbled through oils to solidify them. During this hydrogenation process, double bonds of the *cis*-conformation in the hydrocarbon chain may be converted to double bonds in the *trans*-conformation. This forms a *trans*-fat from a *cis*-fat. The orientation of the double bonds affects the chemical properties of the fat (Figure 1.7.7).

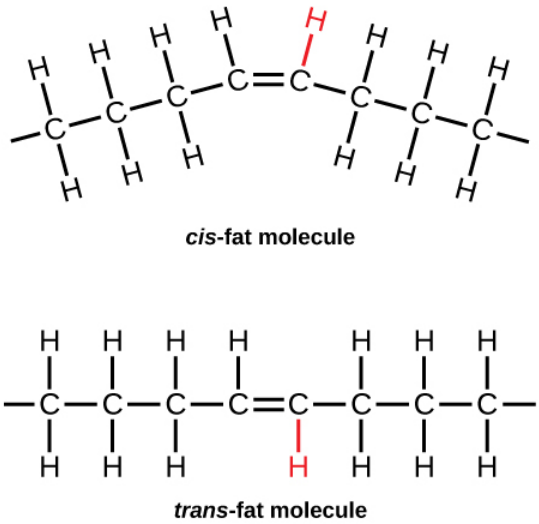


Figure 1.7.7: During the hydrogenation process, the orientation around the double bonds is changed, making a trans-fat from a cis-fat. This changes the chemical properties of the molecule.

Margarine, some types of peanut butter, and shortening are examples of artificially hydrogenated *trans*-fats. Recent studies have shown that an increase in *trans*-fats in the human diet may lead to an increase in levels of low-density lipoprotein (LDL), or “bad” cholesterol, which, in turn, may lead to plaque deposition in the arteries, resulting in heart disease. Many fast food restaurants have recently eliminated the use of *trans*-fats, and U.S. food labels are now required to list their *trans*-fat content.

Essential fatty acids are fatty acids that are required but not synthesized by the human body. Consequently, they must be supplemented through the diet. Omega-3 fatty acids fall into this category and are one of only two known essential fatty acids for humans (the other being omega-6 fatty acids). They are a type of polyunsaturated fat and are called omega-3 fatty acids because the third carbon from the end of the fatty acid participates in a double bond.

Salmon, trout, and tuna are good sources of omega-3 fatty acids. Omega-3 fatty acids are important in brain function and normal growth and development. They may also prevent heart disease and reduce the risk of cancer.

Like carbohydrates, fats have received a lot of bad publicity. It is true that eating an excess of fried foods and other “fatty” foods leads to weight gain. However, fats do have important functions. Fats serve as long-term energy storage. They also provide insulation for the body. Therefore, “healthy” unsaturated fats in moderate amounts should be consumed on a regular basis.

Phospholipids are the major constituent of the plasma membrane. Like fats, they are composed of fatty acid chains attached to a glycerol or similar backbone. Instead of three fatty acids attached, however, there are two fatty acids and the third carbon of the glycerol backbone is bound to a phosphate group. The phosphate group is modified by the addition of an alcohol.

A phospholipid has both hydrophobic and hydrophilic regions. The fatty acid chains are hydrophobic and exclude themselves from water, whereas the phosphate is hydrophilic and interacts with water.

Cells are surrounded by a membrane, which has a bilayer of phospholipids. The fatty acids of phospholipids face inside, away from water, whereas the phosphate group can face either the outside environment or the inside of the cell, which are both aqueous.

Steroids and Waxes

Unlike the phospholipids and fats discussed earlier, steroids have a ring structure. Although they do not resemble other lipids, they are grouped with them because they are also hydrophobic. All steroids have four, linked carbon rings and several of them, like cholesterol, have a short tail.

Cholesterol is a steroid. Cholesterol is mainly synthesized in the liver and is the precursor of many steroid hormones, such as testosterone and estradiol. It is also the precursor of vitamins E and K. Cholesterol is the precursor of bile salts, which help in the breakdown of fats and their subsequent absorption by cells. Although cholesterol is often spoken of in negative terms, it is necessary for the proper functioning of the body. It is a key component of the plasma membranes of animal cells.

Waxes are made up of a hydrocarbon chain with an alcohol (–OH) group and a fatty acid. Examples of animal waxes include beeswax and lanolin. Plants also have waxes, such as the coating on their leaves, that helps prevent them from drying out.

CONCEPT IN ACTION

For an additional perspective on lipids, explore “Biomolecules: The Lipids” through this interactive animation.

Proteins

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective; they may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of different proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of amino acids, arranged in a linear sequence.

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. For example, proteins can function as enzymes or hormones. Enzymes, which are produced by living cells, are catalysts in biochemical reactions (like digestion) and are usually proteins. Each enzyme is specific for the substrate (a reactant that binds to an enzyme) upon which it acts. Enzymes can function to break molecular bonds, to rearrange bonds, or to form new bonds. An example of an enzyme is salivary amylase, which breaks down amylose, a component of starch.

Hormones are chemical signaling molecules, usually proteins or steroids, secreted by an endocrine gland or group of endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that maintains blood glucose levels.

Proteins have different shapes and molecular weights; some proteins are globular in shape whereas others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, found in our skin, is a fibrous protein. Protein shape is critical to its function. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the shape of the protein, leading to a loss of function or denaturation (to be discussed in more detail later). All proteins are made up of different arrangements of the same 20 kinds of amino acids.

Amino acids are the monomers that make up proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom bonded to an amino group ($–NH_2$), a carboxyl group (–COOH), and a hydrogen atom. Every amino acid also has another variable atom or group of atoms bonded to the central carbon atom known as the R group. The R group is the only difference in structure between the 20 amino acids; otherwise, the amino acids are identical (Figure 1.7.8).

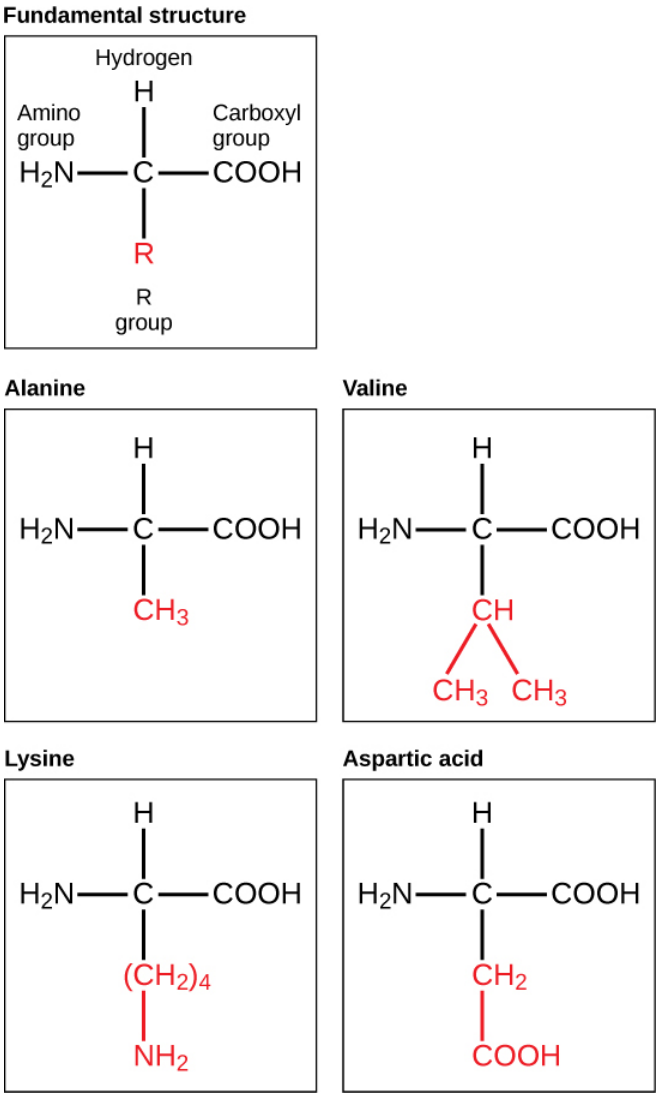


Figure **1.7.8**: Amino acids are made up of a central carbon bonded to an amino group ($–NH_2$), a carboxyl group (–COOH), and a hydrogen atom. The central carbon's fourth bond varies among the different amino acids, as seen in these examples of alanine, valine, lysine, and aspartic acid.

The chemical nature of the R group determines the chemical nature of the amino acid within its protein (that is, whether it is acidic, basic, polar, or nonpolar).

The sequence and number of amino acids ultimately determine a protein's shape, size, and function. Each amino acid is attached to another amino acid by a covalent bond, known as a peptide bond, which is formed by a dehydration reaction. The carboxyl group of one amino acid and the amino group of a second amino acid combine, releasing a water molecule. The resulting bond is the peptide bond.

The products formed by such a linkage are called polypeptides. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, have a distinct shape, and have a unique function.

EVOLUTION IN ACTION: The Evolutionary Significance of Cytochrome c

Cytochrome c is an important component of the molecular machinery that harvests energy from glucose. Because this protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of sequence similarity among cytochrome c molecules of different species; evolutionary relationships can be assessed by measuring the similarities or differences among various species' protein sequences.

For example, scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule that has been sequenced to date from different organisms, 37 of these amino acids appear in the same position in each cytochrome c. This indicates that all of these organisms are descended from a common ancestor. On comparing the human and chimpanzee protein sequences, no sequence difference was found. When human and rhesus monkey sequences were compared, a single difference was found in one amino acid. In contrast, human-to-yeast comparisons show a difference in 44 amino acids, suggesting that humans and chimpanzees have a more recent common ancestor than humans and the rhesus monkey, or humans and yeast.

Protein Structure

As discussed earlier, the shape of a protein is critical to its function. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary (Figure **1.7.9**).

The unique sequence and number of amino acids in a polypeptide chain is its primary structure. The unique sequence for every protein is ultimately determined by the gene that encodes the protein. Any change in the gene sequence may lead to a different amino acid being added to the polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin β chain has a single amino acid substitution, causing a change in both the structure and function of the protein. What is most remarkable to consider is that a hemoglobin molecule is made up of two alpha chains and two beta chains that each consist of about 150 amino acids. The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—that dramatically decreases life expectancy in the affected individuals—is a single amino acid of the 600.

Because of this change of one amino acid in the chain, the normally biconcave, or disc-shaped, red blood cells assume a crescent or “sickle” shape, which clogs arteries. This can lead to a myriad of serious health problems, such as breathlessness, dizziness, headaches, and abdominal pain for those who have this disease.

Folding patterns resulting from interactions between the non-R group portions of amino acids give rise to the secondary structure of the protein. The most common are the alpha (α)-helix and beta (β)-pleated sheet structures. Both structures are held in shape by hydrogen bonds. In the alpha helix, the bonds form between every fourth amino acid and cause a twist in the amino acid chain.

In the β-pleated sheet, the “pleats” are formed by hydrogen bonding between atoms on the backbone of the polypeptide chain. The R groups are attached to the carbons, and extend above and below the folds of the pleat. The pleated segments align parallel to each other, and hydrogen bonds form between the same pairs of atoms on each of the aligned amino acids. The α-helix and β-pleated sheet structures are found in many globular and fibrous proteins.

The unique three-dimensional structure of a polypeptide is known as its tertiary structure. This structure is caused by chemical interactions between various amino acids and regions of the polypeptide. Primarily, the interactions among R groups create the complex three-dimensional tertiary structure of a protein. There may be ionic bonds formed between R groups on different amino acids, or hydrogen bonding beyond that involved in the secondary structure. When protein folding takes place, the hydrophobic R groups of nonpolar amino acids lay in the interior of the protein, whereas the hydrophilic R groups lay on the outside. The former types of interactions are also known as hydrophobic interactions.

In nature, some proteins are formed from several polypeptides, also known as subunits, and the interaction of these subunits forms the quaternary structure. Weak interactions between the subunits help to stabilize the overall structure. For example, hemoglobin is a combination of four polypeptide subunits.

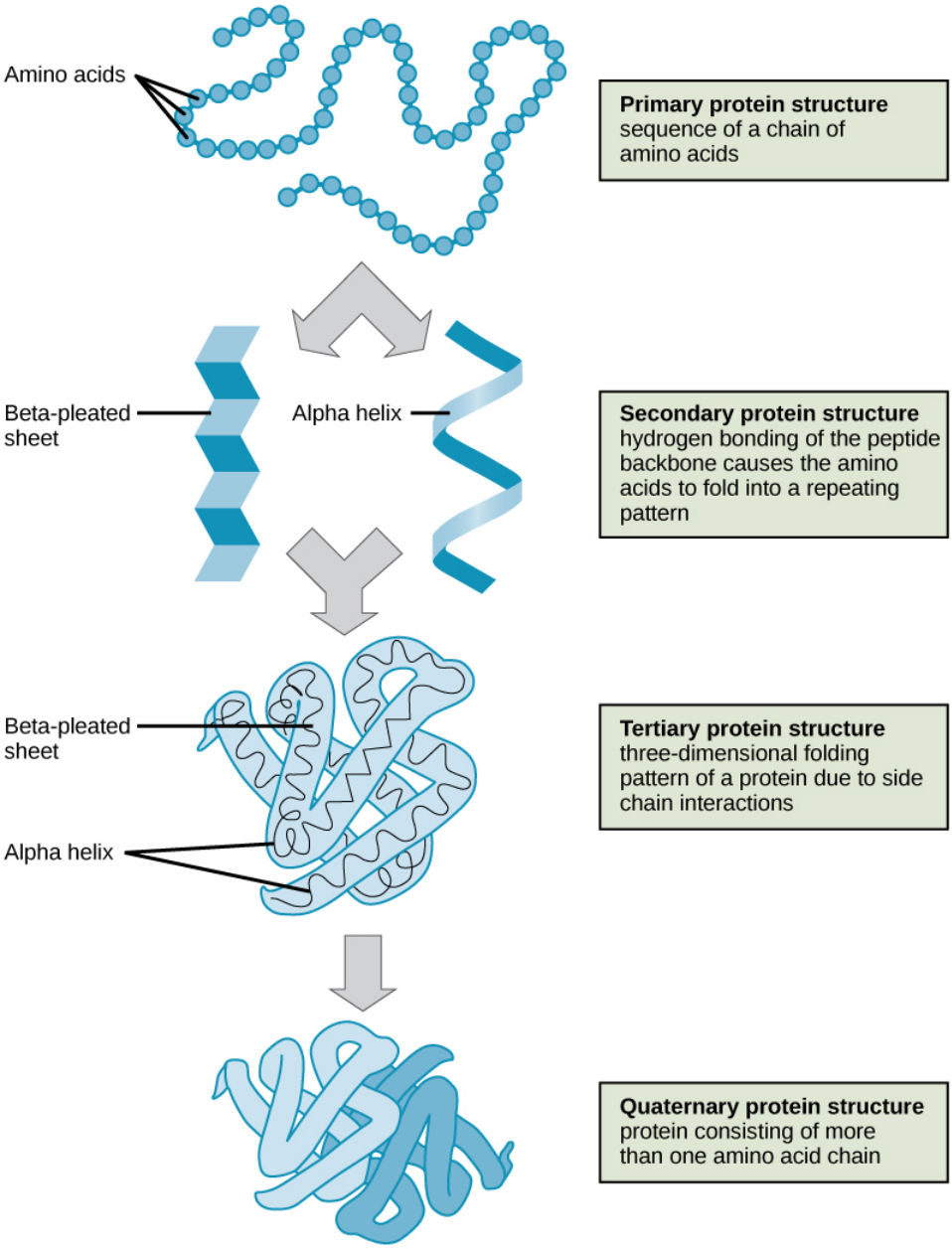


Figure 1.7.9: The four levels of protein structure can be observed in these illustrations. (credit: modification of work by National Human Genome Research Institute)

Each protein has its own unique sequence and shape held together by chemical interactions. If the protein is subject to changes in temperature, pH, or exposure to chemicals, the protein structure may change, losing its shape in what is known as denaturation as discussed earlier. Denaturation is often reversible because the primary structure is preserved if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to a loss of function. One example of protein denaturation can be seen when an egg is fried or boiled. The albumin protein in the liquid egg white is denatured when placed in a hot pan, changing from a clear substance to an opaque white substance. Not all proteins are denatured at high temperatures; for instance, bacteria that survive in hot springs have proteins that are adapted to function at those temperatures.

CONCEPT IN ACTION

For an additional perspective on proteins, explore “Biomolecules: The Proteins” through this interactive animation.

Nucleic Acids

Nucleic acids are key macromolecules in the continuity of life. They carry the genetic blueprint of a cell and carry instructions for the functioning of the cell.

The two main types of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the genetic material found in all living organisms, ranging from single-celled bacteria to multicellular mammals.

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus, but instead use an RNA intermediary to communicate with the rest of the cell. Other types of RNA are also involved in protein synthesis and its regulation.

DNA and RNA are made up of monomers known as nucleotides. The nucleotides combine with each other to form a polynucleotide, DNA or RNA. Each nucleotide is made up of three components: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group (Figure 1.7.10). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to a phosphate group.

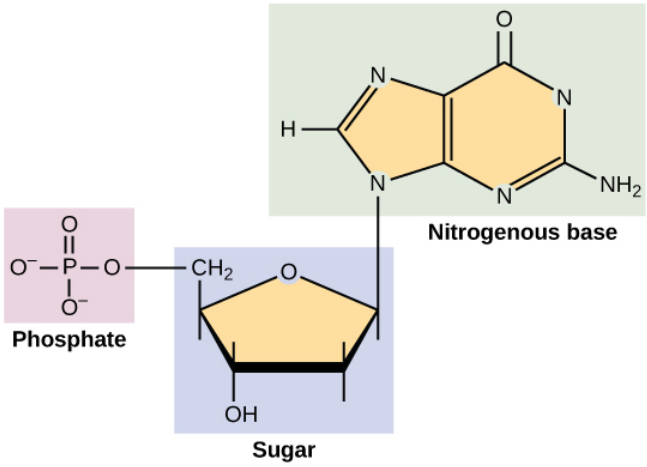


Figure 1.7.10: A nucleotide is made up of three components: a nitrogenous base, a pentose sugar, and a phosphate group.

DNA Double-Helical Structure

DNA has a double-helical structure (Figure 1.7.11). It is composed of two strands, or polymers, of nucleotides. The strands are formed with bonds between phosphate and sugar groups of adjacent nucleotides. The strands are bonded to each other at their bases with hydrogen bonds, and the strands coil about each other along their length, hence the “double helix” description, which means a double spiral.

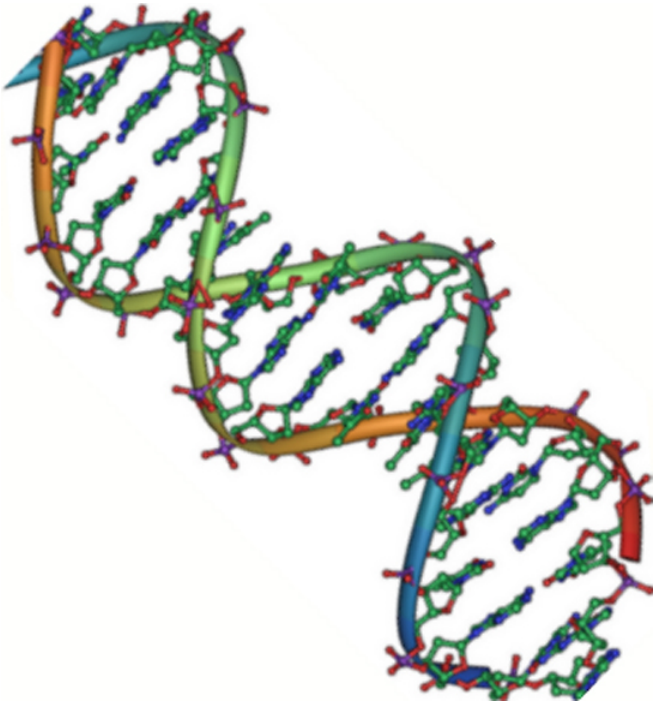


Figure **1.7.11**: The double-helix model shows DNA as two parallel strands of intertwining molecules. (credit: Jerome Walker, Dennis Myts)

The alternating sugar and phosphate groups lie on the outside of each strand, forming the backbone of the DNA. The nitrogenous bases are stacked in the interior, like the steps of a staircase, and these bases pair; the pairs are bound to each other by hydrogen bonds. The bases pair in such a way that the distance between the backbones of the two strands is the same all along the molecule.

Summary

Living things are carbon-based because carbon plays such a prominent role in the chemistry of living things. The four covalent bonding positions of the carbon atom can give rise to a wide diversity of compounds with many functions, accounting for the importance of carbon in living things. Carbohydrates are a group of macromolecules that are a vital energy source for the cell, provide structural support to many organisms, and can be found on the surface of the cell as receptors or for cell recognition. Carbohydrates are classified as monosaccharides, disaccharides, and polysaccharides, depending on the number of monomers in the molecule.

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats and oils are a stored form of energy and can include triglycerides. Fats and oils are usually made up of fatty acids and glycerol.

Proteins are a class of macromolecules that can perform a diverse range of functions for the cell. They help in metabolism by providing structural support and by acting as enzymes, carriers or as hormones. The building blocks of proteins are amino acids. Proteins are organized at four levels: primary, secondary, tertiary, and quaternary. Protein shape and function are intricately linked; any change in shape caused by changes in temperature, pH, or chemical exposure may lead to protein denaturation and a loss of function.

Nucleic acids are molecules made up of repeating units of nucleotides that direct cellular activities such as cell division and protein synthesis. Each nucleotide is made up of a pentose sugar, a nitrogenous base, and a phosphate group. There are two types of nucleic acids: DNA and RNA.

Glossary

amino acid

a monomer of a protein

carbohydrate

a biological macromolecule in which the ratio of carbon to hydrogen to oxygen is 1:2:1; carbohydrates serve as energy sources and structural support in cells

cellulose

a polysaccharide that makes up the cell walls of plants and provides structural support to the cell

chitin

a type of carbohydrate that forms the outer skeleton of arthropods, such as insects and crustaceans, and the cell walls of fungi

denaturation

the loss of shape in a protein as a result of changes in temperature, pH, or exposure to chemicals

deoxyribonucleic acid (DNA)

a double-stranded polymer of nucleotides that carries the hereditary information of the cell

disaccharide

two sugar monomers that are linked together by a peptide bond

enzyme

a catalyst in a biochemical reaction that is usually a complex or conjugated protein

fat

a lipid molecule composed of three fatty acids and a glycerol (triglyceride) that typically exists in a solid form at room temperature

glycogen

a storage carbohydrate in animals

hormone

a chemical signaling molecule, usually a protein or steroid, secreted by an endocrine gland or group of endocrine cells; acts to control or regulate specific physiological processes

lipids

a class of macromolecules that are nonpolar and insoluble in water

macromolecule

a large molecule, often formed by polymerization of smaller monomers

monosaccharide

a single unit or monomer of carbohydrates

nucleic acid

a biological macromolecule that carries the genetic information of a cell and carries instructions for the functioning of the cell

nucleotide

a monomer of nucleic acids; contains a pentose sugar, a phosphate group, and a nitrogenous base

oil

an unsaturated fat that is a liquid at room temperature

phospholipid

a major constituent of the membranes of cells; composed of two fatty acids and a phosphate group attached to the glycerol backbone

polypeptide

a long chain of amino acids linked by peptide bonds

polysaccharide

a long chain of monosaccharides; may be branched or unbranched

protein

a biological macromolecule composed of one or more chains of amino acids

ribonucleic acid (RNA)

a single-stranded polymer of nucleotides that is involved in protein synthesis

saturated fatty acid

a long-chain hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

starch

a storage carbohydrate in plants

steroid

a type of lipid composed of four fused hydrocarbon rings

***trans*-fat**

a form of unsaturated fat with the hydrogen atoms neighboring the double bond across from each other rather than on the same side of the double bond

triglyceride

a fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid

a long-chain hydrocarbon that has one or more than one double bonds in the hydrocarbon chain

Contributors and Attributions

- Samantha Fowler (Clayton State University), Rebecca Roush (Sandhills Community College), James Wise (Hampton University). Original content by OpenStax (CC BY 4.0; Access for free at https://cnx.org/contents/b3c1e1d2-83...4-e119a8aafbdd).

This page titled 1.7: Types of Biological Molecules is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by OpenStax.

- **2.3: Biological Molecules** by OpenStax is licensed CC BY 4.0.
- **2.4: Types of Biological Molecules** by OpenStax is licensed CC BY 4.0.

1.8: Chapter 1 Questions

Section 1 Questions

Question 1.8.1

In Figure 1.2, two examples of types of enzyme-substrate binding are shown: the Lock-and-Key model and Induced-Fit. What are some situations in which one style of the enzyme would be favored over the other?

Answer

Lock and key enzymes are highly specific for their substrate and therefore do not need a transition state to undergo the catalytic reaction. This could be used for substrate channels like Na+/K+ pumps in which a reaction doesn't need to occur.

Induced fit enzymes utilize a transition state, to convert a substrate into a product. The transition state is able to cause a conformational change in the active site of the enzyme and facilitate high-energy reactions such as breaking or forming chemical bonds

Question 1.8.2

Label the following type of import/export mechanisms as passive, active, or facilitated and explain why: endocytosis, ion channels, pores, transporters/permeases. *Some may have more than one answer.*

Answer

Endocytosis: Active, facilitated. Endocytosis or "cell eating" is a multi-enzyme mediated process that allows the cell to uptake large particles from its environment. This involves membrane modification, protein receptors, and digestive enzymes and organelles working across gradients.

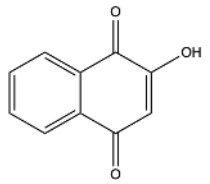
Pores: Passive, facilitated. Once porins establish pores, such as in the nuclear envelope, small molecules like DNA and RNA can passively diffuse in and out of the membrane without the need for carrier proteins.

Ion Channels: Active, facilitated OR passive, facilitated. Active ion channels pump small molecules across a gradient and are typically considered to be "gated," meaning that the enzymes can open and close in a regulated manner to control what is being moved across the membrane. Passive ion channels are permanently open to facilitate transfer and rely on a constantly established concentration gradient to allow for transport to occur.

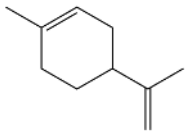
Transporters/Permeases: Active, facilitated. Transporters move larger molecules across a concentration gradient and assist in the movement of soluble proteins and molecules through the hydrophobic membrane

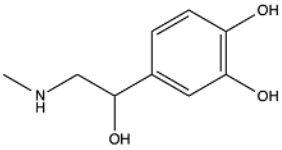
Section 2 Questions

Label the functional groups present in the chemicals shown below:

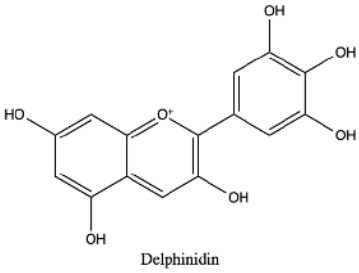


Lawsone

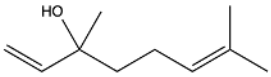

Limonene



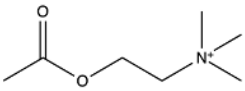
Adrenaline



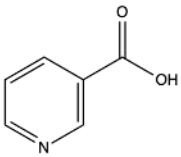
Delphinidin



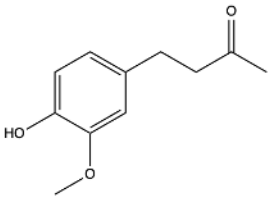
Linalool



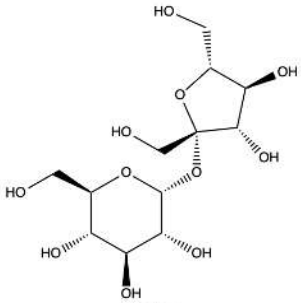
Acetylcholine



Nicotinic Acid



Zingerone



Sucrose

Answers:

Section 3 Questions

1) a. Consider a subset of reactions of glycolysis given below. ΔG'°, substrates, and products are given from colon cancer cells (nmol/g tissue). After examining the conditions of the cell for each enzymatic reaction, predict if the ΔG of the reaction will increase or decrease. (Data from Hirayama A et al. 2009 *Cancer Research*.) The ratio of NAD+/NADH is 10:1 and the concentrations of the cofactors are ATP (110) and ADP (300).

Reaction	ΔG'°	[Substrate]	[Product]
#1 Hexokinase Glucose → Glucose-6-phosphate	-16.6	123	75
#2 Phosphoglucose Isomerase Glucose-6-phosphate → Fructose-6-phosphate	1.67	75	50
#3 Phosphofructokinase Fructose-6-phosphate → Fructose-1,6-bisphosphate	-14.2	50	50
#10 Pyruvate Kinase Phosphoenolpyruvate → Pyruvate	-31.4	5	25
Lactate Dehydrogenase	-25.1	25	25,000

Answer:
Reaction #1 - Increase.
Reaction #2 - Decrease.
Reaction #3 - Increase.
Reaction #10 - Increase.
Lactate Dehydrogenase - Increase.

2) Consider the reaction below along with the thermodynamic properties: ΔH° = -760 kJ/mol, ΔS° = -0.185 kJ/mol K, and ΔG = -705 kJ/mol

$Na^+_{(g)} + Cl^-_{(g)} \rightarrow Na^+_{(aq)} + Cl^-_{(aq)}$

At what temperature would this reaction have an equilibrium constant of 1?

Answer: $\Delta G° = RTln(K_{eq})$

Because we want know know the temperature at which $K_{eq} = 1$, and we know that the ln(1) = 0, $\Delta G° = 0$ when $K_{eq} = 1$.

$\Delta G° = \Delta H° - T\Delta S°$
0 = -760 kJ/mol - T(-0.185 kJ/mol K); Rearrange and solve for T = 4108.1 °K

This page titled 1.8: Chapter 1 Questions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **1.5: Chapter 1 Questions** by Henry Jakubowski and Patricia Flatt has no license indicated.

1.9: Abiotic Origins of Life

"Ugly giant bags of mostly water" -
The crystal life form describing humans on Star Trek Next Generation, Home Soil Episode.

1.9.0.1: **The Start of Life**

This book began with the notion that understandings derived from the study of simple molecules can be applied to complex biological macromolecules and systems. We developed an understanding of the structural, thermodynamic, and kinetic properties of the "simplest" biomolecules, including single chain amphiphiles like fatty acids, and double chain ones like phospholipids, and how these properties could explain the propensity of these molecules to form complex lipid aggregates (micelles and bilayers). We extended these ideas to the process of protein folding and the assembly of biological complexity. Is there something intrinsic to the property of molecules such that their localization together in the right microenvironment could lead to a "living cell"? How did life originate? That is the topic of this last capstone chapter.

Defining life is actually quite difficult. Here is a list of requirements that seem reasonable, but other have noted that this list would exclude the mule. Life can self-replicate, self-sustain, evolve, respond to environmental changes, and die. The earliest known fossils (stromatolites from cyanobacteria) are approximately 3.5 billion years old.

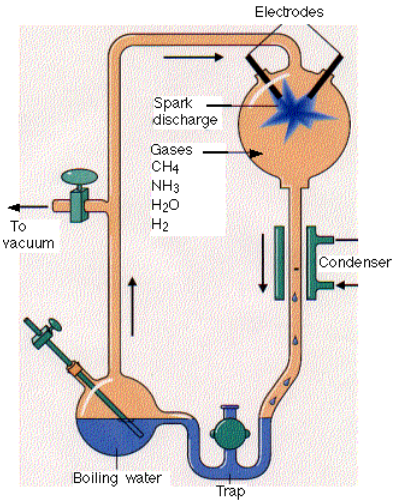
We have just finished studying the complex interactions involved in cell signaling. How could they have evolved? Consider the central dogma of biology. In the present biological world, proteins (DNA polymerase, RNA polymerase, transcription factors) are necessary for DNA synthesis, replication and gene transcription. But you need DNA to encode the proteins. This "chicken vs egg" dilemma has been addressed when it was realized that RNA can both carry genetic information as well as enzymatic activity (even at the level of the ribosome used for protein synthesis).

https://clockwise.software/blog/solv...d-egg-problem/ for new picture

1.9.0.1: Abiotic Synthesis of Amino Acids and Peptides

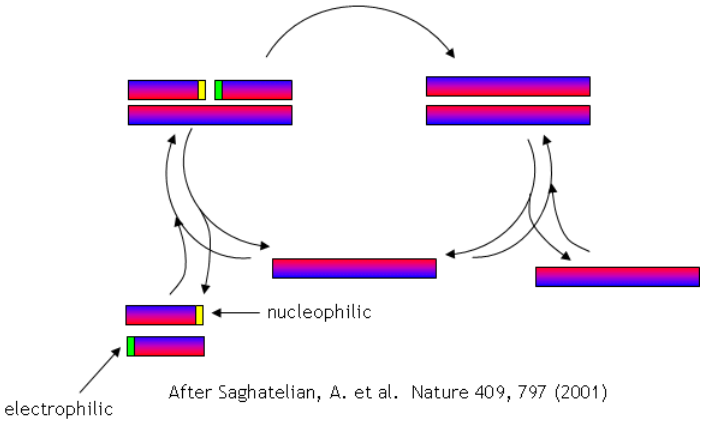
Much work has been done to determine if the building blocks for present biological molecules could have been synthesized early in Earth's history. Amino acids and fatty acids have been found in meteors suggesting the possibility. Earth's early atmosphere would have had little oxygen, so most components

should have been reduced. It probably consisted of methane, ammonia, hydrogen and water similar to the atmospheres of other planets in our solar system. The composition of the early atmosphere is still contentious. In 1953 (the same year that Watson and Crick published the structure of double-stranded DNA), Stanley Miller showed that electric discharges (to simulate lightening) in a reducing atmosphere over a "simulated sea" produced many amino acids. Up to 11 different amino acids have been produced in this fashion along with purines and pyrimidines (these required concentrated reaction mixtures) necessary for nucleic acids. Adenine can be produced just through the reaction of hydrogen cyanide and ammonia in an aqueous solution. Other nucleic acid bases can be made with hydrogen cyanide, cyanogen (C2N2) and cyanoacetylene (HC3N).



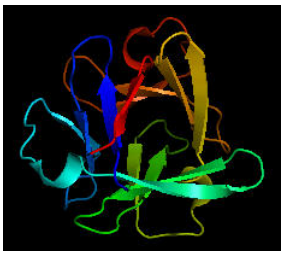
http://www.hencoup.com/Photo%20Stanley%20Miller.jpg

No complex polymers arise through these reactions. However, in 2004, Lehman, Orgel and Ghadiri were able to show that in the presence of carbon disulfide, a gas discharged from volcanoes, homo- and hetero-peptides were produced. Amphiphilic peptides can even catalyze their own formation from peptide fragments, if the fragments are activated. The fragments would bind to the larger "template" peptide through nonpolar actions of the side chains which are oriented along one face of the helical axes. If the fragments bind in a fashion in which the electrophilic end is adjacent to the nucleophilic end of the other peptide fragment, condensation of the two peptide fragments results. The larger template peptide acts as a template (effectively as an "enzyme") in orienting the two fragments for chemical reaction and effectively increasing their local concentration. The reaction of the bound fragments is essentially intramolecular. The reaction even proceeds with amplification of homochirality.



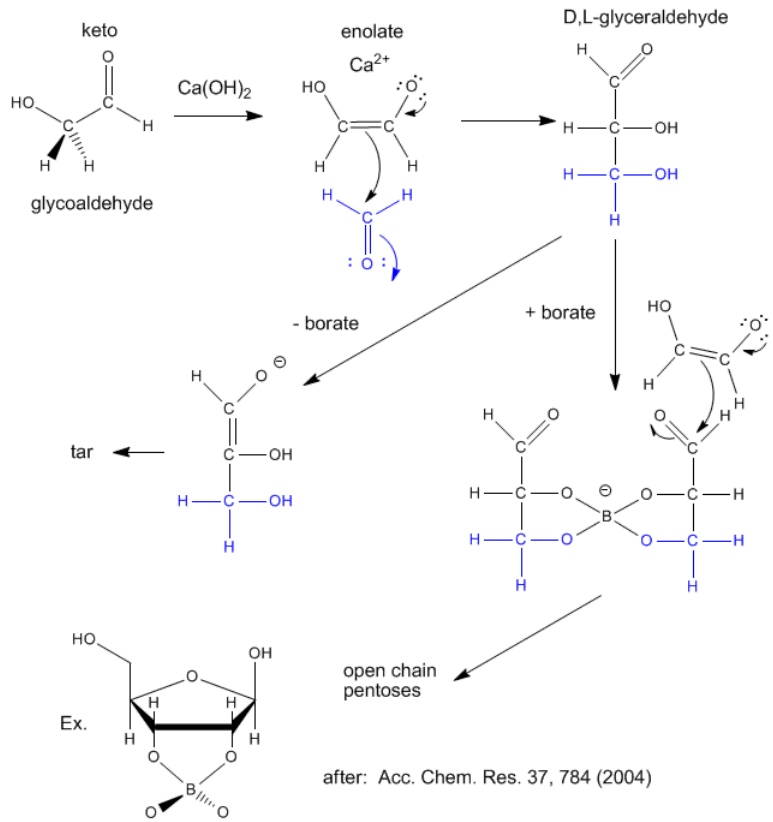
Could the prebiotic amino acids have polymerized into a protein that could fold in a fashion similar to modern proteins? That question has recently been addressed by Longo et al (2013). They asked the question whether the amino acids found in Miller-type prebiotic synthesis mixture and in comets/meteors (Ala, Asp, Glu, Gly, Ile, Leu, Pro, Ser, Thr and Val), a restricted set (10) compared to the present 20 naturally-occurring amino acid, could form a polymer that could fold. Notice that this reduced ensemble of amino acids lacks aromatic and basic amino acids. These proteins would be acidic with a low pI and may have trouble, given the lack of nonpolar aromatic amino acids, in forming a buried hydrophobic core which stabilize proteins. Nevertheless Longo et al were able to synthesize a protein with a slightly expanded set of amino acids (12, including Asn and Gln, with 70% prebiotic amino acids). The structure of one of the proteins, PV2, is shown below. The protein was more stable in 2 M NaCl (compared to 0.1 M) in which it showed a cooperative thermal denaturation with a melting point near 650C using differential scanning calorimetry. The protein had properties similar to those from halophilic organisms that thrive in high salt. These properties include low pIs and high negative charge density, which allows cation-protein interactions in the high salt environment, and lower stability in low salt environments. Earlier oceans were saltier. Halophiles are an example of extremeophiles which are highly represented in archea. Although most halophiles are aerobic, some are anerobic. Perhaps life arose in high salt environments.

Figure: Structure of PV2 protein comprised of a reduced alphabet of mainly prebiotic amino acids.



1.9.0.1: Abiotic Synthesis of Sugars

Sugars are required for present energy production but also as a part of the backbone (ribose, deoxyribose) of present genetic material. Many sugars can be synthesized in prebiotic conditions, using carbon based molecules with oxygen, such as glycoaldehyde and formaldehyde (both found in interstellar gases), as shown in the figure below. Glycoaldehyde was recently found in star forming regions of the Milky Way where planets are likely to form. The presence of borate, which stabilizes vicinal OHs on sugars, is required for the production of sugars instead of tars.



RNA molecules containing sugars such as threose, aldopentopyraonses and hexopyranose can also form stable secondary structures like helices. (Remember, RNA probably preceded DNA as the genetic carrier of information given that is also has enzymatic activity). Is there something special about ribose that made it selected over other sugars for nucleic acids, especially since it is found in low abundancy in the products in synthesis reactions conducted under prebiotic conditions? One probable reason is it unusually high (compared to other sugars) permeability coefficient through vesicles made of phospholipids or single chain fatty acids, as shown below.

Table 1. Solute permeability coefficients (P_s) across fatty acid and phospholipid membranes

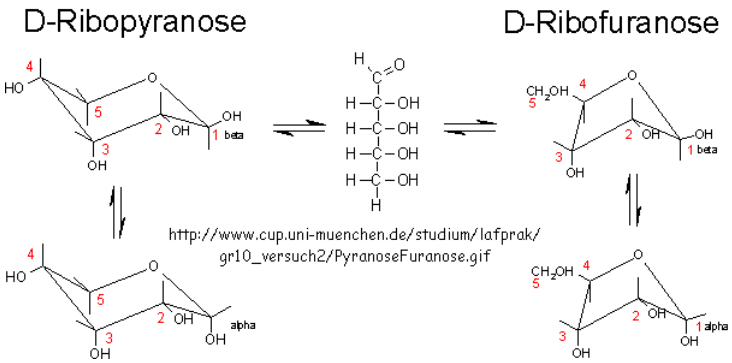
		Membrane composition					
	Solute	Myristoleate*	Palmitoleate*	Oleate*	Oleate†	POPC*	DPPC/C/PA‡
	Glycerol§			490			
4C	Erythritol§	28	14	12	4.8	21	
4C	DL-Threitol	50	23	22	11	40	
	Adonitol	2.9	1.7	0.91	0.38		
	Arabitol	2.1	1.4	0.63	0.39		
	Xylitol	5.0	3.0	1.8	0.95		
	Dulcitol				0.013		
	Mannitol	0.051	0.015	0.0058	0.0058		
	Sorbitol	0.18	0.071	0.026	0.027		
5C	Arabinose	3.5	2.1	1.1	0.49	2.8	1.1
5C	Lyxose	5.7	3.2	1.9	0.84	4.5	1.2
5C	Ribose	31	15	11	2.9	20	9
5C	D-Xylose	3.4	1.9	0.98	0.51	2.8	1.0
5C	L-Xylose	3.3	1.9	0.94		2.7	
5C	Ribulose			31			
6C	Galactose	0.060	0.024	0.011	0.0086		
6C	Glucose§	0.047	0.017	0.0071	0.0050		
6C	Mannose	0.10	0.035	0.018	0.014		
6C	Fructose§	1.1	0.68				
6C	L-Sorbose	0.88	0.56				

Values reported in 10^{-8} cm/s. Chiral solutes are D-enantiomers unless otherwise noted.
*Experiments performed at 23°C in 1 buffer, 0.5 M solute.
†Conditions as above but with 0.1 M solute.
‡Conditions as above, values are relative initial slopes only (D-xylose set to 1.0).
§Literature P_s values across lecithin membranes: glycerol, 540×10^{-8} cm/s (33); erythritol, 75×10^{-8} cm/s (34); glucose, 0.003×10^{-8} cm/s (35); and fructose, 0.04×10^{-8} cm/s (35).

PNAS | April 26, 2005 | vol. 102 | no. 17 | 6005

In general, the greater the number of carbon atoms, the smaller the permeability. However, the table above clearly shows large differences in permeability for sugar isomers with the same number of C atoms, and the difference is not affected by the lipid composition. Ribose has markedly elevated permeability compared to other 5C sugars (as do erythritol and threitol among 4 C sugar alcohols). What is so unique about ribose? 20% of the sugar is in the furanose form. Rate constants for ring opening of

furanoses are elevated, suggesting greater flexibility. The α-furanose anomer is amphiphilic in that one face is hydrophobic and the other hydrophilic. All of these may promote ribose permeability.

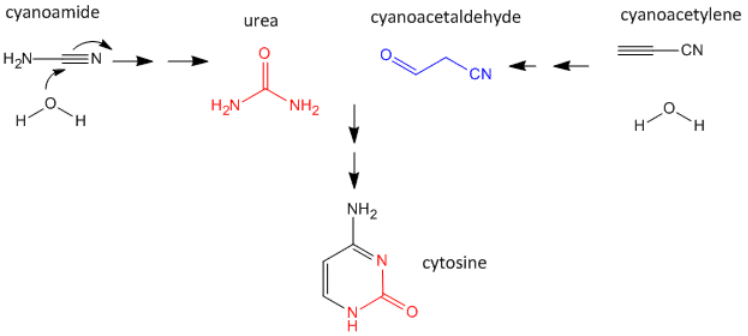


EXTERNAL Jmol: a-D-ribofuranose alpha-D-ribopyranose at

1.9.0.1: Abiotic Synthesis of Nucleobases

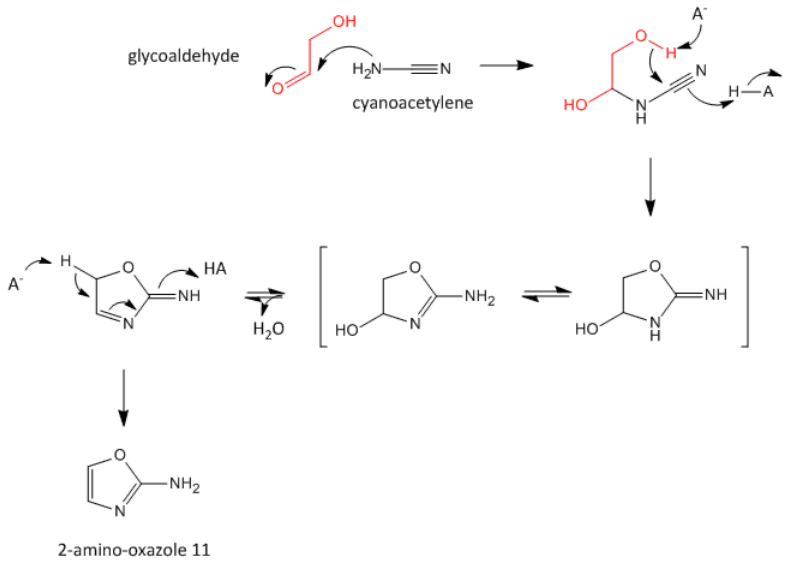
As mentioned above, nucleobases can be made under abiotic conditions with appropriate sources of carbon molecules with nitrogen. These include hydrogen cyanide, cyanoamide (NH2CN) (along with ammonia), hydrogen cyanide, cyanogen (C2N2) and cyanoacetylene (HC3N). Cyanoamide and cyanoacetylene can both react with water to form urea and cyanoacetaldehyde, respectively. The latter two can condense to form cytosine as shown below.

Figure: Abiotic Synthesis of Cytosine



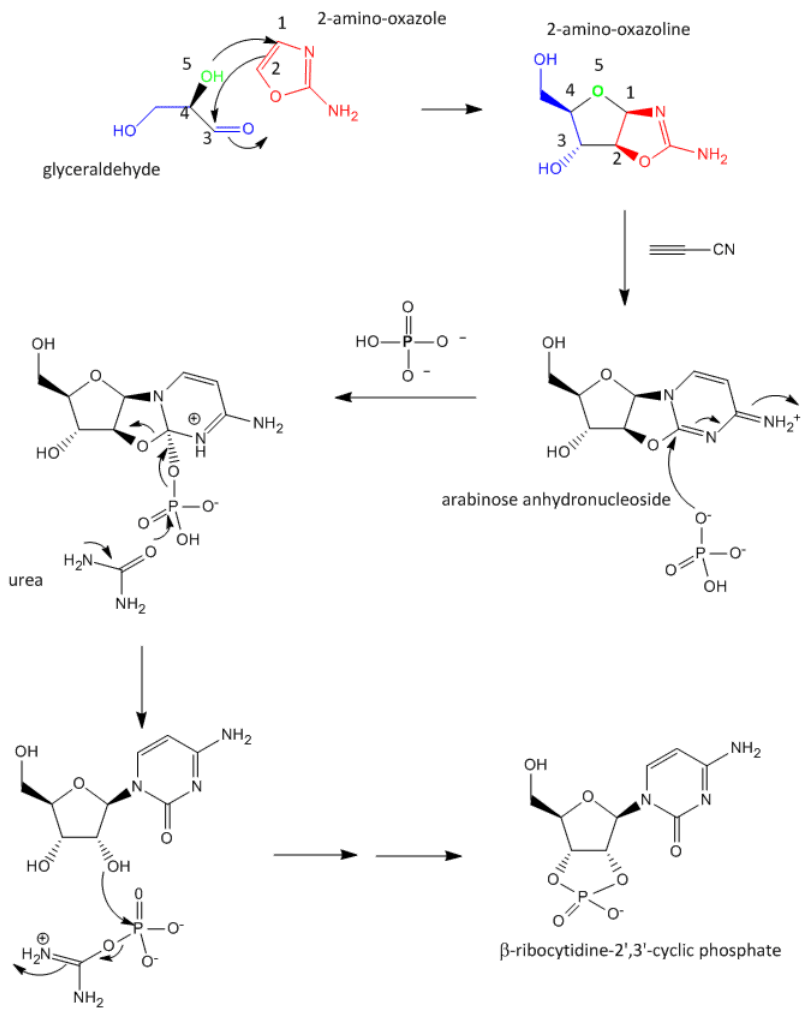
A major problem that has plagued prebiotic researchers is how the product of the carbon-oxygen compound (sugar) links to the product of the carbon-nitrogen compound (nucleobase) to form the nucleoside. Recent research by Powner et al has offered an innovative solution. Instead of forming the sugar and base in separate reaction, and then linking them covalently, the combined molecule could be synthesized in a single set of linked reactions. Inorganic phosphate serves as both a general acid/base catalyst (HA/A- in the figure below) in these new reactions in the formation of an important intermediate, 2-amino-oxazole, and as a nucleophilic catalyst.

Figure: Abiotic Synthesis of 2-amino-oxazole



Glyceraldehyde, 2-amino-oxazole, and inorganic phosphate can react to form a ribocytidine phosphate. Possible reaction mechanisms are shown below.

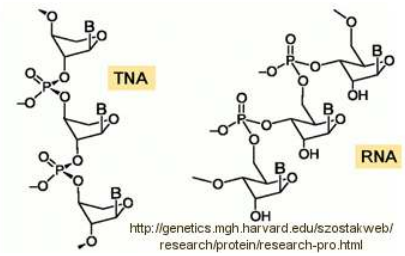
Figure: Abiotic synthesis of a ribocytidine phosphate without condensation of a preformed ribose and cytosine



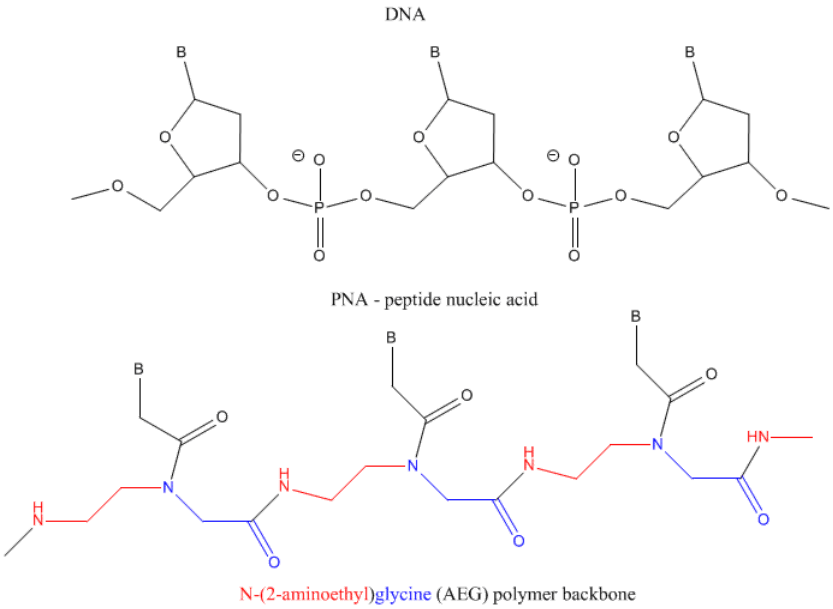
1.9.0.1: Abiotic Synthesis of Genetic Polymers

Abiological synthesis of polymer precursors is a long way from creating genetic polymers like RNA and DNA. These genetic polymers have one property that at first glance seems not conducive to a genetic molecule. Both are polyanions, which must be packed into a cell and folded onto itself to form the classic dsDNA helix and many different RNA structures. This problem is solved to some degree by the presence of counterions that help mask the charge on the negative backbone of the nucleic acids. The presence of phosphate in the phospodiester backbone linkage does confer an important advantage over other possible links (carboxylic acid esters, amides and anhydrides). The electrophilic phosphorous atom is hindered from nucleophilic attack by the negative O attached to the phosphorous. Also, the phosphorous is sp3 hybridized compared to the sp2 hybridization of the electrophilic carbon atom in anhydrides, esters, and amides, and hence is less accessible to nucleophilic attack. Most people now believe that RNA, which can act both as an enzyme and genetic template, preceded DNA as the genetic carrier. The evolution of DNA as the primary genetic carrier required an enzyme to convert ribose to deoxyribose. This would make the nucleic acid less likely to cleave at the phosphodiester bond with the replacement of a nucleophilic 2' OH with an H, and make the genetic molecule more stable. Other types of genetic carriers might have preceded the RNA world, especially if the monomer required could be more readily synthesized from abiological sources. One such alternative are threose nucleic acids (TNA). Synthetics ssTNA can base pair with either RNA, DNA, or itself to form duplexes.

EXTERNAL http://genetics.mgh.harvard.edu/szostakweb/research/protein/research-pro.html



Other possible candidate include peptide nucleic acids (PNA). These can also form double stranded structures with DNA, RNA, or PNA single strands. They were initially designed to bind to dsDNA in the major grove forming a triple-stranded structure. Binding could alter DNA activity, possibly by inhibiting transcription, for example. The structure of a single-stranded PNA is shown. Note that the backbone, a polymer of N-(2-aminoethyl)glycine (AEG) which can be made in prebiotic soups, is not charged, making it easier to bind to dsDNA. AEG polymerizes at 100oC to form the backbone.

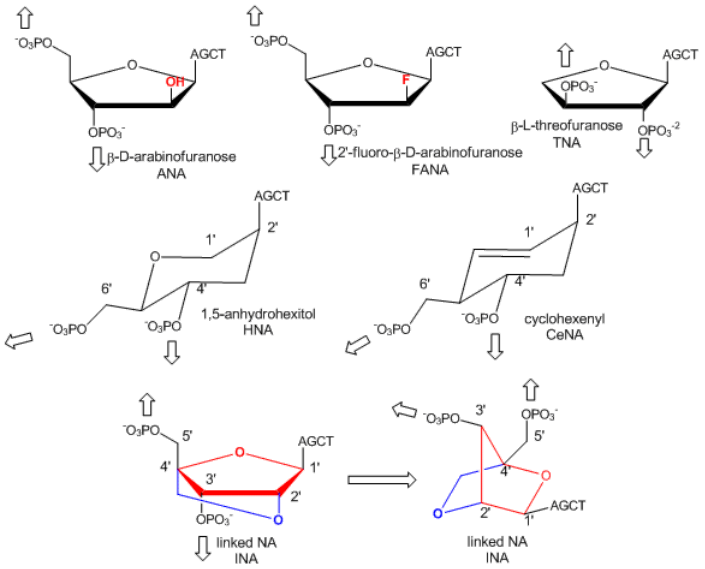


after Gamesh, K. & Nielsen, P. Peptide Nucleic Acids: Analogs and Derivatives. Current Organic Chemistry. 4, 931 (2000)

In addition to changing the backbone, additional bases other than A, C, T, G, and U can be accommodated into dsDNA and ssRNA molecules (Brenner, 2004)

In a recent extension, Pinheiro et al have shown that 6 different foreign backbone architectures can produce xeno-nucleic acids (XNAs) that can be replicated by engineered polymerases which make XNAs from a complementary DNA strand, and a polymerase that can make a complementary copy of DNA from an XNA. XNAs can also be evolved as aptamers to bind specific target molecules. The investigators replaced the deoxyribose and ribose backbone sugar with xenoanalogs (congeners) including 1,5-anhydrohexitol (HNAs), cyclohexene (CeNA), 2'-O,4'-C-methylene-b-D ribose (locked nucleic acids - LNA), L-arabinose (LNA), 2'-fluoro-L-arabinose (FANAs) and threose (TNAs) as shown in the figure below.

Figure: Xeno-nucleic acid sugar congeners



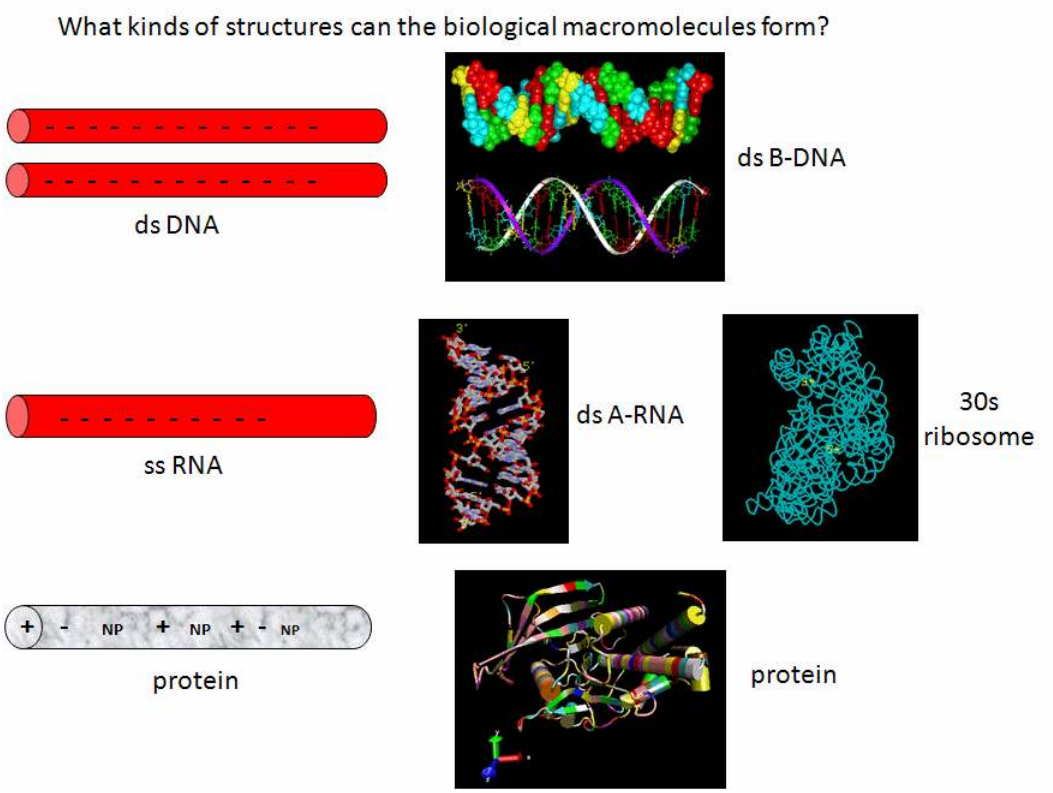
Polymers of these XNA can bind to complementary RNA and DNA and as such act as nuclease-resistant inhibitors of translation and transcription.

Von Kiedrowski, in an experiment similar to the self-replication of peptides described above, has shown that a single stranded 14 mer DNA strand, when immobilized on a surface, can serve as a template for the binding of complementary 7 mers and their conversion to 14 mers. When released by base, this process can occur with exponential growth of the complementary 14 mers. (von Kiedrowski Nature, 396, Nov 1998). Ferris has shown that if the clay montmorillonite is added to an aqueous solution of diadensosine pyrophosphate, polymerization occurs to produce 10 mers which are 85% linked in a 5' to 3' direction.

1.9.0.1: Polyanions as Carriers of Genetic Information

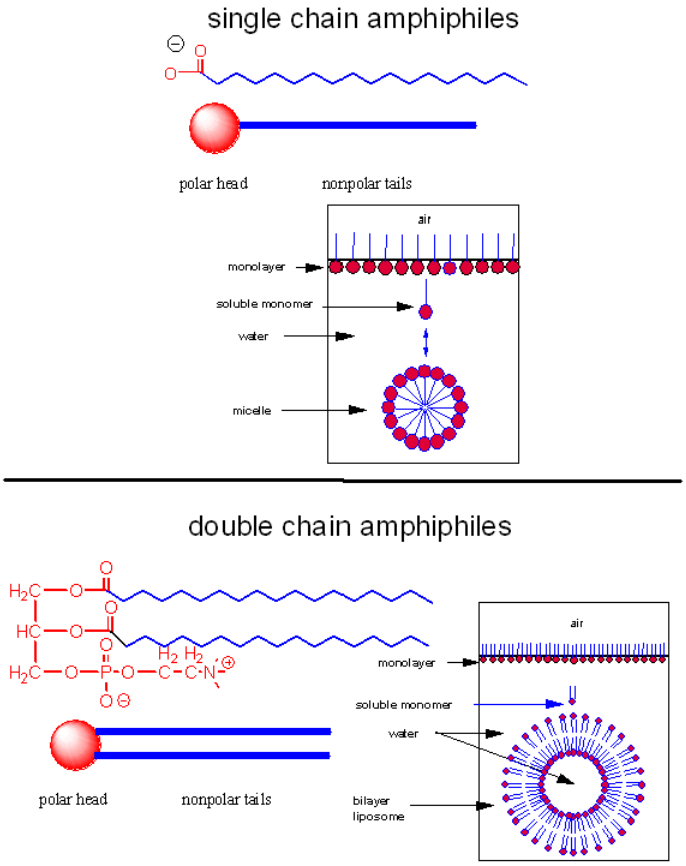
There are other reasons why polyanions are useful genetic molecules, other than their resistance to nucleophilic attack. The biological form of DNA is a large double stranded polyanionic polymer, in contrast to RNA which is a single-stranded polyanion polymer and protein which are polymers with varying combination of anionic, cationic, and hydrophobic properties. Even with counterions, it would be difficult to fold DNA into complicated and compact 3D structures as occurs for proteins, given the large electrostatic repulsions among the charged phosphates. Rather it forms a elongated double stranded rod, not unlike the rod-like structure of proteins denatured with sodium dodecyl sulfate (used in SDS PAGE gels). The elonged rod-shaped structure of ds-DNA is critical for the molecule which is the main carrier of our genetic information since mutations in the bases (leading to a switch in base pairs) causes no change in the overall structure of dsDNA. This enables evolutionary changes in the genetic material to produce new functionalities. A single change an amino acid of a protein, however, can cause a large change in the structure of a whole protein, a feature unacceptable for a carrier of genetic information. RNA structure effectively lies between that of DNA and proteins. Since it has less charge density than dsDNA, it can actually form dsRNA helices, so it can carry genetic information, as well as form complex 3D shapes necessary for its activity as an ribozyme. Perhaps more importantly, steric interference prevents

ribose in RNA from adopting the 2'endo conformation, and allows only the 3'endo form, precluding the occurrences of extended ds-B-RNA helices.

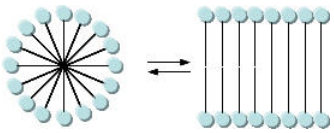


1.9.0.1: The Lipid World

Let's assume that abiological precursors would react to form polymer-like molecules that might be complex enough to fold to structures that would allow binding, catalysis, and rudimentary replication. All this would be worthless unless they could be sequestered in a small volume which would limit diffusion and increase their local concentration. What is required is a membrane structure. Amphiphilic molecules, like lipids, with which we started this book, would be prime candidates since they spontaneously assembly to form micelles and bilayers, as shown in the review diagram below.



As mentioned in Chapter 1, alternative lipid phases are possible. Bilayers can also be formed from single chain amphiphiles, such as certain fatty acids, as illustrated in the equilibrium shown below. This occurs more readily at pH values close to the the pKa of the fatty acid, at which the fatty acids are not all deprotonated with full maximal negative charges. Single chain amphiphiles like fatty acids, which were more likely to formed in abiotic conditions, have been found in meteorites.



Clay surfaces, which have been shown to facilitate the formation of nucleic acids polymers, can also promote the conversion of fatty acid micelles to bilayers (Szostak). One such clay surface, montmorillonite, whose structure is shown below, promotes bilayer formation. It has an empirical formula of Na0.2Ca0.1Al2Si4O10(OH)2(H2O)10.

EXTERNAL Chime Molecule Modeling: montmorillonite EXTERNAL: montmorillonite

The effect of montmorillonite on vesicle formation can be shown by simple measurements of turbidity with time. Microscopy of fluorophore-encapsulated vesicles also shows encapsulated montmorillonite. The fatty acids presumably absorb to the cation layer of the clay particles. Time studies using light scattering also indicate that the vesicles grow in the presence of fatty acid micelles. To differentiate between the formation of new vesicles and the increase in size of pre-existing vesicles (which couldn't be done by simple light scattering without separation of the vesicles), investigators used two different fluorescent molecules to label fatty acid vesicles. The two probes were selected such that if the two probes came in close contact, energy transfer from the excited state of one fluorophore to the other fluorophore could occur, an example of fluorescence resonance energy transfer (FRET). FRET is

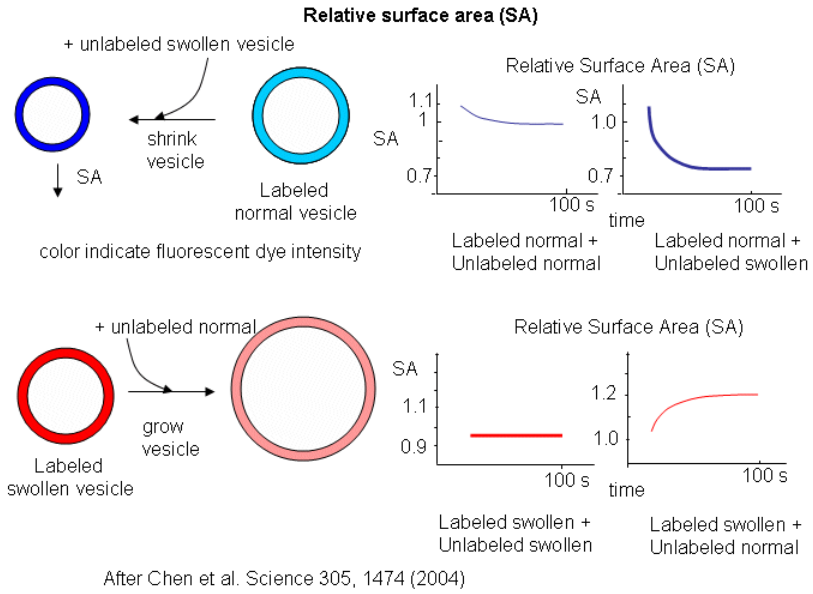
observed when emission of the second dye occurs after excitation of the first dye, at a wavelength outside of the excitation wavelength of the second dye. If unlabeled vesicles were added to either labeled vesicles, no changes in FRET were observed, suggesting that the dyes did not move between vesicles. If fatty acid micelles were added, a decrease in FRET was observed, suggesting that new fatty acids were transferred to the doubly-labeled vesicles, effectively diluting the dye concentrations in the bilayer and their relative proximity, both which would decrease FRET. Most of the new fatty acid was incorporated into pre-existing vesicles which grew. The vesicles could also divide if extruded through a small pore. Later we will see that the energy to grow the vesicles can derive in part from a transmembrane proton concentration collapse. Division of vesicles might be promoted by bilayer assymetries associated with addition of substances to the outer leaflet, causing membrane distortion.

1.9.0.1: Protocells

At some point, early genetic material must have been encapsulated in a membranous vesicles. Would new properties emerge from this mixture that might have a competitive (evolutionary) advantage over either component alone, and thus be a step on the way to the formation of a "living" cell? The answer appears to be yes. Chen et al. have incorporated RNA into fatty acid vesicles with interest effects. They asked the question as whether those vesicles could grow at the expense of vesicles without encapsulated RNA. RNA, with a high charge density and its associated counter ions would create osmotic stress on the vesicles membranes. To relieve that stress they could acquire fatty acids from other fatty acid vesicles (or fatty acid micelles), increasing their surface area, and concomitantly reducing tension in the membrane.

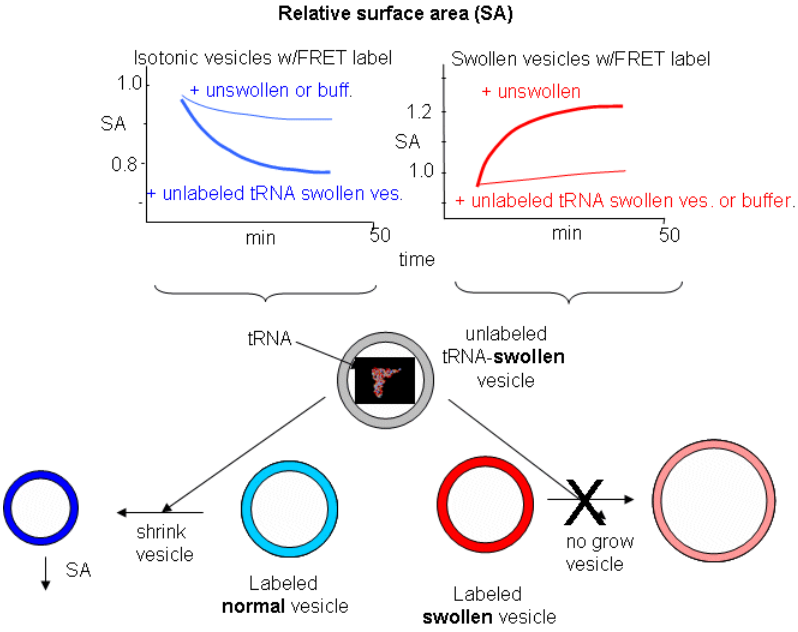
Oleic acids vesicles were first placed under stress by encapsulating 1 M sucrose in the vesicle and then diluting it in hypotonic media. Water would enter and swell the vesicle (but without bursting and resealing, as evident from control experiments). Then they prepared stressed and unstressed oleic acid vesicles in the presence of two nonpolar flurophores, NBD-PE (excitation at 430 nm, emission at 530 nm) and Rh-DHPE (emission at 586 nm). These fluorophores were chosen for fluorescence resonance energy transfer measurements. If the membrane vesicles changed size, the FRET signal would change, based on the relative concentration and proximity of the dual fluorophores. If the separation between probe molecules increased, the FRET signal would decrease. Conversely, if the vesicle shrunk, the FRET signal would increase.

The results showing the effect of adding unlabeled swollen vesicles to labeled normal vesicles, and labeled swollen vesicles to unlabeled normal are shown below. The surface area of normal labeled vesicles decreased by about 25% when unlabeled swollen vesicles were added, but not when unlabeled normal vesicles were added. Labeled swollen vesicles increased 25% in size only if mixed with unlabeled normal vesicles, not with unlabeled swollen vesicles. Hence swollen vesicles win the competition and "steal" lipid from normal vesicles.



After Chen et al. Science 305, 1474 (2004)

Now what about vesicles swollen with encapsulated RNA? RNA, with its associated charge and charged counter ions also placed an osmotic stress on the vesicles. FRET labels (the two fluorophores) were place in vesicles without RNA. Fatty acids were removed from isotonic labeled vesicles in the presence of unlabeled tRNA swollen vesicles (left panel below). Labeled vesicles swollen with glycerol took fatty acids from unswollen vesicles (without tRNA), but not from tRNA swollen vesicles, as both were swollen so no net drive to reduce swelling by lipid exchange was present.

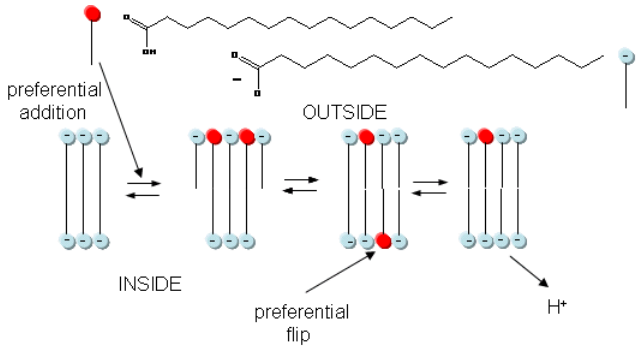


After Chen et al. Science 305, 1474 (2004)

These results show the vesicles with encapsulated RNA have a competitive (evolutionary) advantage over normal vesicles. This data suggests that having a polyanion as the source of genetic material is actually advantageous to the protocell. In addition the move in modern membranes to phospholipids with esterified fatty acids (instead of free ones) may actually have stabilized membranes, given the movement of free fatty acids to different membranes.

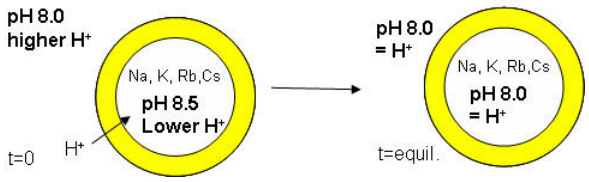
1.9.0.1: Energy Transduction in Protocells

In addition to a genetic macromolecule and a semipermeable membrane, a source of energy to drive intracellular processes must be present. A common source of free energy used in many cells to drive unfavorable processes is a proton gradient, whose formation in modern cells can be coupled to energy input from oxidation, ATP cleavage, light, or the collapse of another gradient. Could a proton gradient be formed in protocells? It can, quite easily, when coupled to the growth of fatty acid vesicles. If a fatty acid vesicle is to grow, more fatty acid must be added to the outer leaflet. The protonated, uncharged form of the fatty acid would preferentially be added, since it would lead to less electrostatic repulsion between adjacent head groups. The protonated, uncharged form of the fatty acid would also be most likely to flip to the inner leaflet to minimize stress asymmetries in the leaflets. Once in the inner leaflet, it could deprotonate to form H+(aq) in the inside of the membrane, creating a transmembrane proton gradient and transmembrane potential. The energy released on growth of the membrane is partly captured in the formation of a proton gradient, as shown in the figure below.



Generation of proton gradient on growth of fatty acid vesicle

The proton gradient would soon inhibit its own formation since further movement of protons into the cell would be attenuated by the positive transmembrane potential unless metal ions inside moved outside. In addition, the gradient would collapse after growth stopped. The investigators made fatty acids vesicles in the presence of pH 8.5 buffers whose pH was adjusted with an alkali metal hydroxide. The external pH was reduced to 8.0, resulting in a 0.5 pH unit proton concentration gradient. (Changes in intravesicular pH were measured with pH-sensitive fluorophore, HPTS.) Inward movement of protons down a concentration gradient, as shown in the figure below, would occur with time, collapsing the imposed concentration gradient.



With fatty acid vesicles, this artificial pH gradient collapsed quickly, suggesting the vesicle permeability to protons was high. The rate was too high for simple flip-flop diffusion. Inward movement of protons appeared to be facilitated by outward movement of the M+ ions. The rate of decay of the proton gradient was exponential, and the resulting first order rate constant was easily determined. A graph of the rate constant for pH gradient collapse vs unsolvated ionic radius of M+ decreased with increasing radius (i.e. kNa > kK > KRb > KCs, suggesting that the pH gradient would be more stable if large, impermeable or otherwise trapped cations were encapsulated. When vesicles were made with encapsulated Arg+, the imposed pH gradient did not collapse for hours. If oleic acid micelles were added to oleic acid vesicles with encapsulated Arg+, with no artificial pH gradient induced across the membrane, the vesicle grew with concomitant movement of protons into the vesicle, producing a pH gradient of 0.3 within seconds.

These experiments show that membrane growth and energy storage could be coupled, and the right composition of encapsulated material could lead to a stable transmembrane pH gradient, a source of energy to drive biological processes. It even suggests that a charge polyanion would be beneficial as a genetic carrier.

1.9.0.1: Hydrothermal Vents or Primoridal Soup?

The case for the origin of life in deep sea hydrothermal vents and not in a primordial "Campbell's" soup has been argued convincingly by Lane et al (2010). What's needed for life are reasonably complex molecules and an energy source to drive unfavorable reactions. It's the latter on which that Lane et al focus. In an early primordial world that was low in oxygen, exergonic oxidation reaction of organic molecules would provide little energy. This can be surmised from the low energy yield (compared to aerobic respiration) achieved in present day glycolytic (fermentative) pathways from all major domains of life, archaea, bacteria, and eukaryotes.

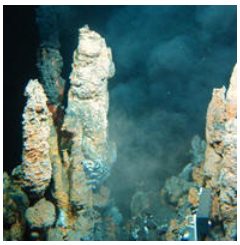
Background: Based on rRNA sequences, a primordial cell evolved into two different types of cells, one that became bacteria, and another that split further into archaea (single cells, similar to bacteria) and eukaryotic cells (complex cells with internal organelles that eventually formed multicellular organisms). Bacterial and archaea are collected called prokaroytic cells.

In addition, these pathways required the evolution of up to dozen different enzymes to produce their relatively meager energy yield which ultimately depends on the oxidation of an organic molecule by another organic molecule instead of by a powerful oxidant like dioxygen. An anisotropic arrangement of molecules in a concentrated soup could lead to transient chemical potential fluctuations but these would be inefficient and impermanent sources of energy. Effectively the primordial soup would be at equilibrium and hardly expected to provide the energy for synthesis of RNA enzymes and replicators. UV light leads to photo-damage and photolysis not replication of complex molecules. What is needed is a way to drive the synthesis of molecule with high chemical potential energy (like sulfur esters and phosphoanhydrides) compared to their lytic products. These could then provide an energy sources to drive ATP synthesis, for example.

A detailed look a the bioworld shows that the earliest organisms used energy from the collapse of the proton gradient (chemisomotic principle elucidated by Peter Mitchell). All present autotrophs (organisms that can fix CO2 and form complex organic molecules) and many heterotrophs (use complex organic molecules of other organisms for fuel) use redox complexes in membranes coupled to membrane gradients. These complexes would take reduced molecules and pass electrons from them to oxidizing agents (electron acceptors), including O2, CO2, and Fe3+ to form H2O, CH4, and Fe2+. Fermentors also use ATPase membrane enzymes to transport nutrients. Yet genomic analysis of bacteria and archaea show that enzymes involved in fermentation differ significantly, suggesting that they evolved separately towards a convergent function. Structure in common include DNA, RNA, ribosomes and membrane ATPases, which Lane et al suggest were in a the Last Universal Common Cell (LUCA).

All autotrophs produce their energy source by fixing CO2 using either H2 directly or indirectly using H2O and H2S. All of the are available in nonhydrothermal deep sea vents. Volcanic vents, however, are extremely hot (not optimal for organic molecule synthesis), very acidic, and lack hydrogen gas. A different type of nonvolcanic vent, an alkaline hydrothermal one, might produce more conducive as a site of the origin of life. In these vents, water chemically reacts with minerals in the crust (such as olivine) leading to their hydroxylation and subsequent fracture, with promotes more water entry into the crust. It has been reported that there is more water found as hydroxylated minerals in the crust, that there is liquid water in the oceans. These processes result in temperatures up to 200 degrees Celsius and release of hydrogen gas into a moderately alkaline vents into the sea water at temperatures more conducive (70 degrees C) to the origin of life.

Figure: Alkaline Vent



1.9.0.1: Fixing CO_2

Of the five different pathways known to fix CO_2, all require ATP except one. That one is present in both methanogens, which produce methane from CO_2 and H_2, and in the acetogens, which produce acetate ($CH_3CO_2^-$) in the form of acetyl-CoA. The simpler reactions of forming acetic acid and methane are shown below:

$2CO_2 + 4H_2 \rightarrow CH_3CO_2H + 2H_2O$.

$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$.

The ΔG^0 values for these reaction (calculated using ΔG^0_f for gas phase H_2, CO_2 and CH_4 and liquid acetic acid and water are -75 and -131 kJ, respectively, at 250C, showing that they are thermodynamically favored. Making AcetylCoA, a "high" energy molecule compared to its hydrolysis products (as is ATP) from acetic acid and CoASH, requires an additional source of energy to drive the reaction. A proton gradient is the likely source.

Some bacteria and Achaea cells (primordial or present) use the reductive acetyl-CoA pathway, also known as the Wood-Ljungdahl pathway, to form, in a noncyclic process, acetyl CoA from CO_2 and at the same time makes ATP. This process is paid for by a

proton gradient. This has been described by Shock as "a free lunch you get paid to eat". The energetics of the present acetyl CoA pathway based on the overall reaction below show an approximate ΔG^0 value of -59 kJ/mol which can drive ATP synthesis.

$2CO_2 + 4H_2 + CoASH \rightarrow CH_3COSCoA + 3H_2O$.

The concentration of carbon dioxide in the primordial ocean was 1000 times higher than now. Vents produced large amounts of methane and hydrogen gas. There was little oxygen and hence lots of Fe^{2+}. The enzymes involved in this acetyl-CoA pathway of carbon fixation have FeS clusters. It has also been shown that bubbles (which are really membrane-bound spaces) of FeS and NiS can be made in deep-sea vents. These could not only encapsulate precursor molecules but also serve as catalysts. Vents also can catalyze the fixation of nitrogen (to ammonia) and laboratory studies show that FeS can catalyze the conversion of formate (found in vents) into pyrimidines and purines. The studies of present methanogens and vent chemistry suggest that the critical ingredients and conditions for the development of the first biological cells probably occurred in the vents.

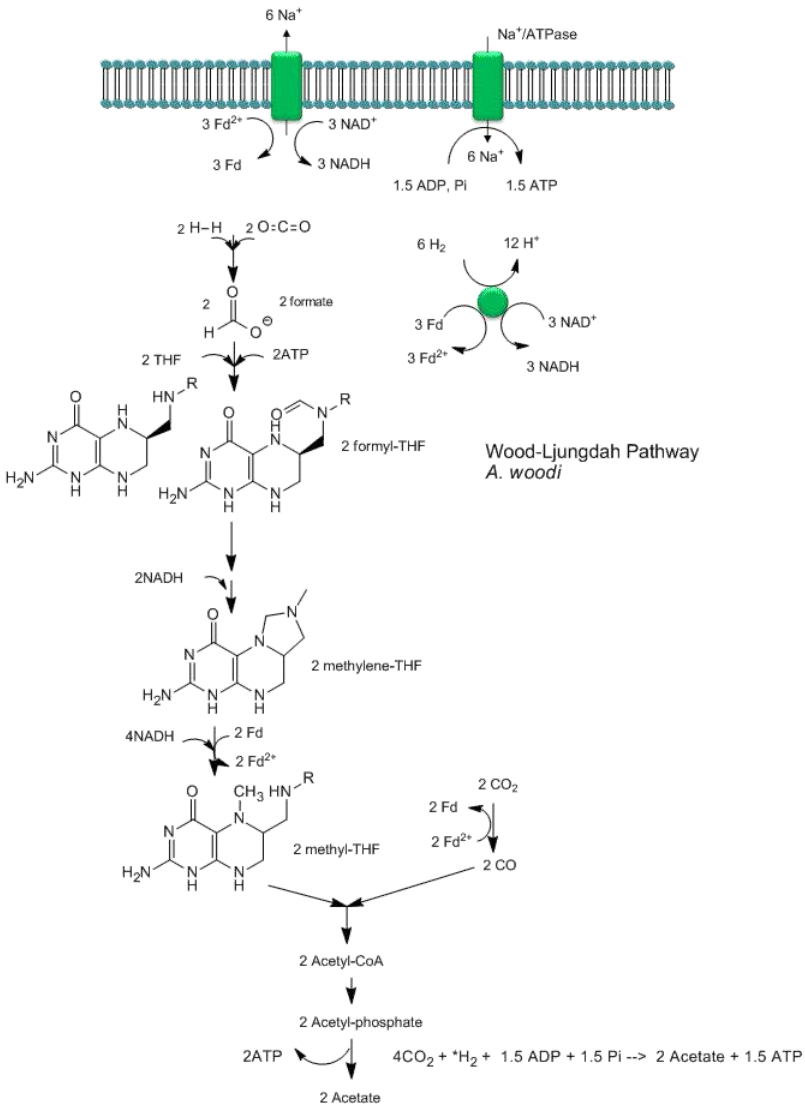
To produce polymers, an energy source and monomers must exist. Concentration gradients found in simulations of vents produce million-fold concentrated molecules. The transient heating and cooling of any double-stranded nucleic acids could lead to concentration amplification by a PCR-like strand separation followed by reannealing. In addition, these vent regions possess a powerful, reoccurring energy source, a pH gradient, as the alkaline vented material entered the acidic oceans that exist with high CO2 concentrations, creating a gradient across an inorganic membrane. This is startlingly analogous to the pH gradient across membranes (acidic outside, alkaline inside) driven by the membrane complexes in the mitochondria and bacteria. Lane et al argue that the existence of membrane proton gradients as an energy source in all cells (eukaryotes, bacteria, and archaea) and in chloroplasts and mitochondria, corroborate their hypothesis. Bacteria and Archaea share homologous ATPases and electron carriers (ferredoxins, quinones, and cytochromes). These similarities contrast to the differences in enzyme structures in fermentative pathways. Arguments that proton pumps evolved to pump proteins (and reduce pH gradients) can't explain their ubiquitous presence even in organisms not subjected to low pH. Hence the ubiquity of proton pumps supports the conjecture that they arose from the first protocells, possibly comprised of inorganic walls and ultimately with amphiphilic molecules synthesized from precursors.

Creationists would argue that it would be impossible to evolve a structure with the complexity of membrane ATPase (which collapse a pH gradient as the power the synthesis of molecules with large negative ΔG^0 of hydrolysis). Lane et al propose that the earliest cells evolved ATPase-like molecules in alkaline vents where pH gradients analogous to those in cells today arose. They envision cell-like columns lined by FeS membrane-like structures with alkaline conditions inside and acid conditions outside. Nonpolar or amphiphilic molecules would line the inside of the cells/columns. An ATPase-like system could then take advantage of the pH gradient which constantly replenishes itself. If structures as complicated as ribosomes evolved from a subsequent RNA world, surely ATPase-like molecules could also. Other chemistry might have evolved earlier to utilize the energy source provided by the pH gradient.

If life originated in the vents, it would need an energy source to leave the vents. Presumably it would have evolved one to utilized pH gradient to replace the one it left in the alkaline vents. The substrate level phosphorylation of glycolysis that requires ATP input to make ATP would not provide the energy source needed. Cells that left would have had to produce their own proton gradient. Perhaps all the was needed initially was concerted conformational changes in proteins that upon exposure of a different pH changed their shape inducing pKa shifts in adjacent proton donors/acceptors leading to vectorial discharge of protons across a membrane. Perhaps the method described above in protocells was sufficient.

Recent analyses by Poehlein et a show that CO2 reduction (fixation) can be coupled to the production of a sodium ion gradient, which could collapse to drive ATP synthesis. Analysis of the genome of a gram positive bacteria, Acetobacterium woodii, an acetogen, shows the it has an ancient pathway for production of acetyl-CoA that can, in an anabolic fashion form biomass or in a catabolic fashion be cleaved to acetate with the production of ATP. It does not require classic electron carriers like ubiquinone or cytochrome C linked to protein gradient formation to drive ATP synthesis. Rather it has only a ferredoxin:NAD+ oxioreductase which couples oxidation to the formation of a sodium ion gradient, which collapses through an sodium ion transporter/ATP synthase to drive ATP synthase. A plausible reaction scheme based on genomic analysis is shown below:

Figure: Acetyl-CoA Synthase and Acetogenesis



1.9.0.1: The Role of Fe/S Centers

Let's return to the chicken and egg dilemma one more time. What is needed for biological polymer formation are monomeric precursors, an energy source, and a way to compartmentalized them all. We discuss how monomeric precursors could form, but wouldn't it be far better if even the synthesis of precursors could be catalyzed? One source of catalysis mostly absent from the "bioorganic" abiotic chemistry in the above discussion is the transition metals. Transition metals can form complexes. Ligands containing lone pairs on O, N, and S atoms can donate them to transition metals ions, which can hold up to 18 electrons in s, p, and d orbitals. Hence as many as 9 lone pairs on ligand molecules (which are often multidentate) could be accommodated around the transition metal ion. Many present small molecule metabolites and their abiotic precursors (H2O, CO, CO2, NH3 and thiols) bind cations as mono- or polydentate donors of electrons. Hence transition metal ions would have a thermodynamic tendencies to be bound in complexes.

Bound ligands that contain potentially ionizable hydrogens could become deprotonated and made better nucleophiles for reactions. Hence the transition state metal ion, acting with the complex, becomes a catalyst as it decreases the pKa of a bound ligand (such as water). In addition, since transition metals ions can have multiple charge and oxidation states, they can easily act as redox centers in the oxidation/reduction of bound ligands that were redox active. Given the relative anoxic conditions of the early oceans, Fe2+ would predominant. It could easily be oxidized to Fe3+ as it reduced a bound ligand. Highly charged transition states would withdraw electron density from bound ligands leading to their possible oxidation.

Metals obviously still play a strong role in catalysis, both indirectly in promoting correct protein folding and directly in stabilizing charge in both the transition state and intermediates in chemical reaction pathways. FeS clusters are of significant importance. Their biosynthesis involves removal by an active site Cys in a desulfurase enzyme of a sulfur from a free amino acid Cys followed by its transfer to an Fe in a growing FeS cluster in a FeS scaffold protein, which then transfers the cluster to an acceptor protein where it acts as a cofactor. FeS clusters can adopt a variety of stoichiometries and shapes, as well as redox states for the participating Fe ions. The continuing importance of FeS clusters in all cells, their involvement in not only redox enzymes in which electron transfer is facilitated by delocalization of electrons over both Fe and S centers, but also in coupled electron/proton transport in mitochondrial electron transport, Fe storage (ferrodoxins), and in regulation of enzyme activity and gene expression, suggests that they were of primordial importance in the evolution of life. T

hey are often found at substrate binding sites of FeS enzymes involved in both redox and nonredox catalysis. A ligand can bind to a particular Fe in the cluster, activating it for hydration or dehydrogenation reactions. Fe 4 of the FeS cluster in the TCA enzyme aconitase can have a coordination numbers of 4, 5, or 6 as it binds water, hydroxide or substrate. It acts to both decrease electron density in the transition state and to change the pKa of bound water as the enzyme catalyzes an isomerization of tricarboxylic acids (citric and isocitric acid) through an elimination/addition reaction with water. In another example it can bind S-adenosylmethionine through its amine and carboxylate groups, which activates the molecule for cleavage and radical formation. In some cases metals other than Fe (Ni for example) are incorporated into the cluster. FeS effects on transcription factors involves facilitation of optimal structure for DNA binding. FeS and FeNi centers in proteins are similar in structure tp FeS units in minerals like greigite and presumably to FeS structure formed when H2S and S2- react with Fe2+ (present in abundance in the early ocean) and other metals in vents Metal sulfides participate in reduction of both CO and CO2. For example the synthesis of CH3SH from CO2 and H2S is catalyzed by "inorganic" FeS.

1.9.0.1: The Minimal Genome

This question is being addressed by eliminating "unnecessary" gene from simple bacteria. Cells placed in a rich nutrient broth with essential lipids, vitamins, and amino acids would need fewer genes than those placed in a more nutrient-poor medium. Bacteria cells like Mycoplasma genetalium, that live within "nutrient rich" eukaryotic cell, have been genetically manipulated to delete unnecessary genes. Based on knockout studies, it may be possible for the cell to survive with only 300-350 genes. Bacillus subtilis has approximately 4100 genes. Estimates have been made that it could survive with as few as 271 genes.

more to be added

This page titled 1.9: Abiotic Origins of Life is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **30.1: Abiotic Origins of Life** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

2: Chemistry of Water

2.1: Water
2.2: The multiple roles of water
2.3: pH and Buffers
2.4: Weak Acids and Bases, pH and pKa
2.5: Buffering against pH Changes in Biological Systems
2.6: Solubility in an aqueous world - noncovalent interactions in depth
2.7: Solubility in an aqueous world - The Hydrophobic Effect
2.8: Chapter 2 Questions

2: Chemistry of Water is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.1: Water

Do you ever wonder why scientists spend time looking for water on other planets? It is because water is essential to life; even minute traces of it on another planet can indicate that life could or did exist on that planet. Water is one of the more abundant molecules in living cells and the one most critical to life as we know it. Approximately 60–70 percent of your body is made up of water. Without it, life simply would not exist.

Water Is Polar

The hydrogen and oxygen atoms within water molecules form polar covalent bonds. The shared electrons spend more time associated with the oxygen atom than they do with hydrogen atoms. There is no overall charge to a water molecule, but there is a slight positive charge on each hydrogen atom and a slight negative charge on the oxygen atom. Because of these charges, the slightly positive hydrogen atoms repel each other and form the unique shape seen in Figure 2.1.6. Each water molecule attracts other water molecules because of the positive and negative charges in the different parts of the molecule. Water also attracts other polar molecules (such as sugars), forming hydrogen bonds. When a substance readily forms hydrogen bonds with water, it can dissolve in water and is referred to as hydrophilic ("water-loving"). Hydrogen bonds are not readily formed with nonpolar substances like oils and fats (*Figure* **2.1.1**). These nonpolar compounds are hydrophobic ("water-fearing") and will not dissolve in water.

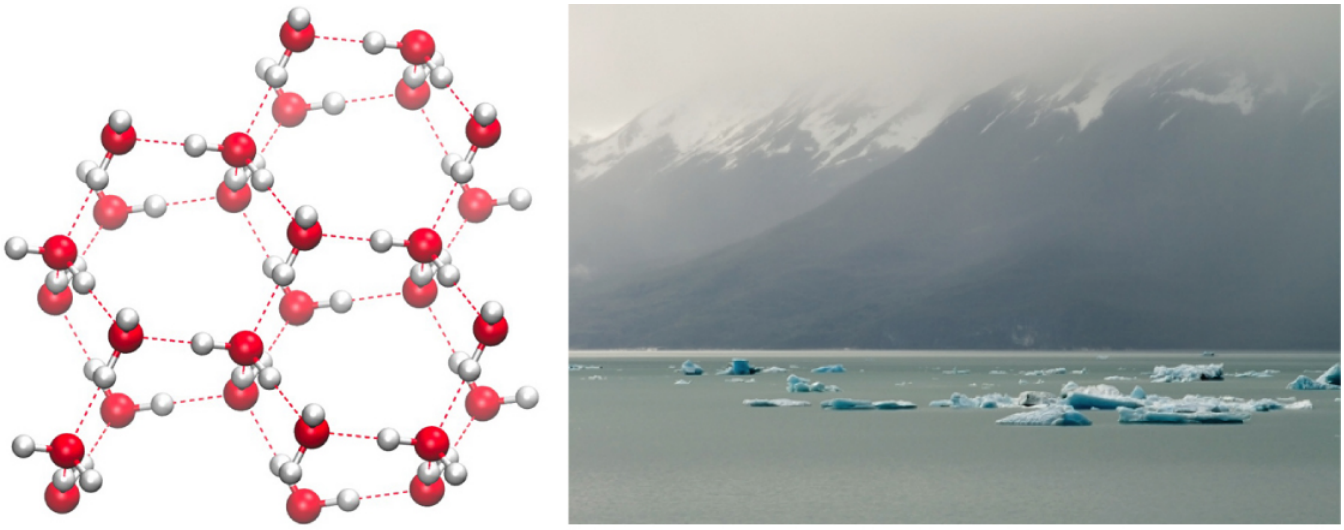
Figure **2.1.1**: As this macroscopic image of oil and water show, oil is a nonpolar compound and, hence, will not dissolve in water. Oil and water do not mix. (credit: Gautam Dogra)

Water Stabilizes Temperature

The hydrogen bonds in water allow it to absorb and release heat energy more slowly than many other substances. Temperature is a measure of the motion (kinetic energy) of molecules. As the motion increases, energy is higher and thus temperature is higher. Water absorbs a great deal of energy before its temperature rises. Increased energy disrupts the hydrogen bonds between water molecules. Because these bonds can be created and disrupted rapidly, water absorbs an increase in energy and temperature changes only minimally. This means that water moderates temperature changes within organisms and in their environments. As energy input continues, the balance between hydrogen-bond formation and destruction swings toward the destruction side. More bonds are broken than are formed. This process results in the release of individual water molecules at the surface of the liquid (such as a body of water, the leaves of a plant, or the skin of an organism) in a process called evaporation. Evaporation of sweat, which is 90 percent water, allows for cooling of an organism, because breaking hydrogen bonds requires an input of energy and takes heat away from the body.

Conversely, as molecular motion decreases and temperatures drop, less energy is present to break the hydrogen bonds between water molecules. These bonds remain intact and begin to form a rigid, lattice-like structure (e.g., ice) (Figure **2.1.2a**). When frozen,

ice is less dense than liquid water (the molecules are farther apart). This means that ice floats on the surface of a body of water (Figure 2.1.2b). In lakes, ponds, and oceans, ice will form on the surface of the water, creating an insulating barrier to protect the animal and plant life beneath from freezing in the water. If this did not happen, plants and animals living in water would freeze in a block of ice and could not move freely, making life in cold temperatures difficult or impossible.



(a) (b)

Figure 2.1.2: (a) The lattice structure of ice makes it less dense than the freely flowing molecules of liquid water. Ice's lower density enables it to (b) float on water. (credit a: modification of work by Jane Whitney; credit b: modification of work by Carlos Ponte)

CONCEPTS IN ACTION

Click here to see a 3-D animation of the structure of an ice lattice. (credit: image created by Jane Whitney using Visual Molecular Dynamics (VMD) software[1])

Water Is an Excellent Solvent

Because water is polar, with slight positive and negative charges, ionic compounds and polar molecules can readily dissolve in it. Water is, therefore, what is referred to as a solvent—a substance capable of dissolving another substance. The charged particles will form hydrogen bonds with a surrounding layer of water molecules. This is referred to as a sphere of hydration and serves to keep the particles separated or dispersed in the water. In the case of table salt (NaCl) mixed in water (Figure 2.1.3), the sodium and chloride ions separate, or dissociate, in the water, and spheres of hydration are formed around the ions. A positively charged sodium ion is surrounded by the partially negative charges of oxygen atoms in water molecules. A negatively charged chloride ion is surrounded by the partially positive charges of hydrogen atoms in water molecules. These spheres of hydration are also referred to as hydration shells. The polarity of the water molecule makes it an effective solvent and is important in its many roles in living systems.

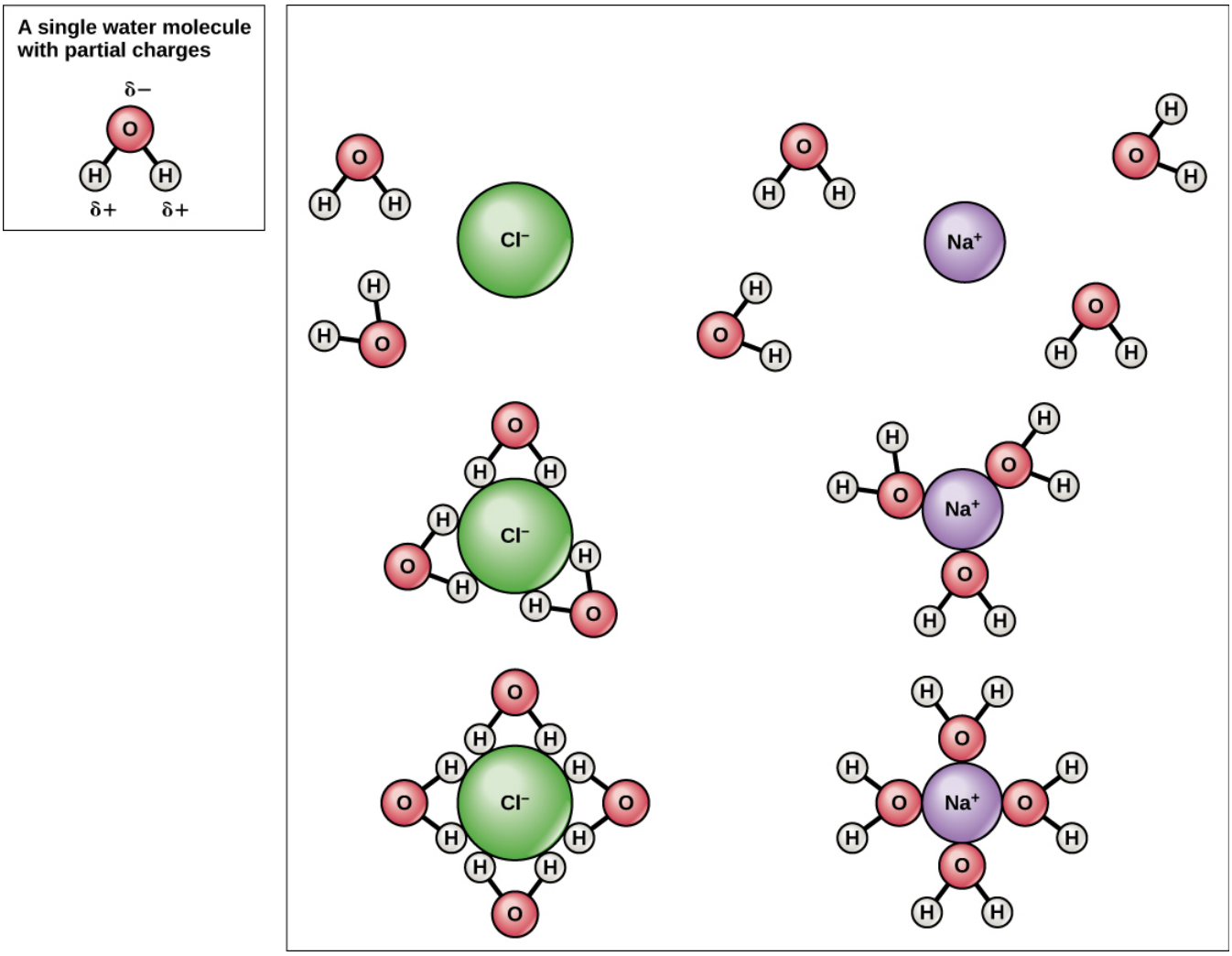


Figure 2.1.3: When table salt (NaCl) is mixed in water, spheres of hydration form around the ions.

Water Is Cohesive

Have you ever filled up a glass of water to the very top and then slowly added a few more drops? Before it overflows, the water actually forms a dome-like shape above the rim of the glass. This water can stay above the glass because of the property of cohesion. In cohesion, water molecules are attracted to each other (because of hydrogen bonding), keeping the molecules together at the liquid-air (gas) interface, although there is no more room in the glass. Cohesion gives rise to surface tension, the capacity of a substance to withstand rupture when placed under tension or stress. When you drop a small scrap of paper onto a droplet of water, the paper floats on top of the water droplet, although the object is denser (heavier) than the water. This occurs because of the surface tension that is created by the water molecules. Cohesion and surface tension keep the water molecules intact and the item floating on the top. It is even possible to “float” a steel needle on top of a glass of water if you place it gently, without breaking the surface tension (Figure 2.1.4).

Figure 2.1.4: The weight of a needle on top of water pulls the surface tension downward; at the same time, the surface tension of the water is pulling it up, suspending the needle on the surface of the water and keeping it from sinking. Notice the indentation in the water around the needle. (credit: Cory Zanker)

These cohesive forces are also related to the water's property of adhesion, or the attraction between water molecules and other molecules. This is observed when water "climbs" up a straw placed in a glass of water. You will notice that the water appears to be higher on the sides of the straw than in the middle. This is because the water molecules are attracted to the straw and therefore adhere to it.

Cohesive and adhesive forces are important for sustaining life. For example, because of these forces, water can flow up from the roots to the tops of plants to feed the plant.

CONCEPT IN ACTION

To learn more about water, visit the U.S. Geological Survey Water Science for Schools: All About Water! website.

Buffers, pH, Acids, and Bases

The pH of a solution is a measure of its acidity or alkalinity. You have probably used litmus paper, paper that has been treated with a natural water-soluble dye so it can be used as a pH indicator, to test how much acid or base (alkalinity) exists in a solution. You might have even used some to make sure the water in an outdoor swimming pool is properly treated. In both cases, this pH test measures the amount of hydrogen ions that exists in a given solution. High concentrations of hydrogen ions yield a low pH, whereas low levels of hydrogen ions result in a high pH. The overall concentration of hydrogen ions is inversely related to its pH and can be measured on the pH scale (Figure 2.1.5). Therefore, the more hydrogen ions present, the lower the pH; conversely, the fewer hydrogen ions, the higher the pH.

The pH scale ranges from 0 to 14. A change of one unit on the pH scale represents a change in the concentration of hydrogen ions by a factor of 10, a change in two units represents a change in the concentration of hydrogen ions by a factor of 100. Thus, small changes in pH represent large changes in the concentrations of hydrogen ions. Pure water is neutral. It is neither acidic nor basic, and has a pH of 7.0. Anything below 7.0 (ranging from 0.0 to 6.9) is acidic, and anything above 7.0 (from 7.1 to 14.0) is alkaline. The blood in your veins is slightly alkaline (pH = 7.4). The environment in your stomach is highly acidic (pH = 1 to 2). Orange juice is mildly acidic (pH = approximately 3.5), whereas baking soda is basic (pH = 9.0).

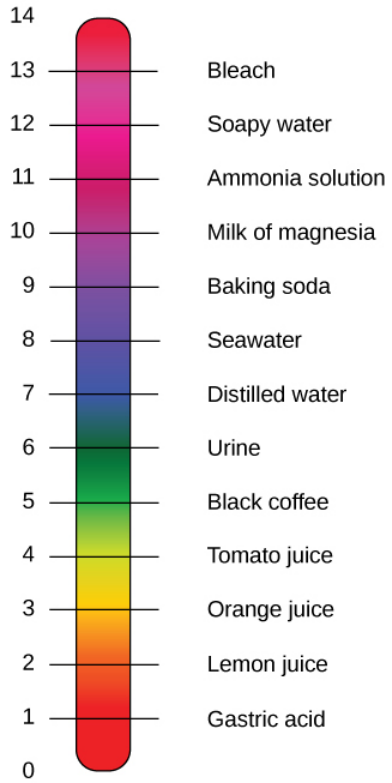


Figure **2.1.5**: The pH scale measures the amount of hydrogen ions (H^+) in a substance. (credit: modification of work by Edward Stevens)

Acids are substances that provide hydrogen ions (H^+) and lower pH, whereas bases provide hydroxide ions (OH^-) and raise pH. The stronger the acid, the more readily it donates H^+. For example, hydrochloric acid and lemon juice are very acidic and readily give up H^+ when added to water. Conversely, bases are those substances that readily donate OH^-. The OH^- ions combine with H^+ to produce water, which raises a substance's pH. Sodium hydroxide and many household cleaners are very alkaline and give up OH^- rapidly when placed in water, thereby raising the pH.

Most cells in our bodies operate within a very narrow window of the pH scale, typically ranging only from 7.2 to 7.6. If the pH of the body is outside of this range, the respiratory system malfunctions, as do other organs in the body. Cells no longer function properly, and proteins will break down. Deviation outside of the pH range can induce coma or even cause death.

So how is it that we can ingest or inhale acidic or basic substances and not die? Buffers are the key. Buffersreadily absorb excess H^+ or OH^-, keeping the pH of the body carefully maintained in the aforementioned narrow range. Carbon dioxide is part of a prominent buffer system in the human body; it keeps the pH within the proper range. This buffer system involves carbonic acid (H_2CO_3) and bicarbonate (HCO_3^-) anion. If too much H^+ enters the body, bicarbonate will combine with the H^+ to create carbonic acid and limit the decrease in pH. Likewise, if too much OH^- is introduced into the system, carbonic acid will rapidly dissociate into bicarbonate and H^+ ions. The H^+ ions can combine with the OH^- ions, limiting the increase in pH. While carbonic acid is an important product in this reaction, its presence is fleeting because the carbonic acid is released from the body as carbon dioxide gas each time we breathe. Without this buffer system, the pH in our bodies would fluctuate too much and we would fail to survive.

Summary

Water has many properties that are critical to maintaining life. It is polar, allowing for the formation of hydrogen bonds, which allow ions and other polar molecules to dissolve in water. Therefore, water is an excellent solvent. The hydrogen bonds between water molecules give water the ability to hold heat better than many other substances. As the temperature rises, the hydrogen bonds between water continually break and reform, allowing for the overall temperature to remain stable, although increased energy is added to the system. Water's cohesive forces allow for the property of surface tension. All of these unique properties of water are important in the chemistry of living organisms.

The pH of a solution is a measure of the concentration of hydrogen ions in the solution. A solution with a high number of hydrogen ions is acidic and has a low pH value. A solution with a high number of hydroxide ions is basic and has a high pH value. The pH scale ranges from 0 to 14, with a pH of 7 being neutral. Buffers are solutions that moderate pH changes when an acid or base is added to the buffer system. Buffers are important in biological systems because of their ability to maintain constant pH conditions.

Footnotes

1. 1 Humphrey, W., Dalke, A. and Schulten, K., "VMD—Visual Molecular Dynamics", *J. Molec. Graphics*, 1996, vol. 14, pp. 33-38. http://www.ks.uiuc.edu/Research/vmd/

Glossary

acid
a substance that donates hydrogen ions and therefore lowers pH

adhesion
the attraction between water molecules and molecules of a different substance

base
a substance that absorbs hydrogen ions and therefore raises pH

buffer
a solution that resists a change in pH by absorbing or releasing hydrogen or hydroxide ions

cohesion
the intermolecular forces between water molecules caused by the polar nature of water; creates surface tension

evaporation
the release of water molecules from liquid water to form water vapor

hydrophilic
describes a substance that dissolves in water; water-loving

hydrophobic
describes a substance that does not dissolve in water; water-fearing

litmus paper
filter paper that has been treated with a natural water-soluble dye so it can be used as a pH indicator

pH scale
a scale ranging from 0 to 14 that measures the approximate concentration of hydrogen ions of a substance

solvent
a substance capable of dissolving another substance

surface tension
the cohesive force at the surface of a body of liquid that prevents the molecules from separating

temperature
a measure of molecular motion

Contributors and Attributions

- Samantha Fowler (Clayton State University), Rebecca Roush (Sandhills Community College), James Wise (Hampton University). Original content by OpenStax (CC BY 4.0; Access for free at https://cnx.org/contents/b3c1e1d2-83...4-e119a8aafbdd).

This page titled 2.1: Water is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by OpenStax.

- **2.2: Water** by OpenStax is licensed CC BY 4.0.
- **2.1: Water** by OpenStax is licensed CC BY 4.0.

2.2: The multiple roles of water

Search Fundamentals of Biochemistry

Learning Goals (ChaptGPT o1, 1/25/25)

- **Interpretation of Duality in Molecular Properties:**
  - Analyze the paradox presented in the riddle ("weak yet strong, strengthens yet destroys") and relate it to water's ability to donate and accept protons/electrons.
  - Explain how water's seemingly contradictory behaviors underpin its role in essential biochemical reactions.
- **Water as a Solvent – Thermodynamic Considerations:**
  - Describe how noncovalent interactions (ion-ion, ion-dipole, hydrogen bonding, etc.) govern the solubility of molecules in water.
  - Evaluate the roles of enthalpy and entropy in solvation, including the formation of clathrate structures and the impact of "holes" in the solvent.
- **Solubility and Amphiphilicity:**
  - Investigate how amphiphilic molecules (like fatty acids) self-assemble into micelles, and the biochemical significance of these aggregates in transporting nonpolar molecules in aqueous environments.
  - Discuss experimental approaches, such as liquid/liquid partitioning, to illustrate the principles of solubility and phase separation.
- **Water's Acid-Base Chemistry:**
  - Examine water's role as both a Brønsted–Lowry acid and base in chemical reactions, and explain the concept of the leveling effect in aqueous systems.
  - Explore acid-base equilibria involving water, including examples with strong acids (e.g., HCl) and weak acids (e.g., acetic acid).
- **Nucleophilic and Electrophilic Reactions Involving Water:**
  - Analyze water's dual functionality as a nucleophile and electrophile in various reaction mechanisms, such as ligand substitution in transition metal complexes and electrophilic hydration of alkenes.
  - Compare and contrast the roles of water and hydroxide ions (OH^-) in hydrolysis reactions.
- **Water in Redox Chemistry:**
  - Explore how water acts as an oxidizing or reducing agent in reactions with metals (e.g., Na, Zn) and in biological processes like photosynthesis.
  - Critically assess how water's redox behavior contributes to generating and controlling reactive oxygen species in biological systems.
- **Molecular Modeling and Visualization:**
  - Utilize interactive molecular models (such as iCn3D) to visualize how water interacts with biomolecules, highlighting the separation of polar and nonpolar regions in structures like micelles and lipoproteins.
  - Interpret structural models of protein complexes (e.g., ferritin, hemoglobin) to understand how water's chemical properties influence the stabilization and function of macromolecular assemblies.
- **Biochemical Implications of Water's Unique Properties:**
  - Develop hypotheses on how water's contrasting properties are essential for life, linking its molecular behavior to its role in processes like nutrient transport, enzyme activity, and cellular structure formation.

- Critically assess experimental data and current research that highlight the centrality of water in maintaining biochemical homeostasis.

These goals are intended to reinforce theoretical knowledge and encourage critical thinking about how fundamental chemical principles manifest in complex biological systems.

"Nothing in the world is as soft and yielding as it,
Yet nothing can better overcome the hard and strong,
For they can neither control nor do away with it.
The soft overcomes the hard,
The yielding overcomes the strong;"

These words come from Lao Tzu's Tao Te Ching. Let's convert this into a chemical riddle and apply it to biochemistry at the nanoscopic level!

"What it loses, it gains,
What it donates, it accepts,
It is weak yet strong,
It strengthens yet destroys;"

What is it? The answer (one of many possible) is water! It gains and loses protons, donates and accepts electrons, can be both a weaker or stronger acid or base, and can act as both an oxidizing and reducing agent, depending on the circumstances. Water, at least on our planet, appears necessary for life. We know of no biological life form that exists without it. This molecule possesses many unique properties, making it distinct from most other liquids and well-suited for the type of life on Earth. It has contrasting and oppositional properties. Let's investigate a few.

2.2.1: Water as a solvent

Solubility is a property that depends on the nature of both solute and solvent. To a first approximation, we tell students in introductory chemistry and biology courses that for a solute to dissolve in a solvent and form a solution (an example of a homogenous mixture), the sum of noncovalent interactions (intermolecular forces) between solute and solvent must be greater than those among solute molecules and those among solvent molecules.

As students advance in chemistry classes, nuance is added to that general understanding as entropic contributions to solubility must be considered. Entropy is often described as the degree of apparent disorder in the system. Given that description, changes in entropy would appear to favor the soluble state as a solution of the solute in solvent would be more disordered. That simple description must be adjusted to account for the ordered state of solvent (a clathrate) surrounding a solute and of "holes" in the solvent that accommodate larger solute molecules. Enthalpy considerations must also be considered. The description of entropy as a measure of disorder is not precise. Rather, it should be described as a measure of the number of microstates of energy or particles available within a system. An increase in entropy would arise from an increase in the number of such available microstates, which could correlate with an increase in the disorder of a system.

Students might often consider a molecule soluble or insoluble in a given solvent. This notion can be reinforced by simple liquid-liquid partitioning experiments in organic chemistry, using two immiscible solvents (for example, water and ether). Yet diethyl ether is partially soluble in water (1 g/100 mL). Nonpolar molecules with no or few bond dipoles are generally considered insoluble. Students would know that acetic acid, a two-carbon molecule, is soluble in water, but how many carbons are necessary for the molecule to become essentially insoluble? Molecules with a single polar group (-OH, CO_2H) and a long alkyl/acyl chain are best described as amphiphilic. Amphiphiles like octanol ($C_8H_{17}OH$) and dodecyl sulfate ($CH_3(CH_2)_{10}CO_2H$) can form multimolecular aggregates called **micelles** even as they exist in a biphasic system, as shown in the following equilibria:

$$\mathrm{C_8H_{17}OH\,(liq)} \leftrightarrow \mathrm{C_8H_{17}OH\,(aq)} \rightleftharpoons \mathrm{C_8H_{17}OH(micelle)}.$$

Figure 2.2.1 shows an interactive iCn3D model of a micelle below, which consists of 54 self-associated molecules of dodecylphosphocholine fatty acids. It has an almost "complete" separation of polar (on the surface) and nonpolar atoms (buried).

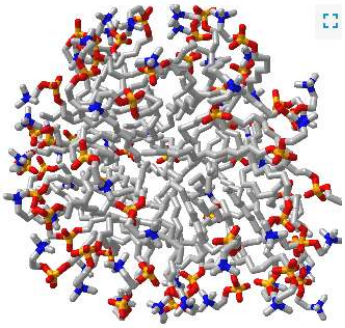


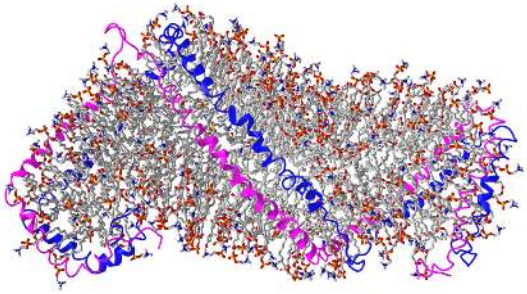
Figure 2.2.1: A Micelle (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...Y7zSYepTQu7yV7.

Note the grey lines representing the nonpolar tails are buried from the surrounding water molecules, which form H bonds with the polar head groups.

Without some limited solubility, the following reaction could not occur:

$$n\,C_8H_{17}OH\,(aq) \leftrightarrow 1\text{-}C_8H_{17}OH(\text{micelle})\,\cdot$$

To solve the general problem of the limited solubility of organic molecules in aqueous-based life, biomolecular structures have evolved to “transport” mostly nonpolar molecules like long-chain carboxylic acids (fatty acids) and cholesterol in circulation. The structure of one such fatty acid and cholesterol-containing particle, nascent high-density lipoprotein (HDL), has been determined by small-angle neutron scattering. Figure 2.2.2 shows an interactive iCn3D model of it. The gray sticks represent the nonpolar acyl tails of the long-chain carboxylic (fatty) acids, while the polar red (oxygen) and blue (nitrogen) atoms surrounding the surface are polar groups connected to the tails. The long magenta and dark blue "helices" represent a protein that wraps around the particle and stabilizes it.



NCBI iCn3D Figure 2.2.2: Nascent HDL (3k2s) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fF85h11SpeYJg6

The same ideas can be applied to the solubility of salts. From introductory chemistry, students will recall general solubility rules (all Group 1 and Group 7 salts are soluble). Salts of divalent cations are less soluble because the attractive ion-ion forces within the solid crystal lattice are too strong for the compensatory ion-dipole interactions between the ion and water. Hence, salts of Ca^{2+} and Fe^{2+} ions, such as $CaCO_3$ and $FeCO_3$, are generally insoluble (K_{sp} values of 1.4 x 10^{-8} and 3.1 x 10^{-11}, respectively). Insoluble calcium salts (carbonates and silicates) are needed for the shells of Foraminifera and the skeletons of vertebrates. Yet, free Ca^{2+} and Fe^{2+} ions are found in extracellular and intracellular compartments. Divalent cations, such as Fe^{2+}, can be toxic at high concentrations, so methods for effectively transporting and sequestering them have evolved. Figure 2.2.3 shows the structure of human heavy-chain ferritin (4zjk), a protein that forms a hollow shell in which Fe^{2+} ions (along with counter ions) are stored. The model below shows a ferritin with 120 Fe^{2+} ions (spheres) inside the hollow ferritin sphere.

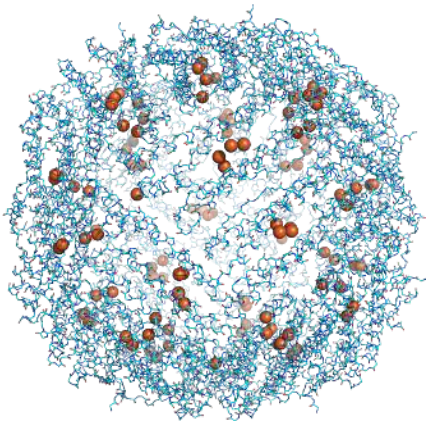


Figure 2.2.3: Human heavy-chain ferritin (4zjk) containing 120 Fe^{2+} ions encapsulated within the hollow sphere formed by the protein.

Finally, let's consider the solubility of gases. The most abundant and relevant ones are O_2 and CO_2, as they are the reactants and products of oxidative respiration. Although the gases contain oxygen atoms, they are nonpolar and therefore have no net dipole moment. Hence, they are quite insoluble in water. However, they must be soluble enough to allow fish to extract them from water. To solve the solubility problem, evolution has produced proteins such as vertebrate hemoglobin, which bind oxygen via a transition metal complex containing the Fe^{2+}-heme complex (found in vertebrate hemoglobin). Some invertebrates use Cu ions in hemocyanins for the same purpose. Figure 2.2.4 shows an interactive iCn3D model of dioxygen (red spheres), bound to a planar heme (yellow highlights), which contains an Fe^{2+} at its center (not shown) in human hemoglobin (6BB5)

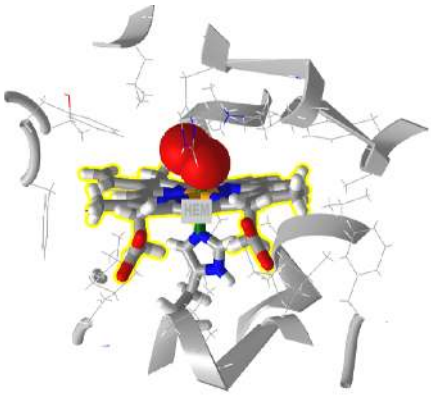


Figure 2.2.4: Oxy-Heme in Human Hemoglobin (6BB5). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...DQ5syZ2XTZ4kR9

Water engages in noncovalent interactions with itself and other molecules. Individual noncovalent interactions are weak, but when combined, they can lead to very strong interactions. You've studied noncovalent interactions before, which may have been described as "intermolecular forces." We prefer the term noncovalent interaction. These include ion-ion, ion-dipole, hydrogen bonds, dipole-dipole, induced dipole-induced dipole, and other variants.

All of these interactions originate in the electrostatic *force* between two charged objects. There is only one law that describes the forces of attraction, and that's Coulomb's Law:

$$F = \frac{kQ_1Q_2}{r^2}$$

From this force, all the electrostatic *interactions* listed above are derived. The magnitude of the attractions for these electrostatic interactions depends on the way the charge is distributed in the attracting species. We will explore these in depth in Chapter 2.4.

2.2.2: Water as a Reactant: Acids and Bases

H_2O, with its shareable lone pairs and slightly positive Hs, is both a Brønsted–Lowry base and acid. Its acid-base chemistry is, hence, among its most important features.

Water, acting as a base, can react with strong and weak acids. Examples of reactions of a strong acid (HCl) and weak acids (acetic acid and ammonium) with water as a base are shown in Figure 2.2.5.

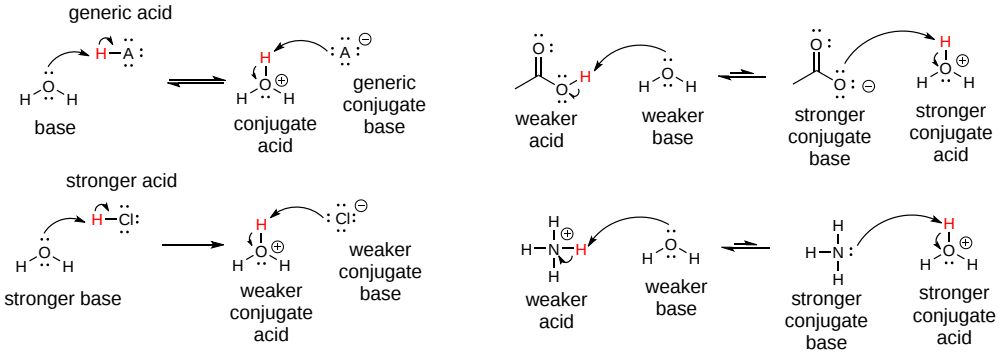


Figure 2.2.5: Reaction of acids with water as a base

Likewise, water can act as an acid as demonstrated in Figure 2.2.6.

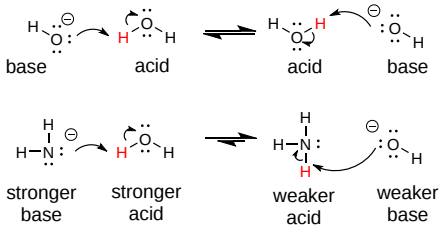


Figure 2.2.6: Reaction of bases with water as an acid

In the first example, no net changes occur. In the second, a negatively charged deprotonated amine (a stronger base than water) can accept a proton from water, which acts as an acid. All acid/base reactions go predominantly in the direction of a stronger acid/strong base to a weaker acid/weaker base. Whether water reacts with a strong acid, such as HCl, or a weak one, like acetic acid, the strongest acid that can exist in an aqueous system is $H_3O^+_{(aq)}$. This is an example of the leveling effect.

2.2.3: Water as a reactant: nucleophile/electrophile

We characterized water as a Brønsted–Lowry acid or base in the reactions above. More generally, we could have said that water is a Lewis acid (electron pair acceptor) or a Lewis base (electron pair donor). In many reactions, we can also refer to water as a nucleophile (when it shares its lone pair) or an electrophile (when its slightly positive hydrogen atoms react with a nucleophile). Here are some examples.

Reaction of water with a transition metal complex.

This reaction below is effectively a nucleophilic substitution reaction in which water displaces ammonia as a ligand, as shown in Figure 2.2.7 and the following chemical equation.

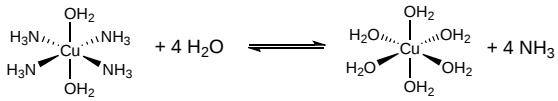


Figure 2.2.7: Reaction of water with transition metal complexes

$$[Cu(NH_3)_4(H_2O)_2]^{2+} + 4\,H_2O \rightleftharpoons [Cu(H_2O)_6]^{2+} + 4\,NH_3$$

Hydration of an alkene

The reaction is catalyzed by adding a proton from an acid (such as H_2SO_4), which can be described as an electrophilic hydration. Once protonated at the carbon, which makes the most stable carbocation, water, as a nucleophile, attacks the positive carbon to produce the alcohol. These steps are illustrated in Figure 2.2.8.

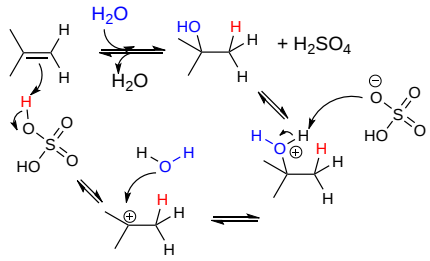


Figure 2.2.8: Mechanism for hydration of an alkene

Nucleophilic substitution at an electrophilic carbonyl

This is a very common reaction. When water acts as the nucleophile, the reaction is called hydrolysis. The reactions in Figure 2.2.9 are shown with OH^- as the nucleophile instead of water for simplicity.

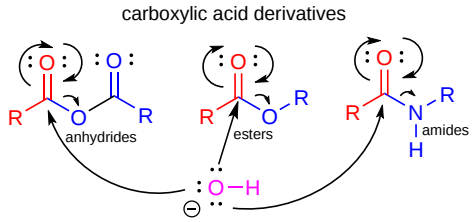


Figure 2.2.9: Reaction of water (OH^-) as a nucleophile

2.2.4: Water as a reactant: Oxidizing/Reducing agent

Everyone knows what happens if you throw a piece of solid Na or K into water. This extremely exothermic reaction releases H_2 gas, which can catch fire and lead to an explosion. The reaction of Na is:

$$2\,Na\,(s) + H_2O \longrightarrow 2\,Na^+\,(aq) + OH^-\,(aq) + H_2\,(g)\,.$$

The oxidation number of elemental sodium is 0, while Na^+ is +1. This indicates that water oxidizes sodium metal, acting as an oxidizing agent.

This reaction occurs with many pure metals, but some that are less reactive (remember the activity series from introductory chemistry?) require acid, a protonated form of water, as shown in the reaction below:

$$Zn\,(s) + 2\,H_3O^+\,(aq) \longrightarrow Zn^{2+}\,(aq) + 2\,H_2O\,(l) + H_2\,(g)$$

As in acid-base reactions, in a redox reaction an oxidizing agent and a reducing agent form new oxidizing and reducing agents. Other reactants can oxidize water to form oxygen. Consider fluorine gas, for example:

$$3\,F_2 + 2\,H_2O \longrightarrow O_2 + 4\,HF$$

F2 is a stronger oxidizing agent (as you would surmise from its electronegativity) than O2, so the reaction proceeds vigorously to the right.

Of more biological relevance is the oxidation of water to produce O_2 in photosynthesis, a complex series of reactions that is effectively the reverse of combustion:

$$6\,CO_2\,(g) + 6\,H_2O\,(l) \longrightarrow C_6H_{12}O_6\,(s) + 6\,O_2(g)\,.$$

This endergonic reaction requires a large energy input and produces the potent oxidizing agent O_2. The special oxygen-evolving complex in photosynthesis is a powerful oxidant that can oxidize H_2O to form the weaker oxidizing agent O_2.

2.2.5: Summary

Chapter Summary

This chapter delves into water's unique and paradoxical role in biochemistry, using a chemical riddle as an engaging entry point. The riddle—“What it loses, it gains; What it donates, it accepts; It is weak yet strong; It strengthens yet destroys”—is a metaphor

for water's ability to act in multiple, seemingly opposing ways. At the heart of the discussion is the recognition that water's versatile chemical behavior underpins many critical processes essential to life.

Key Concepts Covered:

- **Dual Nature of Water:**
  Water's capacity to donate and accept protons and electrons makes it an ideal medium for countless biochemical reactions. Its ability to behave as a weak or strong acid/base and as an oxidizing or reducing agent illustrates its adaptability in various chemical environments.
- **Water as a Solvent:**
  The chapter emphasizes water's role as a universal solvent in biological systems. It explains how the interplay of noncovalent interactions—such as hydrogen bonds, ion-dipole, and van der Waals forces—facilitates the dissolution of solutes. The discussion extends into the nuanced contributions of enthalpy and entropy, including the formation of structured solvent shells (clathrates) and the accommodation of solute molecules.
- **Solubility, Amphiphilicity, and Micelle Formation:**
  The text explores how solubility is not a binary trait but a continuum influenced by the nature of both the solute and the solvent. It introduces amphiphilic molecules and their tendency to form micelles, a key concept for understanding the transport of hydrophobic compounds in aqueous environments. Models of micelles and lipid-protein assemblies, such as nascent high-density lipoproteins (HDL), illustrate this point vividly.
- **Acid–Base and Nucleophilic/Electrophilic Reactions:**
  Water's dual role as a Brønsted–Lowry acid/base and as a Lewis acid/base is examined through its reactions with both strong and weak acids and its participation in nucleophilic substitutions and electrophilic additions. The leveling effect in aqueous solutions is highlighted as a fundamental principle governing these reactions.
- **Redox Chemistry of Water:**
  The chapter outlines how water acts as an oxidizing and reducing agent in reactions with metals, a particularly important feature in the context of biological redox processes. Examples include the reaction of water with alkali and transition metals and its critical role in photosynthetic water oxidation.
- **Molecular Modeling and Structural Insights:**
  To reinforce theoretical concepts, the chapter uses interactive molecular models, including those of micelles, ferritin, and hemoglobin. These visualizations help illustrate how water's chemical properties affect the structure and function of biomolecules, influencing everything from protein stabilization to the transport of essential ions.

In summary, the chapter comprehensively examines water as more than just a solvent—it is a dynamic participant in the chemical processes that sustain life. The text provides junior and senior biochemistry majors with a deeper understanding of how water's contrasting properties are harnessed in biological systems by integrating principles of thermodynamics, solubility, acid-base chemistry, redox reactions, and molecular interactions.

This page titled 2.2: The multiple roles of water is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **2.1: The multiple roles of water** by Henry Jakubowski and Patricia Flatt is licensed CC BY 4.0.

2.3: pH and Buffers

Source: BiochemFFA_1_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Water organization

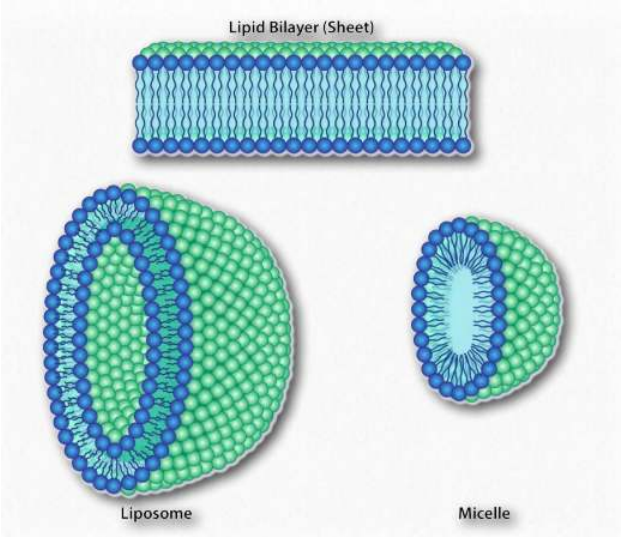


Figure 1.25 - Structures formed by amphiphilic substances in water. Image by Aleia Kim

There is a significant difference, though between surrounding a non-polar molecule with water molecules and surrounding ions (or polar compounds) with water molecules.

The difference is that since non-polar molecules don't really interact with water, the water behaves very differently than it does with ions or molecules that form hydrogen bonds. In fact, around each non-polar molecule, water gets very organized, aligning itself regularly. As any freshman chemistry student probably remembers, entropy is a measure of disorder, so when something becomes ordered, entropy decreases, meaning the ΔS is negative, so the TΔS term in the equation is positive (negative of a negative).

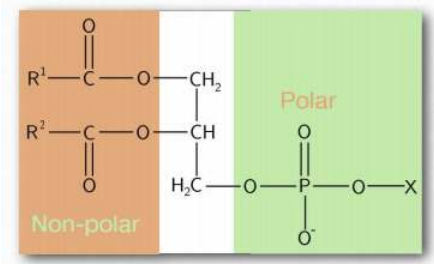


Figure 1.26 - A phospholipid - an amphiphilic substance

Since mixing a non-polar substance with water doesn't generally have any significant heat component, the ΔG is positive. This means, then, that dissolving a non-polar compound in water is not favorable and does not occur to any significant extent. Further, when the non-polar material associates with itself and not water, then the water molecules are free to mix, without being ordered, resulting in an increase of entropy. Entropy therefore drives the separation of non-polar substances from aqueous solutions.

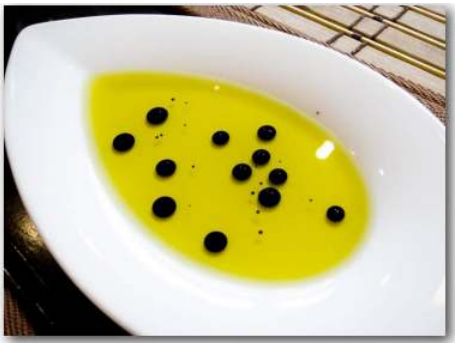


Figure 1.27 - Vinegar (black) and oil (yellow) A mix of polar and nonpolar compounds Wikipedia

Amphiphilic substances

Next, we consider mixing of an amphiphilic substance, such as a soap, with water (Figure 1.24). The sodium ions attached to the fatty acids in soap readily come off in aqueous solution, leaving behind a negatively charged molecule at one end and a non-polar region at the other end. The ionization of the soap causes in an increase in entropy - two particles instead of one. The non-polar portion of the negatively charged soap ion is problematic - if exposed to water, it will cause water to organize and result in a decrease of entropy and a positive ΔG.

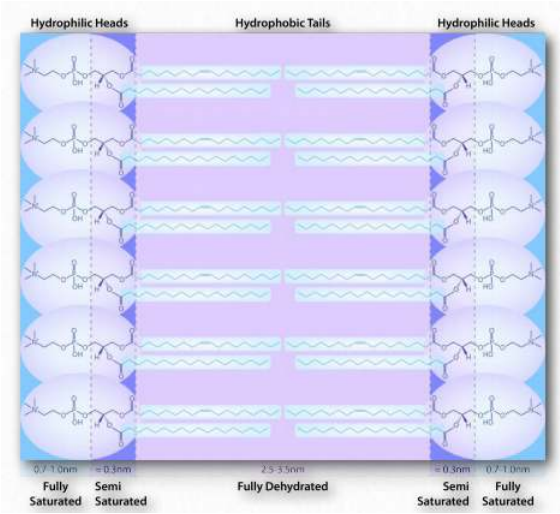


Figure 1.28 - Environment of a lipid bilayer. Water is concentrated away from the hydrophobic center, being saturated on the outside, semi-saturated near the head-tail junction and fully dehydrated in the middle. Image by Aleia Kim

Since we know fatty acids dissolve in water, there must be something else at play. There is. Just like the non-polar molecules in the first example associated with each other and not water, so too do the non-polar portions of the soap ions associate with each other and exclude water. The result is that the soap ions arrange themselves as micelles (Figure 1.25) with the non-polar portions on the interior of the structure away from water and the polar portions on the outside interacting with water.

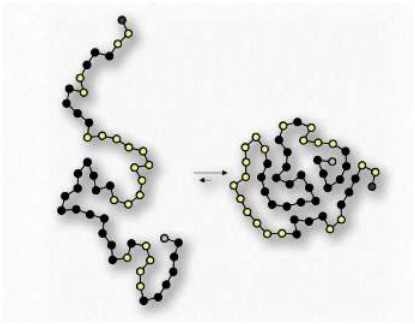


Figure 1.29 - Protein folding arranges hydrophobic amino acids (black dots) inside the protein

The interaction of the polar heads with water returns the water to its more disordered state. This increase in disorder, or entropy, drives the formation of micelles. As will be seen in the discussion of the lipid bilayer, the same forces drive glycerophospholipids and sphingolipids to spontaneously form bilayers where the non-polar portions of the molecules interact with each other to exclude water and the polar portions arrange themselves on the outsides of the bilayer (Figure 1.28).

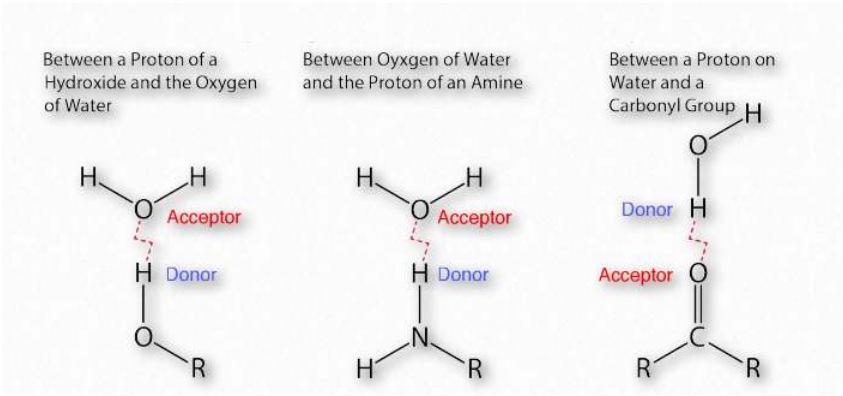


Figure 1.30 - Common hydrogen bonds in biochemistry Image by Aleia Kim

Yet another example is seen in the folding of globular proteins in the cytoplasm. Nonpolar amino acids are found in the interior portion of the protein (water excluded). Interaction of the non-polar amino acids turns out to be a driving force for the folding of proteins as they are being made in an aqueous solution.

Hydrogen bonds

The importance of hydrogen bonds in biochemistry (Figure 1.30) is hard to overstate. Linus Pauling himself said,

" . . . . I believe that as the methods of structural chemistry are further applied to physiological problems it will be found that the significance of the hydrogen bond for physiology is greater than that of any other single structural feature."

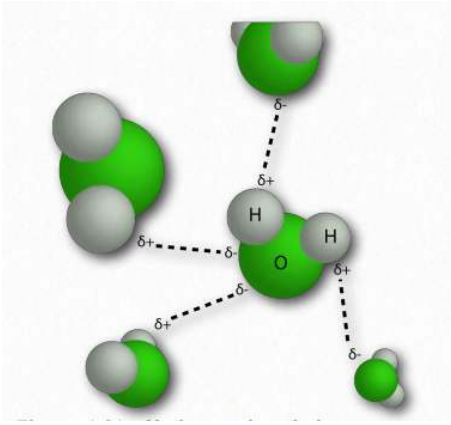


Figure 1.31 - Hydrogen bonds between water molecules Image by Pehr Jacobson

In 2011, an IUPAC task group gave an evidence-based definition of hydrogen bonding that states,

"The hydrogen bond is an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule, in which there is evidence of bond formation."

Partial Charges

The difference in electronegativity between hydrogen and the molecule to which it is covalently bound give rise to partial charges as described above. These tiny charges (δ+ and δ-) result in formation of hydrogen bonds, which occur when the partial positive charge of a hydrogen atom is attracted to the partial negative of another molecule. In water, that means the hydrogen of one water molecule is attracted to the oxygen of another (Figure 1.31). Since water is an asymmetrical molecule, it means also that the

charges are asymmetrical. Such an uneven distribution is what makes a dipole. Dipolar molecules are important for interactions with other dipolar molecules and for dissolving ionic substances (Figure 1.32).

Hydrogen bonds are not exclusive to water. In fact, they are important forces holding together macromolecules that include proteins and nucleic acids. Hydrogen bonds occur within and between macromolecules.

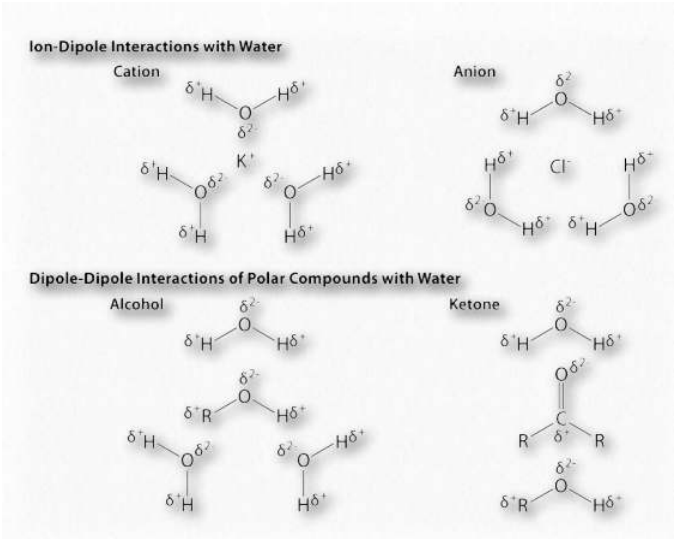


Figure 1.32 - Example dipole interactions in biochemistry Image by Aleia Kim

The complementary pairing that occurs between bases in opposite strands of DNA, for example, is based on hydrogen bonds. Each hydrogen bond is relatively weak (compared to a covalent bond, for example - Table 1.4), but collectively they can be quite strong.

Bond Energies

	Type of Bond	Bond Energy (kJ/mol)
Covalent Bonds	C—H	413
	O—H	460
Noncovalent Bonds	Hydrophobic Interaction	4-12
	Hydrogen Bond	20
	Ion-dipole Interaction	20

Table 1.4 Image by Aleia Kim

Benefits of weak interactions

Their weakness, however, is actually quite beneficial for cells, particularly as regards nucleic acids (Figure 1.33). The strands of DNA, for example, must be separated over short stretches in the processes of replication and the synthesis of RNA. Since only a few base pairs at a time need to be separated, the energy required to do this is small and the enzymes involved in the processes can readily take them apart, as needed. Hydrogen bonds also play roles in binding of substrates to enzymes, catalysis, and protein-protein interaction, as well as other kinds of binding, such as protein-DNA, or antibody-antigen.

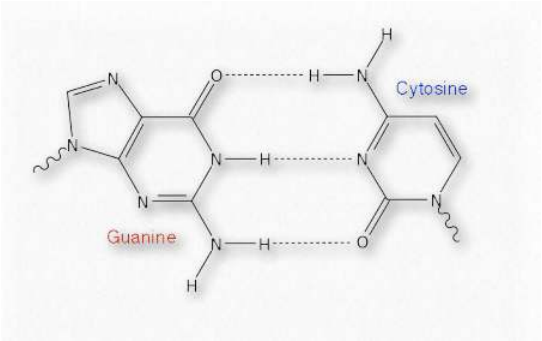


Figure 1.33 - Hydrogen bonds in a base pair of DNA Image by Aleia Kim

As noted, hydrogen bonds are weaker than covalent bonds (Table 1.4) and their strength varies form very weak (1-2 kJ/mol) to fairly strong (29 kJ/mol). Hydrogen bonds only occur over relatively short distances (2.2 to 4.0 Å). The farther apart the hydrogen bond distance is, the weaker the bond is.

The strength of the bond in kJ/mol represents the amount of heat that must be put into the system to break the bond - the larger the number, the greater the strength of the bond. Hydrogen bonds are readily broken using heat. The boiling of water, for example, requires breaking of H-bonds. When a biological structure, such as a protein or a DNA molecule, is stabilized by hydrogen bonds, breaking those bonds destabilizes the structure and can result in denaturation of the substance - loss of structure. It is partly for this reason that most proteins and all DNAs lose their native, or folded, structures when heated to boiling.

Weak Acid pKa Values

Name	Chemical Structure of Acid	Chemical Structure of Salt	pKa
Acetic Acid	CH_3COOH	CH_3COO^-	4.76
Formic Acid	$HCOOH$	$HCOO^-$	3.75
Lactic Acid	$CH_3CHOHCOOH$	$CH_3CH—HCOO^-$	3.86
Pyruvic Acid	$CH_3COCOOH$	$CH_3C—COO^-$	2.50
Oxalic Acid (1)	$HOOC—COOH$	$HOOC—COO^-$	1.23
Oxalic Acid (2)	$HOOC—COO^-$	$OOC—COO^-$	4.19
Carbonic Acid (1)	H_2CO_3	HCO_3^-	6.37
Carbonic Acid (2)	HCO_3^-	CO_3^{2-}	10.20
Malic Acid (1)	$HOOC—CH_2—CHOH—COOH$	$HOOC—CH_2—CHOH—COO^-$	3.40
Malic Acid (2)	$HOOC—CH_2—CHOH—COO^-$	$^-OOC—CH_2—CHOH—COO^-$	5.26
Malonic Acid (1)	$HOOC—CH_2—COOH$	$HOOC—CH_2—COO^-$	2.83
Malonic Acid (2)	$HOOC—CH_2—COO^-$	$^-OOC—CH_2—COO^-$	5.69
Phosphoric Acid (1)	H_3PO_4	$H_2PO_4^-$	2.14
Phosphoric Acid (2)	$H_2PO_4^-$	HPO_4^{2-}	7.20
Phosphoric Acid (3)	HPO_4^{2-}	PO_4^{3-}	12.40
Succinic Acid (1)	$HOOC—CH_2—CH_2—COOH$	$HOOC—CH_2—CH_2—COO^-$	4.21
Succinic Acid (2)	$HOOC—CH_2—CH_2—COO^-$	$^-OOC—CH_2—CH_2—COO^-$	5.63

Image by Aleia Kim Table 1.5

For DNA molecules, denaturation results in complete separation of the strands from each other. For most proteins, this means loss of their characteristic three-dimensional structure and with it, loss of the function they performed. Though a few proteins can readily reassume their original structure when the solution they are in is cooled, most can't. This is one of the reasons that we cook our food. Proteins are essential for life, so denaturation of bacterial proteins results in death of any microorganisms contaminating the food.

The importance of buffers

Water can ionize to a slight extent (10-7 M) to form H+ (proton) and OH- (hydroxide). We measure the proton concentration of a solution with pH, which is the negative log of the proton concentration.

$$pH = -Log[H+]$$

If the proton concentration, [H+]= 10-7 M, then the pH is 7. We could just as easily measure the hydroxide concentration with the pOH by the parallel equation,

$$pOH = -Log[OH^-]$$

In pure water, dissociation of a proton simultaneously creates a hydroxide, so the pOH of pure water is 7, as well. This also means that

$$pH + pOH = 14$$

Now, because protons and hydroxides can combine to form water, a large amount of one will cause there to be a small amount of the other. Why is this the case? In simple terms, if I dump 0.1 moles of H+ into a pure water solution, the high proton concentration will react with the relatively small amount of hydroxides to create water, thus reducing hydroxide concentration. Similarly, if I dump excess hydroxide (as NaOH, for example) into pure water, the proton concentration falls for the same reason.

Acids vs bases

Chemists use the term “acid” to refer to a substance which has protons that can dissociate (come off) when dissolved in water. They use the term “base” to refer to a substance that can absorb protons when dissolved in water. Both acids and bases come in strong and weak forms. (Examples of weak acids are shown in Table 1.5.) Strong acids, such as HCl, dissociate completely in water. If we add 0.1 moles (6.02x1022 molecules) of HCl to a solution to make a liter, it will have 0.1 moles of H+ and 0.1 moles of Cl- or 6.02x1022 molecules of each . There will be no remaining HCl when this happens. A strong base like NaOH also dissociates completely into Na+ and OH- .

Clearing Confusion - Students are often puzzled and expect that

$$[H^+] = [A^-]$$

because the dissociation equation shows one of each from HA. This is, in fact, true ONLY when HA is allowed to dissociate in pure water. Usually the HA is placed into solution that has protons and hydroxides to affect things. Those protons and /or hydroxides change the H^+ and A^- concentration unequally, since A^- can absorb some of the protons and/or HA can release H^+ when influenced by the OH^- in the solution. Therefore, one must calculate the proton concentration from the pH using the Henderson Hasselbalch equation.

$$pH = pK_a + \log([Ac^-]/[HAc])$$

Weak Acids

Weak acids and bases differ from their strong counterparts. When you put one mole of acetic acid (HAc) into pure water, only a tiny percentage of the HAc molecules dissociate into H+ and Ac- . Clearly, weak acids are very different from strong acids. Weak bases behave similarly, except that they accept protons, rather than donate them. Since we can view everything as a form of a weak acid, we will not use the term weak base here.

$$\underset{\text{Acid}}{HA} \rightleftharpoons H^+ + \underset{\text{Salt}}{A^-}$$

Figure 1.34 - Dissociation of a weak acid Image by Aleia Kim

Students are often puzzled and expect that $[H^+] = [A^-]$ because the dissociation equation shows one of each from HA. This is, in fact, true ONLY when HA is allowed to dissociate in pure water. Usually the HA is placed into solution that has protons and hydroxides to affect things. Those protons and /or hydroxides change the H+ and Aconcentration unequally, since A- can absorb

some of the protons and/or HA can release H+ when influenced by the OH- in the solution. Therefore, one must calculate the proton concentration from the pH using the Henderson Hasselbalch equation.

\[pH = pKa + log ([Ac$^-$]/[HAc])\]

Salt/Acid Ratio as a Function of pK_as

pH	[Salt]/[Acid]
$pK_a + 3$	1000
$pK_a + 2$	100
$pK_a + 1$	10
pK_a	1
$pK_a - 1$	1/10
$pK_a - 2$	1/100
$pK_a - 3$	1/1000

Image by Aleia Kim Table 1.6

You may wonder why we care about weak acids. You may never have thought much of weak acids when you were in General Chemistry. Your instructor described them as buffers and you probably dutifully memorized the fact that "buffers are substances that resist change in pH" without really learning what Clearing Confusion - this meant. Buffers are much too important to be thought of in this way.

UPS

Weak acids are critical for life because their affinity for protons causes them to behave like a UPS. We're not referring to the UPS that is the United Parcel Service, but instead, to the encased battery backup systems for computers called Uninterruptible Power Supplies that kick on to keep a computer running during a power failure. The battery in a laptop computer is a UPS, for example.

We can think of weak acids as Uninterruptible Proton Suppliers within certain pH ranges, providing (or absorbing) protons as needed. Weak acids thus help to keep the H+ concentration (and thus the pH) of the solution they are in relatively constant.

Consider the bicarbonate/carbonic acid system. Figure 1.35 shows what happens when H_2CO_3dissociates. Adding hydroxide ions (by adding a strong base like NaOH) to the solution causes the H+ ions to react with OH- ions to make water. Consequently, the concentration of H+ ions would go down and the pH would go up.

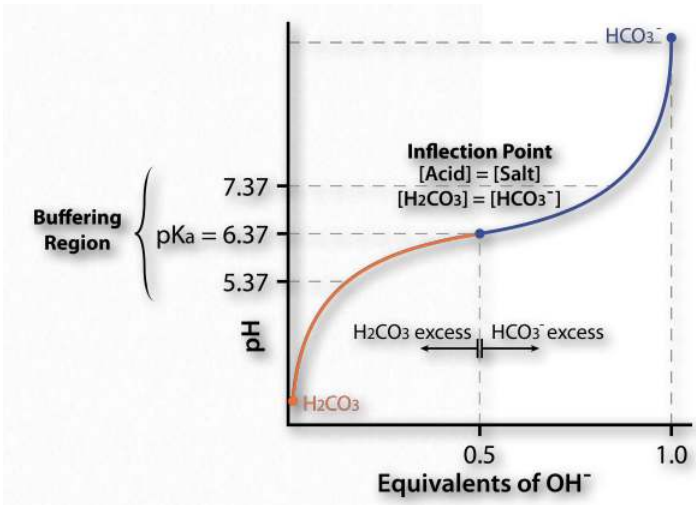


Figure 1.35 - Titration curve for carbonic acid Image by Aleia Kim

However, in contrast to the situation with a solution of pure water, there is a backup source of H+ available in the form of H_2CO_3. Here is where the UPS function kicks in. As protons are taken away by the added hydroxyl ions (making water), they are partly replaced by protons from the H_2CO_3. This is why a weak acid is a buffer. It resists changes in pH by releasing protons to compensate for those "used up" in reacting with the hydroxyl ions.

Why do we care about pH? Because biological molecules can, in some cases, be exquisitely sensitive to changes in it. As the pH of a solution changes, the charges of molecules in the solution can change, as you will see. Changing charges on biological molecules, especially proteins, can drastically affect how they work and even whether they work at all.

Henderson-Hasselbalch

It is useful to be able to predict the response of the H_2CO_3 system to changes in H+ concentration. The Henderson-Hasselbalch equation defines the relationship between pH and the ratio of HCO_3^- and H_2CO_3. It is

$$pH = pKa + \log([HCO_3^-]/[H_2CO_3])$$

This simple equation defines the relationship between the pH of a solution and the ratio of HCO_3^- and H_2CO_3 in it. The new term, called the pKa, is defined as

$$pKa = -\text{Log } K_a,$$

just as

$$pH = -\text{Log } [H^+].$$

The Ka is the acid dissociation constant and is a measure of the strength of an acid. For a general acid, HA, which dissociates as

$$HA \rightleftarrows H^+ + A^-, K_a = [H^+][A^-]/[HA]$$

Thus, the stronger the acid, the more protons that will dissociate from it when added to water and the larger the value its Ka will have. Large values of Ka translate to lower values of pKa. As a result, the lower the pKa value is for a given acid, the stronger the weak acid is.

Constant pKa

Please note that pKa is a constant for a given acid. The pKa for carbonic acid is 6.37. By comparison, the pKa for formic acid is 3.75. Formic acid is therefore a stronger acid than acetic acid. A stronger acid will have more protons dissociated at a given pH than a weaker acid.

Now, how does this translate into stabilizing pH? Figure 1.35 shows a titration curve. In this curve, the titration begins with the conditions at the lower left (very low pH). At this pH, the H_2CO_3 form predominates, but as more and more OH- is added (moving to the 45 Why do we care about pH? Because biological molecules can, in some cases, be exquisitely sensitive to changes in it. As the pH of a solution changes, the charges of molecules in the solution can change, as you will see. Changing charges on biological molecules, especially proteins, can drastically affect how they work and even whether they work at all right), the pH goes up, the amount of HCO3- goes up and (correspondingly), the amount of H_2CO_3 goes down. Notice that the curve “flattens” near the pKa (6.37).

Buffering region

Flattening of the curve tells us is that the pH is not changing much (not going up as fast) as it did earlier when the same amount of hydroxide was added. The system is resisting a change in pH (not stopping the change, but slowing it) in the region of about one pH unit above and one pH unit below the pKa. Thus, the buffering region of the carbonic acid/ bicarbonate buffer is from about 5.37 to 7.37. It is maximally strong at a pH of 6.37.

Now it starts to become apparent how the buffer works. HA can donate protons when extras are needed (such as when OH- is added to the solution by the addition of NaOH). Similarly, A- can accept protons when extra H+ are added to the solution (adding HCl, for example). The maximum ability to donate or accept protons comes when

$$[A^-] = [HA]$$

This is consistent with the Henderson Hasselbalch equation and the titration curve. When [A$^-$] = [HA], pH = 6.37 + Log(1). Since Log(1) = 0, pH = 6.37 = pKa for carbonic acid. Thus for any buffer, the buffer will have maximum strength and display flattening of its titration curve when [A$^-$] = [HA] and when pH = pKa. If a buffer has more than one pKa (Figure 1.36), then each pKa region will display the behavior.

Buffered vs non-buffered

To understand how well a buffer protects against changes in pH, consider the effect of adding .01 moles of HCl to 1.0 liter of pure water (no volume change) at pH 7, compared to adding it to 1.0 liter of a 1M acetate buffer at pH 4.76. Since HCl completely dissociates, in 0.01M (10-2 M) HCl you will have 0.01M H+. For the pure water, the pH drops from 7.0 down to 2.0 (pH = -log(0.01M)).

By contrast, the acetate buffer's pH after adding the same amount of HCl is 4.74. Thus, the pure water solution sees its pH fall from 7 to 2 (5 pH units), whereas the buffered solution saw its pH drop from 4.76 to 4.74 (0.02 pH units). Clearly, the buffer minimizes the impact of the added protons compared to the pure water.

Buffer capacity

It is important to note that buffers have capacities limited by their concentration. Let's imagine that in the previous paragraph, we had added the 0.01 moles HCl to an acetate buffer that had a concentration of 0.01M and equal amounts of Ac- and HAc. When we try to do the math in parallel to the previous calculation, we see that there are 0.01M protons, but only 0.005M A- to absorb them. We could imagine that 0.005M of the protons would be absorbed, but that would still leave 0.005M of protons unbuffered. Thus, the pH of this solution would be approximately

$$pH = -\log(0.005M) = 2.30$$

Exceeding buffer capacity dropped the pH significantly compared to adding the same amount of protons to a 1M acetate buffer. Consequently, when considering buffers, it is important to recognize that their concentration sets their limits. Another limit is the pH range in which one hopes to control proton concentration.

Multiple ionizable groups

Now, what happens if a molecule has two (or more) ionizable groups? It turns out, not surprisingly, that each group will have its own pKa and, as a consequence, will have multiple regions of buffering.

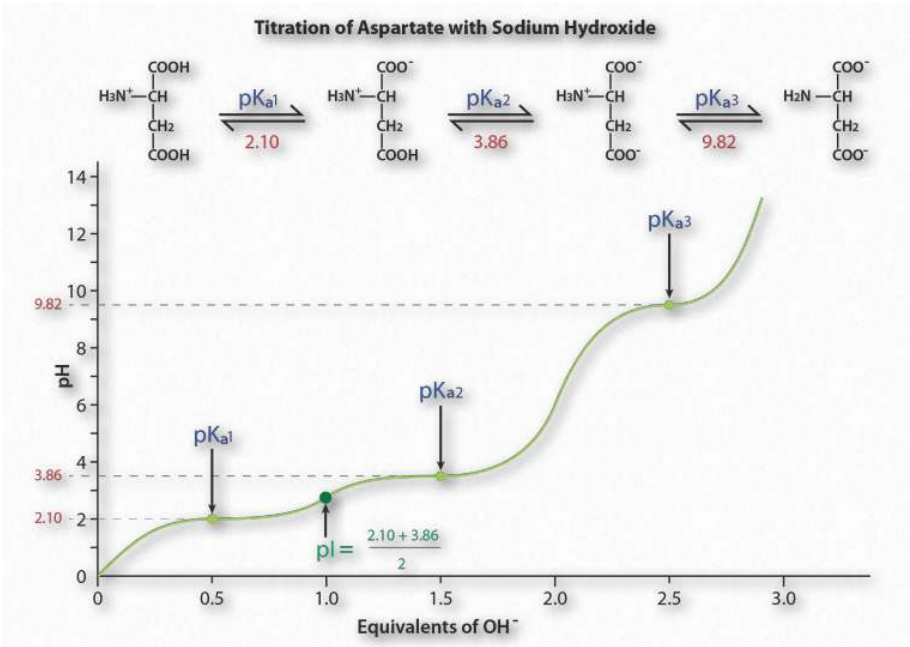


Figure 1.36 - Titration of an acidic amino acid Image by Aleia Kim

Figure 1.36 shows the titration curve for the amino acid aspartic acid. Note that in- stead of a single flattening of the curve, as was seen for acetic acid, aspartic acid's titration curve displays three such regions. These are individual buffering regions, each centered

on the respective pKa values for the carboxyl group and the amine group.

Aspartic acid has four possible charges: +1 (α-carboxyl group, α-amino group, and Rgroup carboxyl each has a proton), 0 (α-carboxyl group missing proton, α- amino group has a proton, R-group carboxyl has a proton), -1 (α-carboxyl group and R-group carboxyl each lack a proton, α-amino group retains a proton), -2 (α-carboxyl, R-group carboxyl, and α-amino groups all lack extra proton).

Prediction

How does one predict the charge for an amino acid at a given pH? A good rule of thumb for estimating charge is that if the pH is more than one unit below the pKa for a group (carboxyl or amino), the proton is on. If the pH is more than one unit above the pKa for the group, the proton is off. If the pH is NOT more than one or less than one pH unit from the pKa, this simple assumption will not work.

Further, it is important to recognize that these rules of thumb are estimates only. The pI (pH at which the charge of a molecule is zero) is an exact value calculated as the average of the two pKa values on either side of the zero region. It is calculated at the average of the two pKa values around the point where the charge of the molecule is zero. For aspartic acid, this corresponds to $pK_{a1}$and $pK_{a2.}$

References

1. http://www.lpi.usra.edu/lunar/missions/apollo/ apollo_12/experiments/surveyor/
2. Arunan, Elangannan; Desiraju, Gautam R.; Klein, Roger A.; Sadlej, Joanna; Scheiner, Steve; Alkorta, Ibon; Clary, David C.; Crabtree, Robert H.; Dannenberg, Joseph J.; Hobza, Pavel; Kjaergaard, Henrik G.; Legon, Anthony C.; Mennucci, Benedetta; Nesbitt, David J. (2011). "Definition of the hydrogen bond". Pure Appl. Chem. 83 (8): 1637–1641. doi:10.1351/PAC-REC-10-01-02

This page titled 2.3: pH and Buffers is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **1.3: Introduction - Water and Buffers** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **2.2: pH and Buffers** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

2.4: Weak Acids and Bases, pH and pKa

Learning Goals (ChaptGPT o1, 1/25/25)

Below is a series of targeted learning goals designed to deepen understanding of acid–base equilibria, titration analyses, and the influence of environmental factors on ionizable groups in biomolecules.

- **Understand Acid–Base Properties of Biomolecules and Water:**
  - Describe how water acts as both a Brønsted–Lowry acid and base, including its autoionization into H_3O^+ and OH^-.
  - Explain how functional groups in biomolecules (e.g., carboxylic acids, phosphoric acids, amines) acquire or lose protons and how this alters their chemical properties.
- **Master the Mathematical Foundations of Acid–Base Equilibria:**
  - Derive and interpret the equilibrium constant expressions for the autoionization of water and for generic acid dissociation reactions.
  - Calculate pH and pKa values using the definitions pH=−log[H3O+]pH = -\log [H_3O^+] and pKa=−logKapK_a = -\log K_a, and understand their physical significance.
- **Apply the Henderson–Hasselbalch Equation:**
  - Use the Henderson–Hasselbalch equation to predict the protonation state of acids in solution under various pH conditions.
  - Interpret titration curves, identify key points such as the inflection point where pH=pKa, and understand the concept of buffering capacity.
- **Analyze Titration Curves and Charge States:**
  - Explain the stepwise protonation/deprotonation processes depicted in titration curves, including the behavior of monoprotic and polyprotic acids.
  - Predict the net charge on biomolecules (e.g., proteins) by applying titration principles to their ionizable functional groups.
- **Explore Polyprotic Acid Behavior:**
  - Define polyprotic acids and explain why each subsequent deprotonation has a higher pKa due to increasing charge repulsion.
  - Analyze complex titration curves for polyprotic acids and discuss the significance of multiple plateaus corresponding to different ionizable groups.
- **Evaluate Environmental Effects on pKa Values:**
  - Discuss how solvent composition (e.g., water vs. ethanol) and the microenvironment (e.g., protein interior vs. surface) influence the pKa of ionizable groups.
  - Use examples such as aspartic acid in different protein contexts to illustrate how nearby charged or hydrophobic groups alter intrinsic acid strength.
- **Connect Acid–Base Chemistry to Biological Function:**
  - Relate the protonation state of biomolecules to their biological functions, such as enzyme activity, ion transport, and protein–protein interactions.
  - Assess how shifts in pH can modulate biomolecular structure and function through changes in charge states.
- **Develop Practical Skills in Quantitative Analysis:**
  - Utilize experimental titration data and interactive models (e.g., spreadsheets, molecular visualizations) to quantitatively determine the pKa values and buffering ranges of various acids and bases.
  - Critically evaluate how titration data supports our understanding of acid–base equilibria in complex biological systems.

These learning goals reinforce theoretical concepts through quantitative reasoning and practical application, ensuring junior and senior biochemistry majors can integrate chemical principles with biological phenomena.

The previous section described the general acid/base properties of water. Many functional groups in both small and large biomolecules act as acids and bases. Common weak acids are carboxylic and phosphoric acid and their derivatives, which become negatively charged after donating a proton. Common weak bases are amines, which become positively charged on protonation. Such charge acquisition changes the properties of the acid or base. A protonated amine is no longer a nucleophile. A deprotonated carboxylic acid can now engage in an ion-ion IMF. The extent of deprotonation depends on the environment's acidity or basicity. We have to turn to a bit of mathematics to determine that extent.

2.4.1: Reaction of water with itself: Autoionization

As shown in the previous section, water can react with itself to produce H_3O^+ and OH^- as illustrated in Figure **2.4.1**.

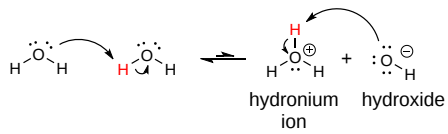


Figure **2.4.1**: Reaction of water with self

This autoionization reaction is often represented in a simpler form:

$$\mathrm{H_2O} \rightleftharpoons \mathrm{H^+} + \mathrm{OH^-} \cdot$$

The equilibrium constant for this simplified reaction can be written as

$$K_{eq} = \frac{[H^+]\,[OH^-]}{H_2O} \tag{2.4.1}$$

Given the known value of K_{eq} and the concentration of water (55 M), this can be simply rewritten as

$$K_a = 55K_{eq} = [H^+]\,[OH^-] = 10^{-14} \tag{2.4.2}$$

(see discussion of the pKa of water below.)

Hence, pure, neutral water has equal but small concentrations, 10^{-7} M of H_3O^+ and OH^-.

You recall from introductory chemistry and everyday life that pure water has a pH of 7. This pH value is derived from the general formulas for both pH and a new quantity, pKa.

$$\begin{aligned} pH &= -\log\left[H_3O^+\right] = -\log\left(10^{-7}\right) = 7 \\ pK_a &= -\log K_a = -\log\left(10^{-14}\right) = 14 \end{aligned} \tag{2.4.3}$$

Note

Some texts incorrectly use 15.7 for the pKa of water. Here is a link to an explanation of why 14 is better. The incorrect value of 15.7 would make the pKa of water higher than that of methanol (15.3), which does not make sense since the methoxide anion is less stable due to electron release by the methyl group than the OH-anion.

Generic Bronsted acids with a dissociable proton can be written using the simplified chemical equation below.

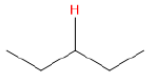
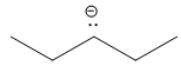
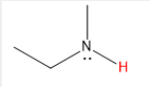
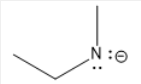
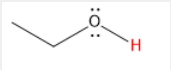
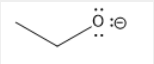
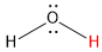
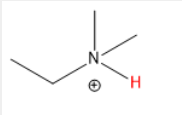
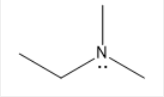
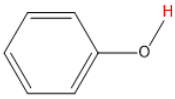
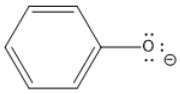
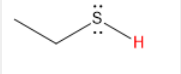
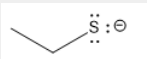
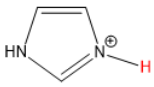
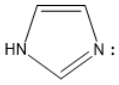
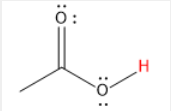
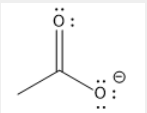
$$\mathrm{HA} \rightleftharpoons \mathrm{H^+} + \mathrm{A^-}$$

The equilibrium constant for this simplified reaction (leaving out water) can be written as

$$\begin{aligned} K_{eq} &= \frac{[H^+]\,[A^-]}{HA} \\ K_a &= [HA]K_{eq} = [H^+]\,[A^-] \\ pK_a &= -\log K_a \end{aligned} \tag{2.4.4}$$

The pKa is a measure of the strength of an acid. The stronger the acid, the larger the Ka and the smaller the pKa.

Here is a table of pK_a values for common acids and functional groups. The pKa values of acids change with different substituents. The stronger the acid, the weaker the conjugate base. This should make sense as a weak base is unlikely to abstract a proton and return to its original acidic form. Likewise, the weakest acids produce the strongest conjugate bases, which would reprotonate to return to the weak acid state.

Group	Example	weaker acid	≈ pKa	Conjugate Base	stronger conj. base
alkane			50		
amine			35		
alkyne			25		
alcohol			16		
water			14		
protonated amine			10		
phenol			10		
thiol			10		
imidazole			7		
carboxylic acid			5		
hydrochloric acid			-8		
		stronger acid			weaker conj. base

2.4.2: The Henderson-Hasselbalch Equation

We can find the pKa for small acids in solution in pKa tables. However, from a biochemical perspective, we often need to know the acid's charge state. Since pH is approximately constant in organisms (more on that later), we know the [H3O+]. Hence, we can calculate the ratio of A^-/HA using the **Henderson-Hasselbalch equation** (Equation 2.2.5), which is derived below.

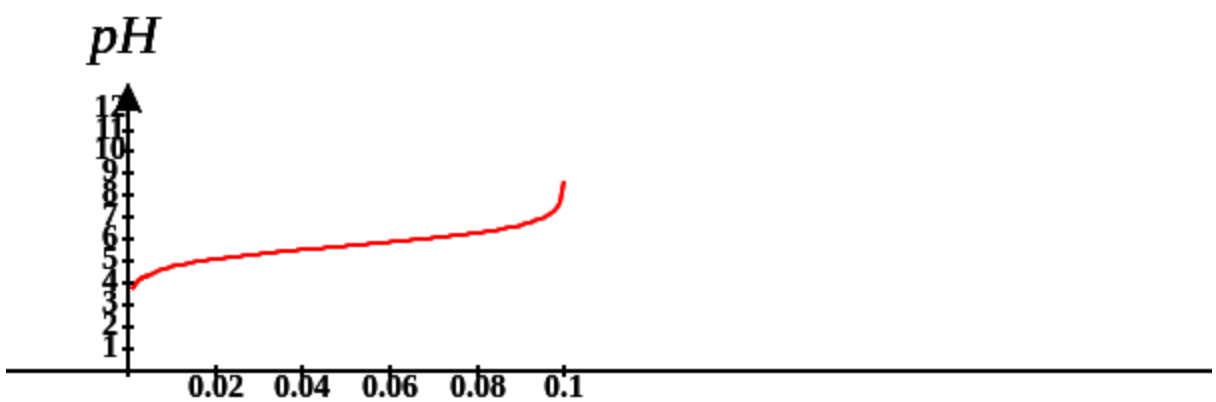
$$\begin{gathered} K_a = \frac{[H^+][A^-]}{[HA]} \\ -\log K_a = -\log [H^+] - \log ([A^-]/[HA]) \\ pK_a = pH - \log ([A^-]/[HA]) \end{gathered} \tag{2.4.5}$$

which gives the traditional Henderson-Hasselbalch equation below.

$$pH = pK_a + \log \frac{[A^-]}{[HA]} \tag{2.4.6}$$

You certainly would have performed titration curve analyses of acids in your chemistry class. What is the chemistry that occurs at each step? Let's assume the pH is low, well below the acid's pKa. From the Henderson-Hasselbalch equation, you would surmise that the ratio of A^-/HA is very small - that is, the acid is essentially fully protonated. That should also make intuitive sense. For a weak acid to be coaxed to give up a proton, a reasonably strong base (like OH^-) should be added. At low pH, the acid exists as HA. Now, consider adding a sufficient amount of NaOH to match the concentration of the ionizable protons. At that point in the titration, the mass balance would suggest that the acid in its protonated state is gone, and all that remains is A^-. What happens if just enough NaOH is added to react with half of the HA? The mass balance would tell us that A^-=HA, and at that point, the pH = pK_a of the acid.

The entire titration curve can be calculated from the Henderson-Hasselbalch equation. A graph of it is shown below.



The graph shifts upward as the pKa increases. The pH starts soaring at the end of the graph after the added hydroxide has reacted with the last ionizable proton. After that, the pH is determined by the concentration of the strong base OH^-. The graph is flattest at the inflection point. Note that at this pH, **pH = pKa.** In the middle of the curve, the pH changes least with the addition of small amounts of OH^-. This is the basis of buffering, which will be covered in the next section.

If you know the pH of a solution and the pKa of the ionizable group, you can very quickly estimate the functional group's average charge (protonation) state. Let's see what the Henderson-Hasselbalch equation (Equation 2.2.5) predicts under three specific pH states:

pH	log (A^-/HA)	(A^-/HA)	protonation state
2 units < pK_a (more acidic, expect protonated)	-2 = log (A^-/HA)	0.01 = (A^-/HA) = 1/100	fn group about 99% protonated
2 units > pK_a (more basic, expect deprotonated)	2 = log (A^-/HA)	100 = (A^-/HA) = 100/1	fn group about 99% deprotontaed
pH = pKa	0 = log (A^-/HA)	1 = (A^-/HA) = 50/59	fn group 50% protonated

From these simple examples, we have illustrated the **±2 rule** to determine the charge state. This rule is used to quickly determine the protonation state and, hence, the charge state of an acid and its conjugate base, and is extremely important to know and use (and easy to derive). If the potential acid HA is a protonated amine (RNH_3^+), the fully deprotonated state (RNH_2) is uncharged. The fully deprotonated state has a -1 charge if it is a carboxylic acid.

2.4.2.1: Polyprotic Oxyacids

Acids that can donate more than one proton are called **polyprotic acids**. They are typically oxyacids, with the ionizable proton on an oxygen atom, which can form a reasonably stable oxyanion (negative on the oxygen) as the oxygen is electronegative and stabilizes the charge. The negative charge on the conjugate base of oxyacids is further stabilized by resonance delocalization involving the doubly bonded oxygen atom. Figure **2.4.2** shows two of the most biologically relevant oxyacids.

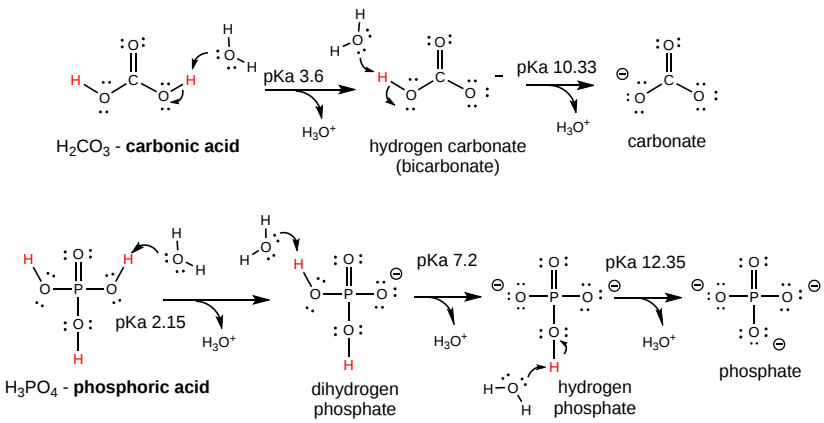


Figure **2.4.2**: Reactions of polyprotic acids with water

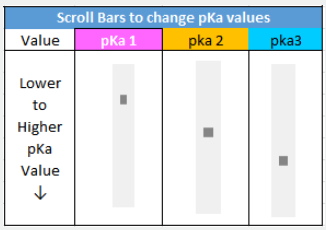
The pKa of each subsequent ionization is higher because removing a proton from an increasingly charged molecular ion is more difficult. The titration plot of pH vs NaOH is similar to the graph above, but has multiple plateaus at pH=pKa,

Derivatives of phosphoric acid are found in all major classes of biomolecules. Nucleic acids contain a sugar-phosphate link in their backbone. Many proteins are phosphorylated after synthesis. Membrane lipids usually contain a phosphate group. A whole class of phospholipids is found in biomembranes.

Titration Curves for Polyprotic Acids

As we will see in a subsequent chapter, all amino acids have an amine and carboxylic acid group, and some have an additional ionizable side chain. Each has its pKa values. Those with three ionizable groups are triprotic acids, much like phosphoric acid. Titration curves for polyprotic acids are more complex than those for monoprotic acids. If the pKa values are separated enough, three general plateaus, each centered at the pKa value of the ionizable group, can be seen in their titration curves. If two of the groups are carboxylic acids, no clear plateau will be observed in the region of the titration curve for those groups.

Click on the Excel file link to download a spreadsheet for triprotic acid titrations with sliders to interactively change pKa values, as shown in the image below.



2.4.2.2: Charge State of Biomolecules

The Henderson-Hasselbalch equation can be used to determine the charge state of ionizable functional groups (carboxylic and phosphoric acids, amines, imidazoles, guanidino groups), even in large macromolecules such as proteins, which contain carboxylic acids (weak acids) and amines (weak bases). Figure **2.4.3** shows how the weakly acidic aspartic and glutamic acids, two common amino acids found in proteins, contribute a negative charge to the protein and how the amine of the amino acid lysine, a weak base, contributes a positive charge.

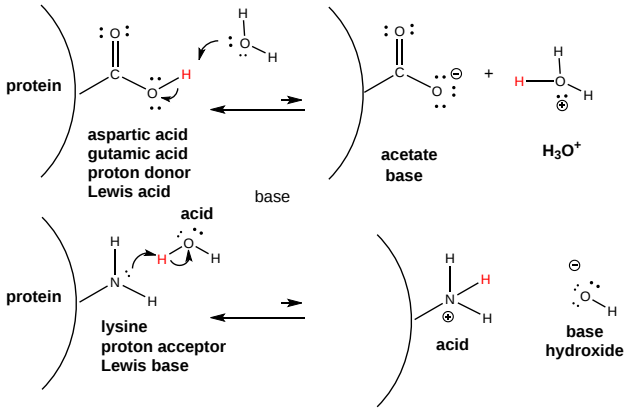


Figure **2.4.3**: Deprotonation/Protonation of weak amino acid side chains contributes to protein charge.

Other amino acids that contain an alcoholic function group can also become phosphorylated to produce a phosphoprotein, which converts a neutral ROH group to a phosphoester with a negative two charge, as shown in Figure **2.4.4**.

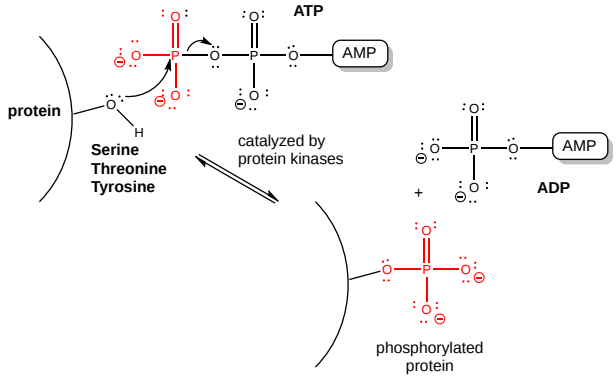


Figure **2.4.4**: Phosphorylation of -OH containing side chains of proteins by ATP

2.4.2.3: pKa and Environment

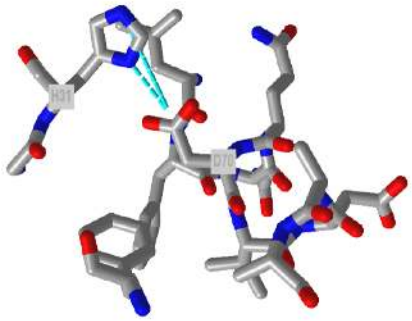
The pKa is a measure of the equilibrium constant for the reaction. And, of course, you remember that $\Delta G^o = -RT \ln K_{eq}$. Therefore, pKa is independent of concentration and depends only on the intrinsic stability of the reactants relative to

the products. However, this is true only under specific conditions, such as temperature, pressure, and solvent composition.

Consider, for example, acetic acid, which in aqueous solution has a pKa of about 4.7. This weak acid dissociates only slightly to form H^+ (in water, the hydronium ion, H_3O^+, is formed) and acetate (Ac^-). These ions are moderately stable in water but reassociate readily to form the starting product. The pKa of acetic acid in 80% ethanol is 6.87. This can be attributed to the decrease in stability of the charged products, which are less shielded from each other due to the less polar ethanol. Ethanol has a lower dielectric constant than water. The pKa increases to 10.32 in 100% ethanol and a whopping 130 in air!

The pKa values of ionizable groups in proteins vary enormously as they depend on the microenvironment of the group. Consider the amino acid aspartic acid (Asp, D), which has a $-CH_2CO_2H$ R-group or "side chain" similar to acetic acid. In a given protein, a given Asp side chain might be on the surface, but another in the same protein might be buried in the protein away from water. You would expect the pKa values for these two Asp side chains to differ. The average pKa for Asp side chains in 78 different proteins was 3.5, less than that of acetic acid (4.7), but not dramatically. However, the range of pKa values for Asp in these proteins was wide, with the lowest at 0.5 (a buried Asp in the protein T4 Lysosome) and the highest at 9.2 in the protein thioredoxin from *E. coli*.

Figure **2.4.5** shows an interactive iCn3D model of the surrounding environment of Asp 70 (D70) in T4 Lysoszyme. Its pKa has been experimentally determined to be 0.5, which is significantly stronger than that of acetic acid. The dotted cyan lines show ion-ion interactions between the $-CH_2CO_2^-$ side chain of Asp 70 (D7) and the positively charged imidazolium group of histidine (H31) in the protein. The distance between the two charged groups is about 3.4 A.



NCBI iCn3D Figure **2.4.5**: Surrounding environment of Asp 70 (D70) in T4 Lysoszyme (2b6z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RRiVn6saN92tU6

The following model shows the surroundings of Asp 26 (D26) in *E. coli* thioredoxin. This particular aspartic acid has a pKa of 9.2. The dark blue group represents a surface-exposed, positively charged lysine side chain that can stabilize the negative charge on Asp 26. Note, however, that it is much farther away than the imidazolium group in T4 lysozyme that stabilizes the negative charge on D70. The rest of the model is colored by hydrophobicity, indicating that nonpolar groups surround the Asp26 side chain. These would destabilize the negative charge on D26, thereby enhancing the stability of protonated (neutral) Asp and elevating its pKa to 9.2.

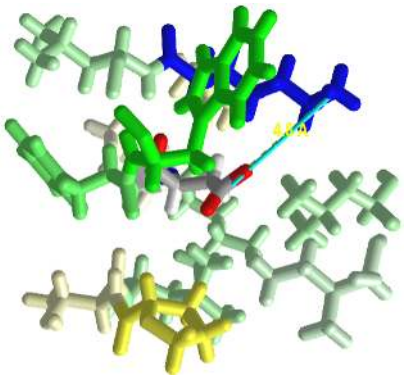


Figure 2.4.6: Surrounding environment of Asp 26 (D26) in E. Coli thioredoxin (5HR2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...i9BNNdbA2bmP5A

2.4.3: Summary

Chapter Summary

This chapter deepens our understanding of acid–base chemistry as it applies to both water and biomolecules. The discussion begins by reinforcing the fundamental acid–base properties of water, highlighting its autoionization into hydronium (H_3O^+) and hydroxide (OH^-) ions. This sets the stage for quantitatively treating acid–base equilibria, introducing essential concepts such as pH, pKa, and the equilibrium constant (K_a). Students are reminded that the concentration of H_3O determines pH^+, while the pKa value indicates the strength of an acid—the lower the pKa, the stronger the acid.

A major focus is placed on the Henderson–Hasselbalch equation, which provides a practical tool for predicting the ratio of deprotonated to protonated forms of an acid in solution. The chapter illustrates how this equation explains the behavior observed in titration curves, where the point at which pH = pKa corresponds to a 50/50 mixture of ionized and unionized species. Through graphical representations and titration analyses, the text clarifies that small pH changes around the pKa have minimal effect on the pH and buffering capacity of biological systems.

The material expands to cover polyprotic acids—molecules that can donate more than one proton. It explains that each successive proton donation generally occurs at a higher pKa because it becomes more difficult to remove a proton from a more negatively charged species. This section underscores the complexity of titration curves for molecules with multiple ionizable groups, such as phosphoric acid and amino acids, which may exhibit multiple plateaus.

Another critical topic is the effect of environmental factors on pKa values. Using examples such as aspartic acid in different protein microenvironments, the chapter demonstrates that the intrinsic pKa of a functional group can shift dramatically depending on local interactions, solvent polarity, and nearby charged residues. These variations are essential for understanding the precise control of protein structure, function, and enzyme activity in vivo.

Overall, the chapter integrates mathematical derivations with biochemical applications, equipping students with the tools to predict the charge states of biomolecules under various conditions. By connecting acid–base theory to biological function, the text reinforces the fundamental role of pH and protonation dynamics in many processes in living systems—from enzyme catalysis to cellular signaling and metabolism.

This page titled 2.4: Weak Acids and Bases, pH and pKa is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **2.2: Weak Acids and Bases, pH and pKa** by Henry Jakubowski and Patricia Flatt is licensed CC BY 4.0.

2.5: Buffering against pH Changes in Biological Systems

Learning Goals

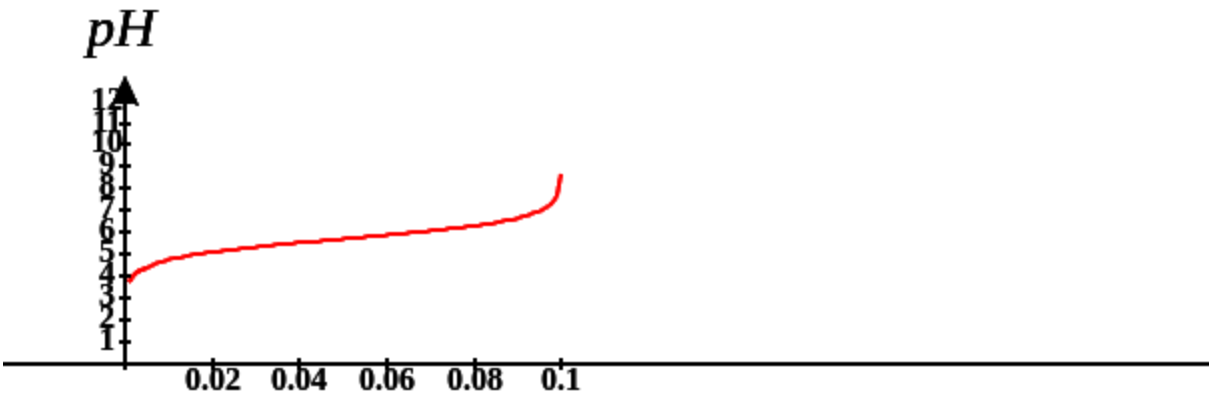
- **Understand pH Homeostasis in Physiology:**
  - Describe how humans maintain blood pH within a narrow range (7.35–7.45) and explain the consequences of deviations (acidosis and alkalosis).
  - Compare pH environments in different cellular compartments (e.g., lysosomes at pH ~4.5 vs. blood).
- **Apply the Henderson–Hasselbalch Equation to Buffer Systems:**
  - Derive and use the Henderson–Hasselbalch equation to predict the ratio of conjugate base to weak acid in buffering systems.
  - Explain why buffering capacity is maximal when pH equals pKa and how small changes in $[A^-]$ or [HA] impact the pH.
- **Analyze the Carbonic Acid/Bicarbonate Buffering System:**
  - Describe the sequential reactions involving CO_2, H_2CO_3, H_3O^+, and HCO_3^- that underlie the carbonic acid/bicarbonate buffering system.
  - Calculate the effective pKa for the system and interpret how the CO_2-to-bicarbonate ratio primes the buffer for rapid response to metabolic changes.
  - Explain how the respiratory system and kidneys work in tandem to regulate this buffer system and, hence, blood pH.
- **Evaluate Other Biological Buffering Agents:**
  - Examine the roles of the phosphate buffering system and protein buffers in maintaining intracellular and extracellular pH.
  - Discuss why certain buffering systems (like the phosphate pair) are less influential in blood compared to the carbonic acid system.
- **Develop Laboratory Buffer Preparation Skills:**
  - Identify strategies for preparing buffered solutions in the lab (mixing separate acid and conjugate base solutions, titrating with strong acid/base, and adjusting volumes).
  - Understand the importance of selecting appropriate buffers (e.g., dihydrogen phosphate/monohydrogen phosphate, HEPES, Tris) based on their pKa values and potential interactions (e.g., ion binding).
- **Link Buffer Systems to Broader Physiological and Environmental Contexts:**
  - Discuss how imbalances in buffering systems can lead to clinical conditions like respiratory acidosis or alkalosis, and explain the physiological responses to such imbalances.
  - Evaluate the impact of increased atmospheric CO_2 on ocean pH and consider how buffering reactions relate to carbon capture strategies in climate change mitigation.

These goals encourage a comprehensive understanding of both the quantitative and qualitative aspects of biological buffers, linking fundamental chemistry with physiological regulation and broader environmental challenges.

2.5.1: Introduction

To maintain homeostasis, humans regulate their pH levels between 7.35 and 7.45. (Much lower pH values, ≈ 4.5, are found in the lysosome.) Lower pH values are associated with metabolic and respiratory acidosis, while higher pH values are characteristic of metabolic and respiratory alkalosis. pH is maintained by buffering systems that consist of a weak acid and a base. If you understand the Henderson-Hasselbalch equation from the previous section, buffer systems become easy to understand.

$$\mathrm{pH} = \mathrm{pK_a} + \log \frac{[\mathrm{A^-}]}{[\mathrm{HA}]} \tag{2.5.1}$$



At the curve's inflection point, pH = pKa; at this pH, the system is most resistant to changes in pH when adding either acid or base. At this pH, [HA]=[A^-].

If a bit of a strong acid is added, it will react with the strongest base in the solution, which would be the conjugate base of the weak acid:

$$HCl + A^- \rightarrow HA + Cl^-$$

The reaction goes from a strong acid, HCl, to a weak acid, HA. Its concentration would increase slightly, but it will only ionize to a small extent since it's a weak acid. The [HA] in the Henderson-Hasselbalch equation increases slightly, but not enough to significantly change the pH. If the same amount of HCl were added to pure water, it would react completely, forming an equal amount of H_3O^+, significantly altering the pH of pure water (7.0).

If a bit of a strong base is added, it will react with the strongest acid in the solution, which would be HA:

$$HA + OH^- \rightarrow H_2O + A^-$$

The reaction goes from a strong base to a weak base A^-. Its concentration would increase slightly, but it wouldn't significantly affect the pH since it's a weak base. The [A^-] in the Henderson-Hasselbalch equations increases slightly but not enough to change the pH significantly. If the same amount of NaOH were added to pure water, it would react, making the solution basic and significantly altering its pH (7.0).

To review, buffer solutions contain a weak acid and its conjugate base. They have a maximal buffering capacity at a pH equal to the weak acid's pKa. Generally, a buffered solution can best withstand a change in pH of only $\pm$ 1 pH unit from the pKa.

2.5.2: Biological Buffering Agents

The most relevant biological systems are the carbonic acid/carbonate buffering system, which controls blood and cell pH, and the phosphate buffering system. Proteins with many weak acid and base functional groups can also act as buffering agents.

Carbonic acid/carbonate buffering system: At first glance, the reaction of carbonic acid, H_2CO_3, with water can be written as follows:

$$\mathrm{H_2CO_3(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + HCO_3^-(aq) \quad pKa = 3.6} \tag{2.5.2}$$

where H_2CO_3 (carbonic acid) is the weak oxyacid, and HCO_3^-(aq) (bicarbonate aka hydrogen carbonate) is its weak conjugate base.

However, this system is more complex since we must add to it another reaction for the formation of H_2CO_3 (aq) in the blood:

$$\mathrm{CO_2(\ g) \leftrightarrow CO_2(aq) + H_2O(l) \leftrightarrow H_2CO_3(aq)} \tag{2.5.3}$$

The [CO_2(aq)] >> [H_2CO_3 (aq)] and their ratio is around 340/1. This makes sense since CO_2 is a very stable molecule. CO_2 in the aqueous form can be readily transported through the blood. Combine the reactions to give the equation below:

$$\mathrm{CO_2(\ g) \leftrightarrow CO_2(aq) + H_2O(l) \leftrightarrow H_2CO_3(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + HCO_3^-(aq)} \tag{2.5.4}$$

How can carbonic acid, with a pKa of 3.6, buffer an aqueous solution at pH 7.5 in the blood and cells? An astute student might have picked up this conundrum. The solution to this problem involves re-examining the complete set of reactions for the components of the buffer system. Let's simplify **Equation 2.3.4** since there would be no free "gas bubbles" in blood, so CO_2 (g) = CO_2(aq):

$$\mathrm{CO_2(aq) + H_2O(l) \leftrightarrow H_2CO_3(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + HCO_3^-(aq)} \tag{2.5.5}$$

$H_2CO_{3(aq)}$ participates in two distinct reactions.

Rightwards from H_2CO_3 (aq) :

$$\mathrm{H_2CO_3(aq) + H_2O(l) \leftrightarrow H_3O^+(aq) + HCO_3^-(aq)} \tag{2.5.6}$$

Using the simplified equation with H^+ gives

$$\mathrm{K_a = \frac{[H^+]\,[HCO_3^-]}{[H_2CO_3]}} \tag{2.5.7}$$

Hence,

$$\mathrm{[H_2CO_3] = \frac{[H^+]\,[HCO_3^-]}{K_a}} \tag{2.5.8}$$

Leftwards from H_2CO_3 (aq) :

$$\mathrm{H_2CO_3(aq) \leftrightarrow CO_2(aq) + H_2O(l)} \tag{2.5.9}$$

$$\mathrm{K_2 = \frac{[CO_2]}{[H_2CO_3]}} \tag{2.5.10}$$

so

$$\mathrm{[H_2CO_3] = \frac{[CO_2]}{K_2}} \tag{2.5.11}$$

Since there can be only 1 H_2CO_3 concentration, set Equations 2.3.8 and 2.3.11 equal to each:

$$\mathrm{[H_2CO_3] = \frac{[H^+]\,[HCO_3^-]}{K_a} = \frac{[CO_2]}{K_2}} \tag{2.5.12}$$

Solving for [H^+] gives:

$$[H^+] = \frac{[CO_2]\,(K_a)}{[HCO_3^-]\,(K_2)} \tag{2.5.13}$$

Now take the -log of each side to produce an equation similar to the Henderson-Hasselbalch equation.

$$-\log[H^+] = -\log\left(\frac{[CO_2]}{[HCO_3^-]}\right) - \log\left(\frac{K_a}{K_2}\right) \tag{2.5.14}$$

where

$$K_{aEFFECTIVE} = \frac{K_a}{K_2} \tag{2.5.15}$$

This Henderson-Hasselbalch-like equation indicates that pH is determined by the K_a/K_2 ratio. $pK_{a\ EFFECTIVE}$ = 6.3. This gives a ratio of CO_2/HCO_3^- of 0.08 = 8/100. There is effectively 12-13 times as much HCO_3^-(aq) as CO_2, making the system primed to react with acid produced metabolically. Yet a second conundrum exists. The pH of the blood (7.4) is outside the optimal range for a buffer system (in this case, $\pm$1 pH unit from the pKa, which is 6.3). Again, the system is primed to react with acid, as this would move the pH closer to the optimal buffering pH of 6.3. Other biological systems also must be involved in maintaining pH.

The respiratory system can quickly adjust pH by increasing CO_2 exhalation. The kidneys can respond more slowly to remove H_3O^+ and retain HCO_3^-. The carbonic acid/bicarbonate buffering system can help us understand how shifting equilibria caused by excessive CO_2 released during rapid, deep breathing or decreased CO_2 release associated with pulmonary disease or shallow, rapid breathing can lead to respiratory alkalosis and acidosis, respectively.

- Respiratory alkalosis can be caused by “hyperventilation” or breathing rapidly. This can lead to breathing out (removing) too much CO_2, shifting the above equilibrium to the left, consuming H_3O^+, and increasing pH, making the blood more alkaline. To increase your CO_2 levels, you could breathe into a bag.
- Respiratory acidosis is caused by increased CO_2, which can occur when the lungs aren’t working well and you can’t get rid of the CO2 you produce during respiration. Respiratory acidosis can occur with asthma, pneumonia, lung disease, or any condition that decreases the respiratory rate.

Inhaling CO_2 can lead to panic. This makes sense as it would mimic suffocation, which is lethal to humans. A suffocation response follows. High CO_2 would drive the equilibrium to the right, leading to H_3O^+ production. An acid-sensing ion channel-1a (ASIC1a) in the amygdala, a center of emotion regulation in the brain, has been identified and appears to mediate panic. Panic attacks are sometimes associated with hyperventilation, which leads to alkalosis, not acidosis. Less noted is that when some people panic, they take short, shallow breaths (in a way, almost stopping their breath). This would lead to a buildup of CO_2, as it wouldn’t be released during exhalation. The acid channel in the amygdala would be activated, and a panic response would ensue.

Phosphate buffering system: Phosphates, specifically dihydrogen ($H_2PO_4^-$) and monohydrogen phosphate (HPO_4^{2-}), are also present in the blood. Given the pKa of HPO_4^{2-}, why is PO_4^{3-} **not** present to any significant degree? Since the concentration of phosphates in blood is low, this system plays a minor role in blood.

Proteins: Proteins are found in all cellular and extracellular fluids and, with their weak acid substituents, act as buffer components. Proteins contain two amino acids, aspartic acid and glutamic acid, that contain carboxylic acid side chains. Each comprises about 6% of the proteins. In blood, hemoglobin is the most abundant protein by far. Its role in buffering and in O_2 and CO_2 will be discussed in a subsequent chapter.

2.5.3: Making Buffers in the Lab

When studying biomolecules, such as proteins and nucleic acids, in the lab, the pH of the solution is typically maintained under physiological conditions. These macromolecules are either dissolved in or diluted into a buffer solution. Sometimes, it's essential to study their properties and activities as a function of pH. A wide variety of buffer systems have been developed for the lab study of these molecules. The dihydrogen phosphate ($H_2PO_4^-$)/monohydrogen phosphate (HPO_4^{2-}) pair is commonly used as the pKa of $H_2PO_4^-$ is 7.21, making them physiologically relevant. Care must be taken when selecting buffer systems, as some might bind

calcium ions. The pKa of some weak acids also varies significantly with temperature. The following are some standard biological buffers.

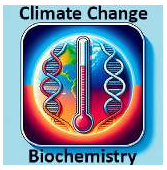
Buffers	**pKa (at 25°C)**
MES	6.10
Bis-Tris	6.50
ACES	6.78
PIPES	6.76
MOPSO	6.90
MOPS	7.20
HEPES	7.48
Tris	8.06
Tricine	8.05
Gly-Gly	8.20
Bicine	8.26
TAPS	8.40
AMPSO	9.00
CAPS	10.40

There are three general ways to make a buffered solution:

1. Make separate and equal concentration solutions for both the weak acid (for example, NaH_2PO_4) and its conjugate base (for example, Na_2HPO_4). Use the Henderson-Hasselbalch equation to calculate how much of each should be added to give the correct $[A^-]/[HA]$ ratio (in the case $[HPO_4^{2-}]/[H_2PO_4^{-1}]$) to give the correct pH.
2. Use a pH meter to monitor the pH as you add both solutions together until the desired pH is achieved.
3. Prepare a solution of one of the components (weak acid or its conjugate base) to the correct volume for the desired molarity. Monitor the pH as you add a concentrated solution of either HCl or NaOH to get the desired pH. Then, bring the solution to the correct volume in a volumetric flask. With this method, you will add counter ions (Cl^- with HCl and Na^+ with NaOH), which you may not want in your buffer solution. Often, it is not a problem.

2.5.4: Climate Change and Carbon Capture

Perhaps humanity's greatest challenge is the effects of global warming and climate change on the biosphere and its health. CO_2 in the atmosphere, HCO_3^-(aq), its soluble form, and CO_3^{2-}, which is mainly precipitated as insoluble carbonate, are key players in the Carbon Cycle and climate change. Follow the link below to see how increasing CO_2 in the atmosphere due to the production and use of fossil fuels affects the pH of the oceans and their health, and how CO_2 captured by oceans could be used to decrease atmospheric CO_2 and help stop and reverse global warming.

	32.3: Climate Change - Part 1 - The Carbon Cycle and Carbon Chemistry

2.5.5: Summary

Chapter Summary

This chapter explores the critical role of buffering systems in maintaining pH homeostasis within the human body and in laboratory settings. It begins by discussing how blood pH is tightly regulated between 7.35 and 7.45, with deviations leading to metabolic or respiratory acidosis and alkalosis. The chapter emphasizes that this regulation is achieved through buffer systems—combinations of weak acids and their conjugate bases—that resist significant pH changes when small amounts of strong acids or bases are added.

A central tool introduced is the Henderson–Hasselbalch equation, which quantitatively relates pH, pKa, and the ratio of the concentrations of the conjugate base and the weak acid. The equation explains the buffering capacity at pH = pKa and underpins the behavior of titration curves and the concept of buffering ranges.

The chapter then focuses on the carbonic acid/bicarbonate buffering system, a primary mechanism for pH regulation in blood and cells. Through a series of interrelated reactions involving CO_2, H_2CO_3, H_3O^+, and HCO_3^-, the system is shown to rapidly adjust pH. The effective pKa of the system, derived from the ratio of CO_2 and bicarbonate, explains how the buffer is primed to neutralize metabolic acids. At the same time, respiratory and renal adjustments fine-tune the pH balance.

Additional biological buffers, including the phosphate system and proteins, are discussed. Although the phosphate buffer plays a minor role in blood due to its low concentration, it is vital in other cellular contexts. Proteins, especially abundant ones like hemoglobin, also contribute to buffering through their ionizable side chains.

Finally, the chapter provides practical insights into buffer preparation in the laboratory. Various strategies for creating buffered solutions are presented, with attention to selecting buffers based on their pKa values and potential interactions with other ions. The discussion concludes with a broader perspective on how buffer chemistry connects to global issues, such as the role of CO_2 in climate change and carbon capture efforts.

The chapter integrates fundamental chemical principles with biological applications, enabling students to understand and predict how buffer systems stabilize pH in physiological and experimental contexts.

This page titled 2.5: Buffering against pH Changes in Biological Systems is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **2.3: Buffering against pH Changes in Biological Systems** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **4.9: Protein Stability - Thermodynamics** by Henry Jakubowski and Patricia Flatt has no license indicated.

2.6: Solubility in an aqueous world - noncovalent interactions in depth

Learning Goals (ChaptGPT o1, 1/25/25)

Foundational Understanding

1. **Define** and **differentiate** the major types of noncovalent interactions, including:
   - Ion-ion
   - Ion-dipole
   - Dipole-dipole
   - Hydrogen bonds (conventional, short, and very strong)
   - Induced dipole-induced dipole (London dispersion or hydrophobic interactions)
   - Pi-stacking
   - Cation-pi interactions
   - Halogen bonds
   - Nontraditional hydrogen bonds
   - n → π* interactions
2. **Explain** how all noncovalent interactions fundamentally originate from electrostatic forces and are governed by Coulomb's Law.
3. **Distinguish** between the terms "force" and "interaction" in the context of molecular chemistry and physics.

Quantitative and Visual Analysis

4. **Describe** how noncovalent interactions vary in strength, distance dependence, and directionality using numerical data (e.g., kcal/mol, 1/r dependencies).
5. **Interpret** potential energy vs. distance (E vs r) graphs for different interactions and relate them to physical stability and molecular packing.
6. **Compare** the Lennard-Jones potential with electrostatic potential functions and explain how these models capture attractive and repulsive interactions.

Application in Biological Contexts

7. **Identify** examples of each type of noncovalent interaction in biological molecules such as proteins, DNA, and ligand-binding complexes.
8. **Explain** how noncovalent interactions contribute to:
   - Protein folding and stability (e.g., salt bridges, hydrogen bonds, hydrophobic clusters)
   - Ligand binding specificity and affinity
   - Molecular recognition and assembly
   - DNA stability (e.g., pi-stacking)
9. **Distinguish** between intermolecular and intramolecular noncovalent interactions and recognize both in small and large biomolecules.

Advanced Conceptual Integration

10. **Discuss** the relevance of **directionality** and **geometry** in hydrogen bonds, including short hydrogen bonds and proton sharing.

11. **Evaluate** the impact of **solvent environment** (e.g., aqueous solution vs protein interior) on the strength and persistence of noncovalent interactions.
12. **Analyze** electrostatic potential maps and quantum calculations to explain less intuitive phenomena (e.g., σ-holes in halogen bonds, absence of lone pair "rabbit ears").
13. **Relate** modern quantum mechanical models (e.g., MO theory for n → π* interactions) to observed secondary interactions in proteins and ligands.

Critical Thinking and Synthesis

14. **Use molecular visualization tools** (e.g., iCn3D) to explore and annotate examples of noncovalent interactions in protein structures from the Protein Data Bank.
15. **Critically assess** the terminological ambiguity (e.g., "van der Waals forces") in chemical and biochemical literature and adopt IUPAC-recommended definitions.
16. **Predict** how changes in molecular structure (e.g., substitution with halogens, ionization state) might alter the types and strengths of noncovalent interactions in biological systems.

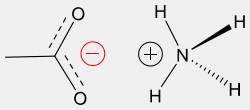
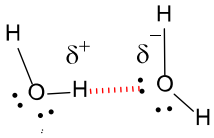
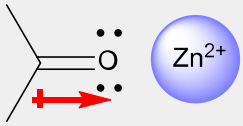
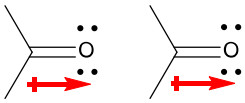
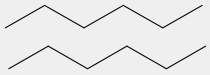
2.6.1: Introduction

In section 2.1, we explored the role of water as a solvent. Using the adage "like dissolves like" that you may have learned in introductory chemistry and biology courses, we can rationalize what substance might dissolve in water. We related this to the types and strengths of attractive interactions between solute and solvent. If, in sum, they are stronger than self-interactions (solute-solute and solvent-solvent), the solute would dissolve (to a reasonable extent) in the solvent. We also discussed entropic contributions to the dissolution process. For now, we will refocus on the noncovalent interactions.

In introductory science courses, noncovalent interactions are often described as intermolecular forces. This term is ambiguous when applied to biochemistry. Take, for example, hydrogen bonds. They occur *between* two water molecules, for example, and *within* larger molecules (like proteins) if hydrogen bond donors and acceptors within the molecule get close enough to each other in space.

The table below summarizes the common noncovalent interactions/“intermolecular forces” you studied in introductory science classes. It is hard enough for students to recognize and identify these interactions between two small molecules, let alone in large molecules like proteins. We will explore these in more detail below and give examples of noncovalent interactions *between* small molecules and *within* large ones, such as proteins. We'll also add a few more specific examples of interactions.

Noncovalent Interactions - "Intermolecular Forces"

Interaction Type	Example	Distance Dependence	Relative Strength Kcal/mol (kJ/mol)	Direction Dependence
Ion-Ion		1/r	60 (250)	nondirectional
H-Bond			3-15 (12-63)	directional
Ion-dipole		$1/r^2$	3-5 (12-21)	directional
Dipole-dipole		$1/r^3$	0.5-1 (2-4)	directional
Induced Dipole-Induced Dipole		$1/r^6$	0.5 (2) (depend on size)	nondirectional

Although there are many noncovalent interactions, one fundamental principle governs them all. They all originate in the electrostatic *force* between two charged objects. There is one simple law, Coulomb's Law, which you have discussed in introductory science courses, and one simple equation that describes the electrostatic force:

$$F = \frac{kQ_1Q_2}{r^2} \tag{2.6.1}$$

F is the force (attractive or repulsive) between two particles of charge Q_1 and Q_2 with their centers separated by some distance r. Replace the charges with the masses of two objects, and you will have Newton's Law of Gravitation. Both are inverse-squared laws.

All the *interactions* described in the table above arise from the electrostatic force. The magnitude of the attraction depends on the distribution of charge within the attracting species. Each interaction has different dependencies on distance.

Different words are used to describe noncovalent interactions. This can be distressing for learners who might hear chemists and biologists use different terms for the same noncovalent interactions. Some use van der Waals forces to describe induced-dipole-induced-dipole interactions, while others use London dispersion forces or hydrophobic interactions. Others use van der Waals forces to describe *all* noncovalent interactions, except ion-ion interactions. To avoid ambiguity, consult the IUPAC Gold Book, Compendium of Chemical Terminology, which defines van der Waals forces.

Definition: van der Waals Forces

"The attractive or repulsive forces between molecular entities (or between groups within the same molecular entity) other than those due to bond formation or to the electrostatic interaction of ions or ionic groups with one another or with neutral molecules. The term includes dipole–dipole, dipole-induced dipole, and London (instantaneous induced dipole-induced dipole) forces. The term is sometimes used loosely for the totality of nonspecific attractive or repulsive intermolecular forces". IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson.

Blackwell Scientific Publications, Oxford (1997). Online version (2019-) created by S. J. Chalk. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook.

Figure 2.6.1 summarizes covalent and noncovalent interactions, using that definition.

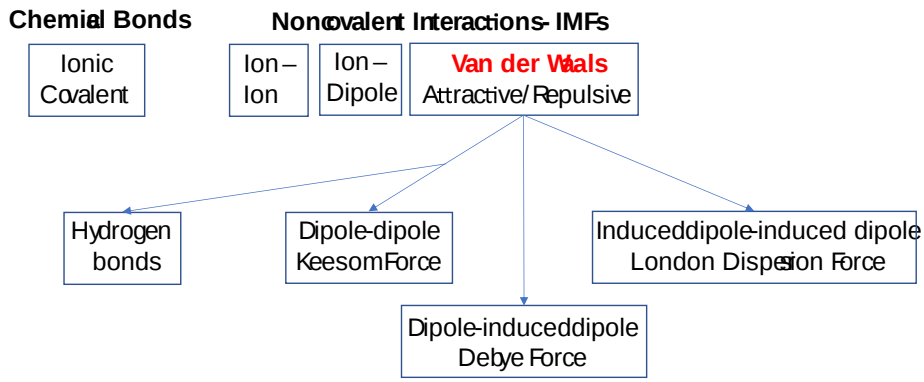


Figure 2.6.1: Noncovalent Interactions

Hydrogen bonds are usually considered a type of dipole-dipole interaction using this definition. Historically, several noncovalent interactions have alternative names based on the person associated with them. Only the names van der Waal and London are commonly used in introductory chemistry courses

Even the word "force" is also used loosely. To a physicist, there are only four known forces:

- gravitational, between two objects with mass
- electromagnetic (between static charges - the electrostatic force, and moving charges - the magnetic force)
- the strong force (holding the nucleus together)
- the weak force (also nuclear and involved in radioactive decay).

We'll try to use the word interaction throughout this book.

Interactions *within* small molecules, such as covalent bonds, and *between* molecules, such as induced dipole-induced dipole, vary as some function of r, the distance between the two interacting particles. Only ion-ion interactions vary as $1/r^2$ ($F= k/r^2$). Attractions lower the overall energy, while repulsions raise it. The system is in its most favored, lowest energy state at some optimal distance.

We just switched from discussing forces to discussing energy, another complex term. The relationship between the potential energy for covalent bond formation and the noncovalent attraction of two atoms as a function of distance is shown in general form in Figure 2.6.2 below.

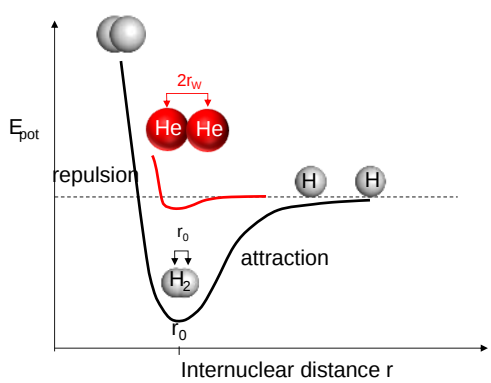


Figure 2.6.2: Potential Energy for Covalent and Noncovalent interactions

The black curve shows the shape of E_{pot} vs. r for forming a covalent bond between H atoms. The Morse potential energy function models energy as a function of r for simple diatomic molecules. The red line shows the shape of E_{pot} vs. r for the noncovalent attraction of two He atoms through induced dipole-induced dipole interactions. It is modeled using the Lennard-Jones (6-12) potential function (see below). Each has an optimal r_0 (the bond length for H_2 and two times the van der Waals radius, r_W, of each He in 2He). The energy required to break the induced dipole-induced dipole interactions between He atoms is very small, which accounts for the fact that liquid He, in which many He atoms are interacting, only exists at very cold temperatures (boiling point = -269 °C).

Although the graph for H_2 shows the relationship between the potential energy and r_0 for the covalent bond, in reality, the sources of stability of any covalent bond are complex and require, in addition, a term for the kinetic energy of the electron. Fundamentally, the strength of a covalent bond is best described using quantum wave functions for the system. The average single covalent bond strength depends on the atoms bonded and varies between 30-120 kcal/mol (125-500 kJ/mol), a factor of 4.

Another confusing feature in discussing noncovalent interactions is that, when discussing forces (like electrostatic forces), we often plot energy E vs. r, the distance between two interacting particles. Let's briefly examine the relationship between potential energy (E_{pot}) and force for the electrostatic force given by Coulomb's Law by using a more familiar example, the next gravitational force of a stationary ball placed at various points on a hill, as illustrated in Figure 2.6.3.

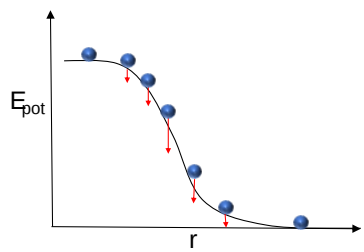


Figure 2.6.3: Potential Energy vs. r for a ball on a hill

Assume that the ball is motionless at each position in the diagram, and only potential energy is considered. The red arrows (vectors) represent the relative *net* downward force on the ball at each position. The *net* downward force at the top and bottom of the hill is zero. As the *slope* of the hill increases, the net downward force increases. The force is directly proportional to the slope (dE/dr), or simply:

$$F = -\frac{\Delta \mathrm{E}}{\Delta \mathrm{r}} = -\frac{\mathrm{dE}}{\mathrm{dr}} \tag{2.6.2}$$

Now let's apply this same relationship to Coulomb's Law for the force. Rearranging gives

$$dE = -\mathrm{Fdr} = -\frac{\mathrm{kq_1q_2}}{\mathrm{r}^2}\mathrm{dr} \tag{2.6.3}$$

Using calculus and integrating both sides of the equations gives this general relationship between E and r for the electrostatic forces:

$$E = \mathrm{kq_1q_2}\left(\frac{1}{\mathrm{r}}\right) \tag{2.6.4}$$

A graph of E_{pot} vs r for the electrostatic force is shown in Figure 2.6.4. Note that the curves are hyperbolic (1/r) functions of r. There are attractive OR repulsive components.

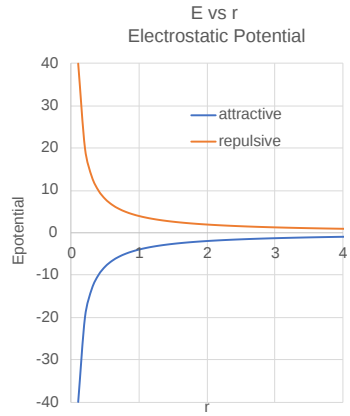


Figure 2.6.4: Electrostatic potential vs distance r

An equation for E_{pot} vs r for induced dipole-induced-dipole interactions can also be derived. For this interaction, E_{pot} depends on r differently. It has both an attractive (E_{pot} α $-1/r^6$) AND a repulsive term (E_{pot} α $+1/r^{12}$), which are added together. This potential is called the Lennard-Jones or 6-12 potential. Figure 2.6.5 shows the total and component attractive and repulsive terms E_{pot} vs r. Note how similar these curves are to the graphs for electrostatic energy.

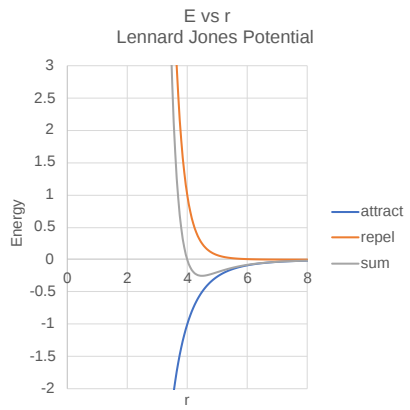


Figure 2.6.5: Lennard Jones Potential

Now, let's look at noncovalent interactions more carefully, using examples of both small and large molecules.

2.6.2: Ion-Ion

All introductory chemistry and biology textbooks differentiate between ionic and covalent bonding. Ionic bonding occurs between fully charged species. Some ions are monatomic (like Na^+ or Cl^-), formed from gaining or losing electrons. Others are polyatomic (like ammonium - NH_4^+ or acetate - CH_3COO^-, generally formed from molecules gaining or losing protons in Brønsted acid/base reactions. Polyatomic ions are also called molecular ions. An example of the monatomic salt NaCl and the molecular salt ammonium acetate are shown in 2D Lewis structure and molecular modeling representations (spheres and sticks) in Figure 2.6.6 below.

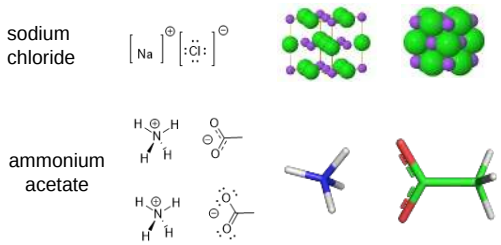


Figure **2.6.6**: Representations of monatomic and polyatomic/molecular salts

Now, an intramolecular ion-ion interaction can form within a larger molecule if a negatively charged group in the molecule comes close enough in 3D space to a positively charged group in the same molecule. In contrast to the above examples, ion-ion interactions within large molecules like proteins are not the same as those within a large lattice of ions held together by numerous similar ionic bonds. Rather, a single ion-ion interaction could persist within a larger molecule held together by many other noncovalent interactions. An ionic bond between a single monatomic or polyatomic cation and an anion would not exist in an aqueous solution as long as the species would dissociate into separate ions solvated by water. Hence, the ion-ion interaction between charged groups within a large molecule, like a protein, exists in a different environment than a solid crystal lattice. We often give it a different name, a ***salt bridge***, as the ion-ion interaction bridges the distal parts of the larger molecule. We also categorize it as an **ion-ion** noncovalent attraction.

Figure **2.6.7** shows a salt bridge/ion-ion interaction (represented as a yellow line) between the side chains of two amino acids, aspartic acid (Asp) 67 ($-CH_2COO^-$, similar to acetate) and lysine (Lys) 69 ($-RCH_2NH_3^+$, similar to NH_4^+) in a protein, human lysozyme.

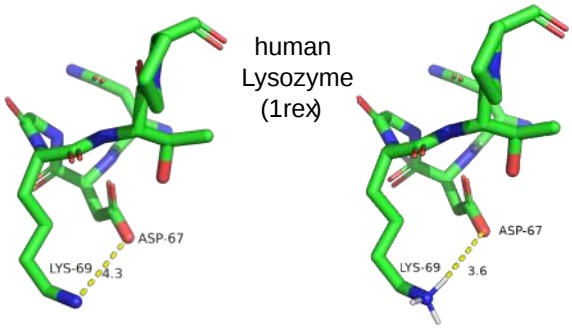
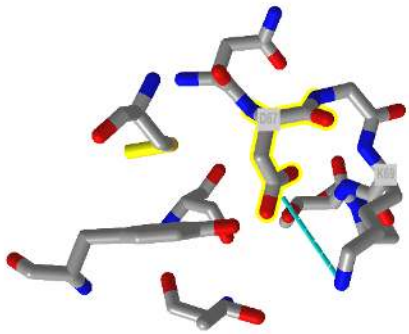


Figure **2.6.7**: Salt bridge in human lysozyme

Figure **2.6.8** shows an interactive iCn3D model of a salt bridge/ion-ion interaction between the carboxylate side chain of Asp 67 and the amine side chain of Lys 69 in human lysozyme (1REX).



NCBI iCn3D Figure **2.6.8**: Salt bridge (represented as a yellow line) between Asp 67 and Lys 69 in human lysozyme (1REX). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1qpAtSs3CuvVs8

Most of the protein's atoms have been removed to simplify the structure. We haven't studied proteins yet, but to a first approximation, they are polymers consisting of amino acid monomers. The polymer's backbone contains a repeating amide group with an N-H hydrogen bond donor and a C=O hydrogen bond acceptor. Each amino acid contains an R group side chain oriented away from the backbone. The R groups can be fully charged, polar, or nonpolar.

This protein, containing 129 amino acids and a large polymer of over 1000 atoms, has just 10 salt bridges/ion-ion interactions within its most stable structure. The structure files that contain the x,y, and z coordinates of the atoms in a large biomacromolecule, like a protein, usually don't give coordinates for hydrogen atoms in the structure since they are too small to detect by techniques such as X-ray crystallography or cryoelectron microscopy, which are used to determine the structure of large biomacromolecules. Computer programs can be used to add them so they can be visualized in modeling programs. The left molecule in Figure 2.4.7 shows a stick model of a small protein part containing a single salt bridge/ion-ion interaction. The blue represents nitrogen with a +1 formal charge in the side chain of lysine.

Hydrogen atoms have been added to the right molecule to illustrate the distance between adjacent atoms. Quantum calculations of actual electron density in molecular ions such as H_3O^+ and NH_4^+ (and charged amines) show that the electron density in these cations is shifted to the electronegative O and N atoms with electron deficiencies over the bonded H atoms (in contrast to the simpler ideal of formal charge), even though we state that the N in a charged amine has a positive formal charge.

The attractions decrease as the distance r between interacting groups increases past the optimal interaction distance. When modeling most noncovalent interactions in large molecules, programs generally use cutoff values of 5-6 Angstroms, beyond which the interactions do not contribute to stabilization. Given a fixed distance for comparison, the ion-ion interaction is the strongest.

2.6.3: Hydrogen Bond (H-bond)

The term hydrogen bond is somewhat ambiguous, leading students to misunderstand it. It is **not** a covalent bond between two atoms, X and H, such as C-H and O-H in methane and water, respectively. Rather, it is a noncovalent interaction between a slightly positive H on an electronegative atom X and a slightly negative electronegative atom Y on another molecule or part of a large molecule. X and Y are electronegative atoms, such as F, O, or N, with lone pairs. The H on X-H (for example, O-H or N-H) is slightly positive (δ^+) since the X-H bond is polar covalent, and electron density in the bond is drawn toward the electronegative atom (for example, O or N). Given its small size compared to all other atoms, a slightly positive H can get very close to a lone pair on another molecule's slightly negative (δ-) electronegative atom Y (for example, O or N). Since r, the distance between the δ^+ H and δ^- N or O on two separate molecules is small, Coulomb's Law informs us that the attractive force is significant. This interaction is highly directional and distance-dependent, which accounts for the large range in relative strength (3-15 kcal/mol, 12-63 kJ/mol) for hydrogen bonds within large molecules.

Hydrogen bonds occur between hydrogen bond donors and acceptors. This is determined by looking at the slightly negative electronegative atoms in the two interacting molecules. An alcohol (ROH) can be either a hydrogen bond donor or acceptor, while a ketone (R)C=OR can only be an acceptor since it has no slightly positive H. In a hydrogen bond between an alcohol and a ketone, the O-H on the alcohol is the hydrogen bond **donor,** while the O=C on the ketone is the hydrogen bond **acceptor**. This is illustrated in Figure 2.6.9 below.

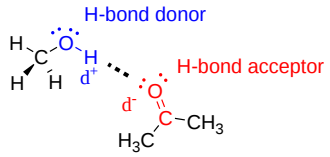


Figure 2.6.9: A hydrogen bond between a donor and acceptor

In a given hydrogen bond, the donor is the X-H with the slightly positive H.

Figure 2.6.10 shows multiple representations of a central water molecule hydrogen-bonded to four other water molecules. The left image shows lone pairs as purple spheres.

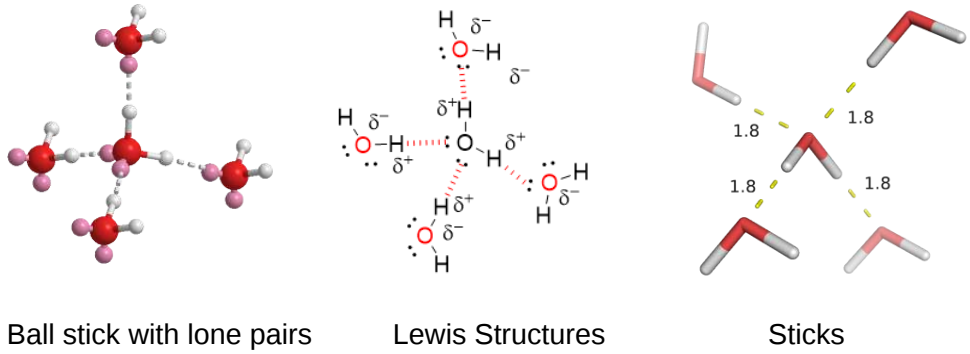


Figure 2.6.10: Multiple representations of hydrogen bonding among 5 water molecules

A common difficulty for students is identifying which hydrogen atoms in a given structure can form hydrogen bonds. One way is to circle all δ^+ Hs in structures (i.e., those covalently attached to N or O) and see if there are any nearby δ^-: N or :O atoms close enough to form a hydrogen bond. Figure 2.6.11 shows a methanol molecule forming two hydrogen bonds with two different water molecules. Only 1 of the 4 Hs on methanol is δ^+ (circled in green). The others are bonded to the carbon through nonpolar covalent bonds.

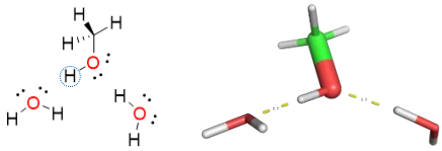


Figure 2.6.11: Hydrogen bond between methanol and waters

Hydrogen bonds are abundant in large molecules, such as proteins. They occur between backbone atoms, backbone and side chain atoms, side-chain atoms, and protein atoms with water. Their strength depends on the magnitudes of the δ+ and δ- charges on the H-bond donor and acceptor atoms, respectively, the distance r between them, and the bond angle. Three types of H bonds have been categorized based on their relative strengths, based in large part on the distance between the donor and acceptor:

- conventional (weakest of these three, so sometimes called weak), 2.4 to 12 kcal/mol (10-50 kJ/mol)
- strong or low barrier, 12 to 24 kcal/mol (50-100 kJ/mol), often called **short hydrogen bonds (SHB)**
- very strong or no barrier >24 kcal/mol (100 kJ/mol), (Frey et al).

In large proteins with known 3D structures, conventional H bonds are calculated by locating all donors and acceptors within 3 +/- 0.2 Å of each other. Most structural files do not include H atoms, so the 3 Angstrom distance is measured from the centers of the electronegative atoms, typically N and O, involved in the hydrogen bond, as shown in Figure 2.6.12 (purple bracket).

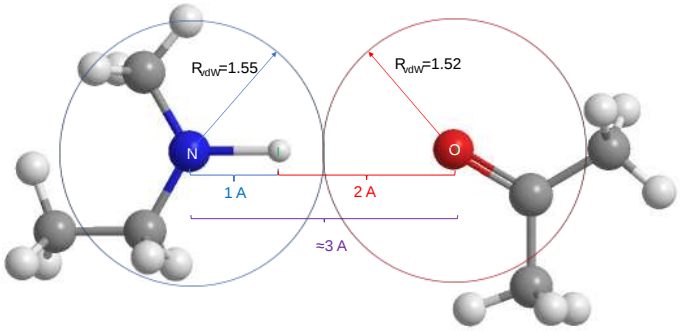


Figure 2.6.12: Hydrogen bonding distances between the center of a participating N with a H donor and an O with a H bond acceptor

Conventional H bonds vary between 2.8-3.2 Å, which gives a distance range from the actual δ^+ hydrogen on the donor to the acceptor δ^- N or O (the red line below) of 1.8 to 2.2 A. Short H bonds are < 2.7 Å, which is *smaller than the sum of the van der Waals radii of N and O* (blue and red circles above), suggesting that the bond has a covalent character (see below). Those between 2.5 and 2.7 A are strong, low-barrier, or short hydrogen bonds. Analysis of a large number of Protein Data Bank (PDB) structures of proteins shows many short hydrogen bonds characterized by these properties:

- the donor and acceptor electronegative atoms A and B are N or O
- r, the separation distance, is 2.3 A to 2.7 A
- the A–H–B angle is 135 °.

Detailed analyses of high-quality protein structures show one short hydrogen bond for every 16 conventional ones. They are found in proteins, protein-ligand complexes, and in DNA. They are involved in many aspects of molecular function.

It would seem likely that the δ^+ H atom, which is covalently attached to a heteroatom like O or N (A), and which is attracted to another heteroatom B, could be exchanged between the two heteroatoms as shown in the chemical equation below, where ---- represents an H bond.

A-H ----B ↔ A ----H-B

A very strong/no barrier H bond occurs if A and B are very close, have similar δ^- charges, and have similar pKas so that the H atom could be equally shared between A and B. It is represented by the representation below.

A ||| H ||| B

An example is FHF^- $(F|||\ H\ |||F)^-$ in which there is no barrier for H to move from one heteroatom to another.

Thus, strong and very strong hydrogen bonds have some covalent bond character.. Quantum calculations show an overlap between the unoccupied antibonding σ* molecular orbital of X-H (the hydrogen donor) and the non-bonding lone electron pair molecular orbital of the hydrogen bond acceptor molecule.

Even though water is a simple and ubiquitous molecule, scientists still struggle to understand its properties. Lewis structures for water can explain only so much of its physical and chemical properties. However, look at Figure **2.6.13**, which shows the electron density around water calculated using quantum theory.

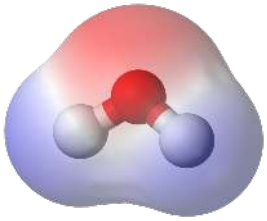


Figure **2.6.13**: Electron density around a water molecule. https://commons.wikimedia.org/wiki/F...balls.png#file

Do you see any "rabbit ears" (i.e., lone pairs) emanating from the oxygen atom? Don't think so! Nevertheless, everyone still uses Lewis structures with lone pairs to explain the chemistry of water and other molecules. We present this figure before a discussion at the end of this section on the halogen bond, which requires understanding the electron density around bonded atoms.

Now, let's look at some hydrogen bonds within a single protein molecule. Figure **2.6.14** hydrogen bonds (yellow dotted line) between serine (Ser) 24 (side chain $-CH_2OH$) and asparagine (Asn) 27 (side chain $-CH_2(C{=}O)NH_2$ of hen egg white lysozyme (1REX). As in the figures above showing salt bridges, two images are shown, one with polar H atoms added. Find the hydrogen bonds between side chains, side chains and backbone, and between backbone hydrogen bond donors and acceptors.

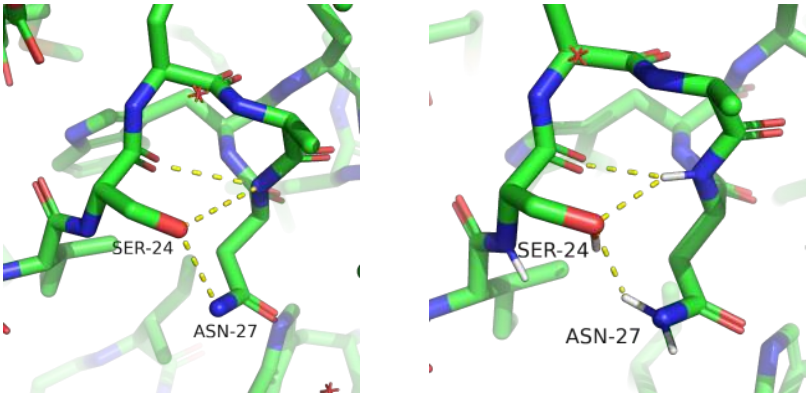


Figure 2.6.14: More hydrogen bonds in hen egg white lysozome (1REX)

Proteopedia has an excellent review of hydrogen bonds.

2.6.4: Hydrogen Bond or Ion-Ion Interaction

If a negatively-charged carboxyl group (such as the side chain of aspartic or glutamic acid - Asp, Glu) is close in proximity to a positively-charged nitrogen group (quaternary amine such as the protonated side chain of lysine - Lys, or a guanidinium group such as in the side chain of arginine - Arg) in a folded protein, it would seem a no-brainer to name it as an ion-ion interaction. However, the research literature abounds with examples that describe them as hydrogen bonds. Which is true? Effectively, both are, but an example from the literature might clarify why the term' hydrogen bonds' is often used among structural biologists.

Let's look at a noncovalent interaction diagram from a protein called lanmodulin, shown in Figure 2.6.*i* below.

Figure 2.6.*i*: J.A., Jung, J.J., Lin, CY. *et al*. Enhanced rare-earth separation with a metal-sensitive lanmodulin dimer. *Nature* **618**, 87–93 (2023). https://doi.org/10.1038/s41586-023-05945-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Only one resonance structure is shown in the right-hand figure. The green sphere is an ion (in this case, La^{3+}). The amino acid Arg100 has a positively charged guanidinium side chain that interacts with negatively charged carboxyl groups on Asp85, Asp93, and Glu91. The paper shows the interactions as dotted lines and refers to them as hydrogen bonds. Why?

The reasons are that the interactions are short (≤ 3.2 Å), directional (linear, non-bent between aligned lone pairs and δ^{+}Hs), and constrained by the geometry of the interacting atoms/ions. If the interacting carboxylate and guanidinium groups were farther apart, the interactions would be called ion-ion or salt bridges. Hydrogen bonds have optimal distances and geometries that are usually met in the interactions of a folded protein. The protein folds to maximize hydrogen-bonding interactions.

As described above, hydrogen bonding arises from electrostatic interactions and is a subset of those interactions. Their orientation and distance are set, in part, by ion-ion interactions that attract the side chains in both the folding and the folded states. The term **charge-assisted hydrogen bond (CAHB)** has been used to describe these types of interactions in proteins, especially among physical and structural chemists. These are likely stronger than "neutral" hydrogen bonds and common in folded proteins. The term is not widely used in biochemistry, but it helps clarify the terminology for distinguishing between biological ion-ion and hydrogen-bond interactions.

For a structural biologist, the word ion-ion probably does not adequately describe the observed interactions. Writers could choose between ion-ion, salt bridges, or hydrogen bonds based on what they wish to emphasize. If the charged groups are interacting over a long distance, the term ion-ion is best; if they are close and meet the geometric constraints of a hydrogen bond, then that term is likely best. Figure $2.6.ii$ below illustrates different ways to name the same effective interaction.

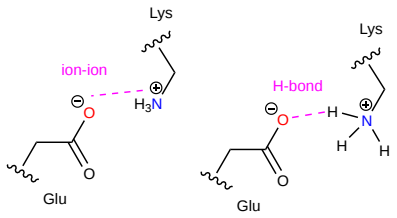


Figure $2.6.ii$: Different ways to present and emphasize noncovalent interactions between a positive and negative group in close approximation.

Is it an ion-ion interaction or hydrogen bond? - An example from the literature

If a negatively-charged carboxyl group (such as the side chain of aspartic or glutamic acid - Asp, Glu) is close in proximity to a positively-charged nitrogen group (quaternary amine such as the protonated side chain of lysine - Lys, or a guanidinium group such as in the side chain of arginine - Arg) in a folded protein, it would seem a no-brainer to name it as an ion-ion interaction. However, the research literature abounds with examples that describe them as hydrogen bonds. Which is true? Effectively, both are, but an example from the literature might clarify why hydrogen bonds if often used among structural biologists.

Let's look at a noncovalent interaction diagram from a protein called lanmodulin, shown in Figure $2.6.i$ below.

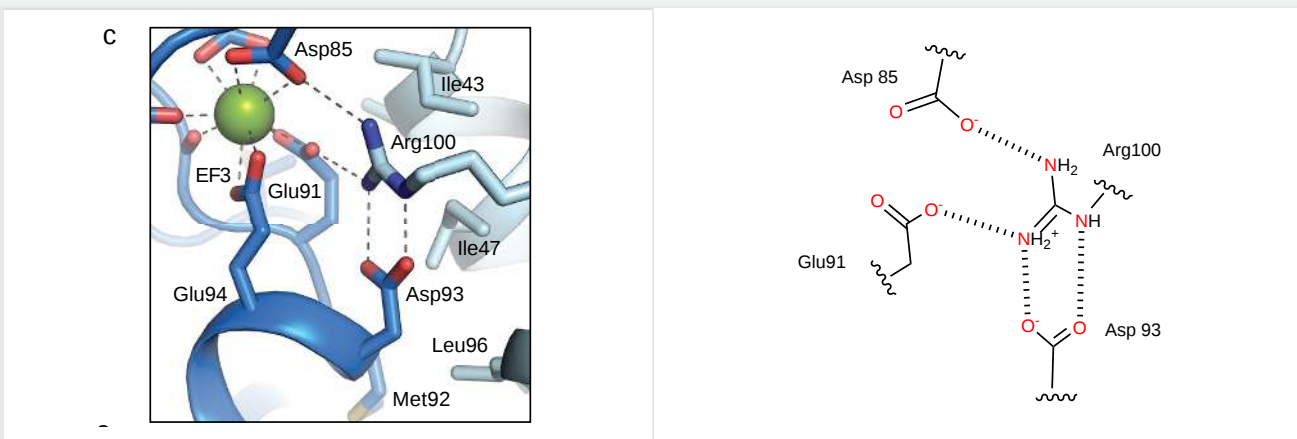


Figure $2.6.i$: J.A., Jung, J.J., Lin, CY. *et al.* Enhanced rare-earth separation with a metal-sensitive lanmodulin dimer. *Nature* **618**, 87–93 (2023). https://doi.org/10.1038/s41586-023-05945-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Only one resonance structure is shown in the right-hand figure. The green sphere is an ion (in this case, La^{3+}). The amino acid Arg100 has a positively charged guanidinium side chain that interacts with negatively charged carboxyl groups on Asp85, Asp93, and Glu91. The paper shows the interactions as dotted lines and refers to them as hydrogen bonds. Why?

The reasons are that the interactions are short (≤ 3.2 Å), directional (linear, non-bent between aligned lone pairs and δ^+Hs), and constrained by the geometry of the interacting atoms/ions. If the interacting carboxylate and guanidinium groups were farther apart, the interactions would be called ion-ion or salt bridges. Hydrogen bonds have optimal distances and geometries that are usually met in the interactions of a folded protein. The protein folds to maximize hydrogen-bonding interactions.

As described above, hydrogen bonding arises from electrostatic interactions and is a subset of those interactions. Their orientation and distance are set, in part, by ion-ion interactions that attract the side chains in both the folding and the folded states. The term **charge-assisted hydrogen bond (CAHB)** has been used to describe these types of interactions in proteins, especially among physical and structural chemists. These are likely stronger than "neutral" hydrogen bonds and common in folded proteins. The term is not widely used in biochemistry, but it helps clarify the terminology we use to differentiate between biological ion-ion and hydrogen-bond interactions.

For a structural biologist, the word ion-ion probably does not adequately describe the observed interactions. Writers could choose between ion-ion, salt bridges, or hydrogen bonds based on what they wish to emphasize. If the charged groups are interacting over a long distance, the term ion-ion is best; if they are close and meet the geometric constraints of a hydrogen bond, then that term is likely best. Figure **2.6.*ii*** below illustrates different ways to name the same effective interaction.

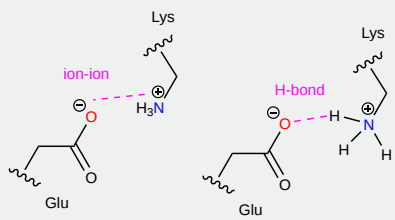


Figure **2.6.*ii***: Different ways to present and emphasize noncovalent interactions between a positive and negative group in close approximation.

2.6.5: Dipole-Dipole

This interaction involves the alignment of permanent dipoles in molecules so that the geometric center of the δ+ of one permanent dipole on one molecule is close to and aligned with the geometric center of the δ- of another permanent dipole. Figure **2.6.15** shows two acetone molecules interacting through dipole-dipole interactions. These molecules can't form hydrogen bonds with each other because they both contain only hydrogen-bond acceptors.

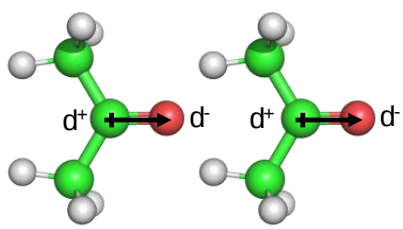


Figure **2.6.15**: Attractions of two acetones through dipole-dipole interactions

The arrow represents the *molecular* dipole moment vector (as opposed to the individual bond dipole moment for each polar covalent bond in the molecule). Note the difference in Figure **2.6.16**. The molecular dipole is the vector sum of the bond dipoles.

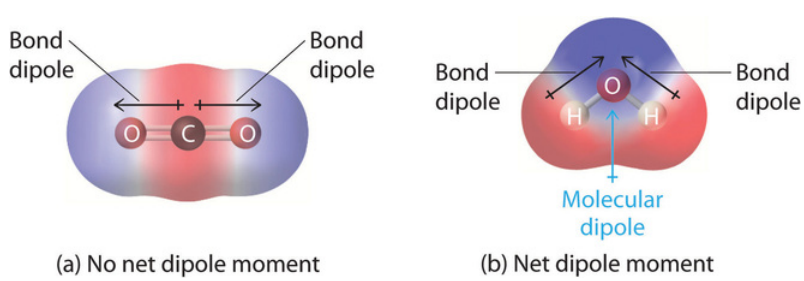


Figure **2.6.16**: Difference between bond and molecular dipoles. https://chem.libretexts.org/Bookshel...bond%20dipoles.

None of the H atoms bonded to carbon in acetone are $\delta^{+,}$ so the molecules contain no H bond donors. Although they contain a δ^{-} oxygen, a hydrogen bond acceptor, two acetone molecules cannot hydrogen bond to each other. However, they can form hydrogen bonds with water. Pure liquid acetone evaporates readily (BP 56^0 C) due to this lack of strong hydrogen bonds.

You can imagine two water molecules forming dipole-dipole interactions as well. However, tilting the molecule to align the lone pair on an O with the δ^{+}H on another water molecule, and presto, you have a hydrogen bond. H bonds are often viewed as a special case of a dipole-dipole interaction.

Modeling programs can calculate the charge on each atom of a large molecule, such as a protein, and determine the geometric center and magnitude of the overall + and - charge. A line between them represents the entire protein's permanent "dipole" moment. More simply, the molecular dipole is the vector sum of all of the individual bond dipole moments. Entire proteins have a net dipole moment, which probably facilitates their interactions with other proteins or ligands. Figure **2.6.17** shows the net dipole moment for the protein carboxypeptidase A1 (2v77). This was calculated using the Protein Dipole Moments Server. Proteins, however, do have net charges (not considering any bound counterions), so the molecular dipole for a protein is a bit different conceptually from that of a small molecule. Nevertheless, it is a good way to quantify asymmetric charge distribution in large biomolecules. Asymmetric charge distributions would influence molecular properties.

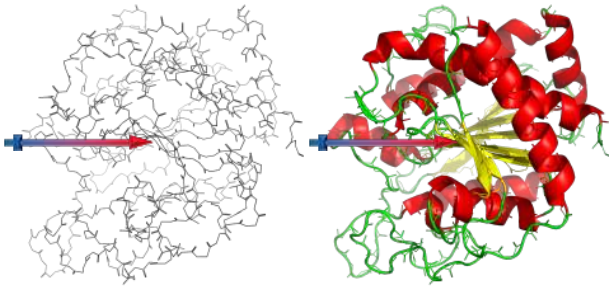


Figure **2.6.17**: Net dipole moment of the protein carboxypeptidase A1 (2v77)

2.6.6: Ion-Dipole

Figure **2.6.18** shows interactions between a Na^+ ion and the dipoles of multiple water molecules.

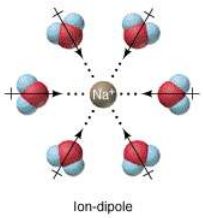


Figure **2.6.18**: Ion-dipole interactions between Na^+ and water. https://chem.libretexts.org/Courses/...-Dipole_Forces

Figure **2.6.19** shows an interactive iCn3D model of the molecular ion sulfate SO_4^{2-} bound to a protein through its hydrogen bonding and ion-dipole noncovalent interactions with protein side chain and backbone groups in the sulfate binding protein from Salmonella typhimurium.

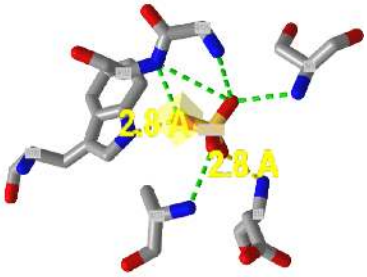


Figure 2.6.19: Sulfate bound to sulfate binding protein from S. typhimurium (1sbp). Hydrogen bonds are shown as green and yellow lines. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...JvqQuZZnvf3AU8

The SO_4^{2-} is buried within the protein. The green dotted lines show hydrogen bonds between the sulfate and the amide N-Hs on the protein chain surrounding it, as well as between the sulfate and the protein's side chains. Modeling programs don't show lines depicting dipole-x interactions. The SO_4^{2-}, through its oxygen, can form hydrogen bonds with nearby donors.

Figure 2.6.20 shows an interactive iCn3D model of another example of a protein backbone and side chains ion-dipole interactions, this time with a Na^+ ion, a simple non-transition state metal ion, which can not form hydrogen bonds. The protein is tryptophan synthase from Salmonella typhimurium (6dz4). The red spheres represent water oxygen atoms (no hydrogen atoms shown).

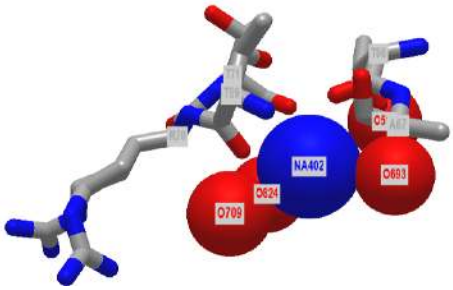


Figure 2.6.20: Sodium ion binding site in tryptophan synthase from Salmonella typhimurium (6dz4). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pDnU36vXXPeyX7

The ions illustrated in these last two cases are **not** transition metal ions, whose interactions with ligands can best be described using ligand field theory and the formation of covalent (dative) bonds between electron-pair donors on nucleophilic side-chain or main-chain atoms and d orbitals on the transition metal.

2.6.7: Induced Dipole - Induced Dipole/Hydrophobic Interactions.

These noncovalent interactions occur when a temporary dipole, created by random fluctuations in electron density in one molecule, induces a temporary dipole in a nearby molecule. These interactions are weak and can be easily disrupted by raising the temperature. Induced dipole-induced dipole interactions allow nonpolar gases like He, N_2, O_2, and CH_4 to be liquefied, but only at higher pressures and/or low temperatures to force the molecules close enough and slow them down enough for sufficient interactions to occur to liquefy the molecules. Although individually weak, the larger the molecules, the greater the extent of induced dipole-induced dipole interactions and the stronger the interactions among molecules. This is reflected in the fact that methane (CH_4) is a gas at room temperature, octane (C_8H_{18}) is a liquid, and $C_{30}H_{62}$ is a solid.

Figure 2.6.21 shows induced dipole interactions between two molecules.

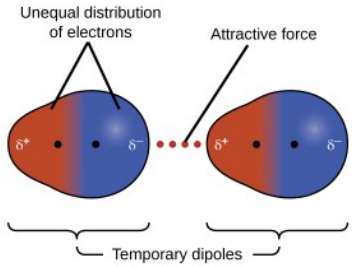


Figure 2.6.21: Induced dipole-induced dipole interactions. https://boisestate.pressbooks.pub/ch...ecular-forces/

Induced dipole-induced dipole interactions are important among large biomolecules as well. Most biologists and probably biochemists prefer to use the words hydrophobic interactions (but not hydrophobic forces) instead of the longer and more formal induced dipole-induced dipole interaction. We will also try to use the more commonly used term within the biochemistry community.

Figure 2.6.22 shows an interactive iCn3D model of a hydrophobic cluster around the side chain of a hydrophobic amino acid, valine 143, in human **carbonic anhydrase II (4ca2)**. Val 143 is highlighted in yellow and shown with normal atom (CPK) colors. White to green indicates nonpolar amino acids, while dark blue indicates polar ones.

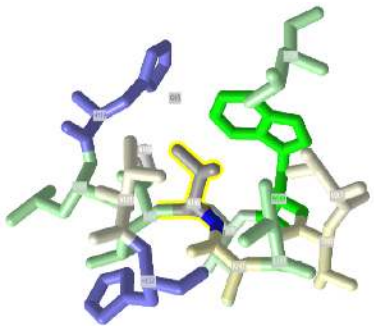


Figure 2.6.22: Hydrophobic cluster around the side chain of a hydrophobic amino acid, valine 143 in human carbonic anhydrase II (4ca2). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...6rgjajJsg3Xvg8

You can see that the side chain of Val 143 (highlighted in yellow) is surrounded by nonpolar amino acids. If the structure were rendered in spacefill instead of sticks, Val 143 would be closely packed to maximize induced dipole-induced dipole (hydrophobic)interactions.

Induced dipole-induced dipole interactions also occur between polar molecules, but they are weaker than the hydrogen bonding and dipole-dipole interactions between them.

2.6.8: Pi stacking

Aromatic rings stacked over each other can interact through induced dipole -induced dipole (hydrophobic) and dipole-induced dipole interactions. These interactions can depend on the presence of heteroatoms in the aromatic ring. Figure 2.6.23 shows an example with benzene in which a staggered arrangement of the rings is more attractive.

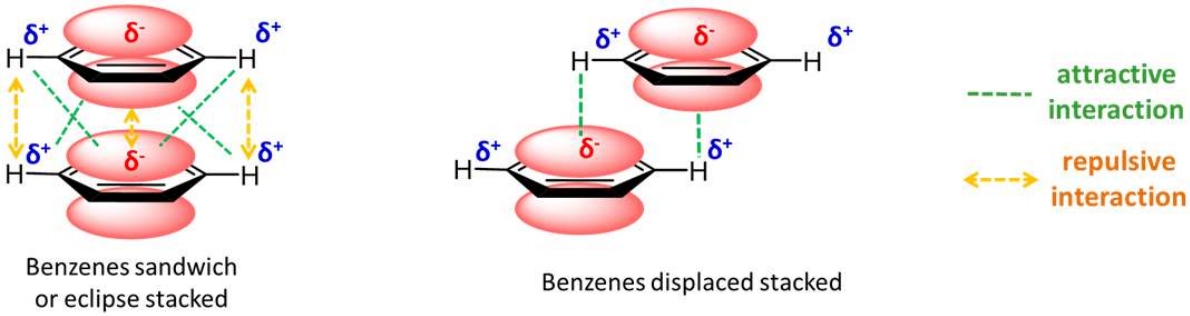


Figure 2.6.23: Pi stacking in benzene. https://chem.libretexts.org/Bookshel...g_interactions

For a biological example, everyone is familiar with the structure of B-DNA in which the bases A, G, C, and T point inward, perpendicular to the double helix axis, and are stacked over each other.

Figure 2.6.24 shows an interactive iCn3D model of a short stretch of DNA with a sugar-phosphate backbone and bases colored in magenta and cyan. Five bases on one strand are shown in stick and atomic color to show the pi-stacking interactions of the aromatic ring.

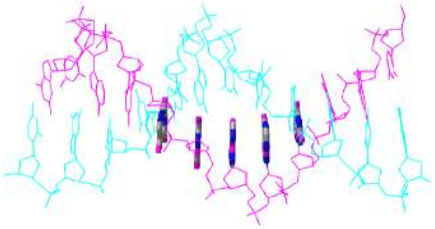


Figure 2.6.24: Pi stacking in B-DNA (5t4w). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...euHaVfXwk18JJ8

Pi stacking also occurs in proteins. Figure 2.6.25 shows an interactive iCn3D model of two sets of pi stacking interactions in the protein arginine kinase (1M15). The aromatic side chains involved in pi stacking are shown in cyan.

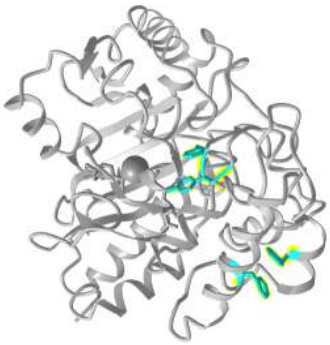


Figure 2.6.25: Pi stacking in Arginine Kinase (1M15). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...owSDJ2DUqtwk8A

2.6.9: Cation - Pi

Figure 2.6.26 shows an interactive iCn3D model of a specific example of an ion-induced dipole interaction (called a cation-pi interaction) between a sodium ion (blue sphere) and the aromatic ring of the side chain tryptophan (cyan) in hen egg white lysozyme (1lpi).

Figure 2.6.26: Cation-Pi stacking (ion-induced dipole) in hen egg white lysozyme (1lpi). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...13gw1hmVYrBL8A

Example 2.6.1

For another example of a cation-pi interaction, open up iCn3D with 1REX and view the interaction of lysine (K1) side chain with the nonpolar aromatic ring of phenylalanine (F3).

Solution

https://structure.ncbi.nlm.nih.gov/i...4d3wXXsSEYhgv7

Here are some more examples.

Exercise 2.6.1

Select the link below to answer the following questions.

1. What type of noncovalent interaction best describes the red dotted line in the structure?

2. What type of noncovalent interaction best describes the red dotted line in the structure?

Answer

1. cation-pi
2. pi stacking

2.6.10: Halogen Bond

Lastly, we come to the halogen bond. You might ask whether halogens are found in proteins. The answer is no (until one is found!), but halogenated molecules (drugs, xenobiotics, toxins) bind proteins. Consider the C-X bond where X is a halogen. The electronegativity of C is 2.56, while the halogens have these electronegativity values: F (3.98), Cl (3.16), Br (2.96), and I (2.66). Compare these to oxygen (3.44) and N (3.04). Covalent bonds between two bonded atoms whose electronegativity differences are between 0.4 and 1.8 are considered polar covalent, so C-F, C-Cl, and C-Br are considered polar covalent. The C-I bond is the longest, and iodine is the most polarizable of these halogens. An alkyl halide with a C-I bond can undergo S_N2 nucleophilic substitution reactions, with I^- an excellent leaving group. Hence, the C-I bond behaves somewhat as a polar covalent bond.

Nevertheless, quantum calculations show that the electron density is not uniformly distributed around the X atom in a C-X bond but is pulled more toward the C, leaving the distal *end* of the halogen depleted in electron density and slightly positive. This region of relatively depleted electron density is called the σ-hole. Color-coded renderings of the electron density of the halogen involved in a C-X bond show the halogen atom to have bands (like Jupiter) with the more negative electrostatic potential (represented in blue) closest to C and the more positive potential, the σ-hole (represented in red), at the end farthest from the C atom. Calculations show that this effect is greatest for the heavier halogens (Br, I) with longer C-X bonds. The halogen's slightly positive σ-hole can act analogously to a hydrogen bond donor in its interactions with nearby δ^- :O and :N atoms/lone pairs. This might take a while to grasp. You have always heard that, in general, the halogens are more electronegative than carbon and would hence always be δ^- when bonded to it. This case is similar to our chemical intuition about lone-pair "rabbit ears" on oxygen, which quantum calculations show is not an accurate representation of the electron density (see Fig. 2.4.13).

Figure 2.6.27 shows the electrostatic potential on a halogen X atom covalently attached to a carbon in two different molecules, CF_3-I and :NC-Br. The red distal end is the σ-hole, relatively depleted in electron density and with a higher, more positive electrostatic potential. (This is opposite the usual coloration that biochemists use in which oxygen (δ^- or fully -) is colored red and nitrogen (in a protonated amine with a positive charge) is shown in blue.)

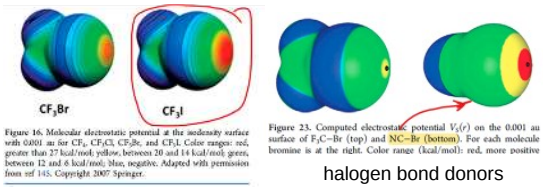


Figure 2.6.27: Electrostatic potential on a halogen X atom on CF_3-I and :NC-Br

Figure 2.6.28 shows a molecule with a carbonyl (a hydrogen bond acceptor with a δ^- :O) interacting with another molecule through either a hydrogen bond or a halogen bond. Again, the red distal end of the halogen X is the σ-hole relatively depleted in electron density.

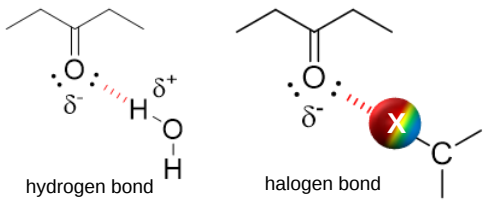


Figure 2.6.28: Comparison of a hydrogen and halogen bond (noncovalent interaction)

Medicinal chemists use halogen substituents on drug molecules to alter drug binding specificity, membrane diffusion, and $t_{1/2}$. Increasingly, they use halogen bonds in rational drug design to increase drug affinity to target proteins.

Figure 2.6.29 shows an interactive iCn3D model illustrating the interaction of a haloaminopyrimidine inhibitor bound to its binding site on the c-Jun N-Terminal Kinase (JNK) protein (2P33).

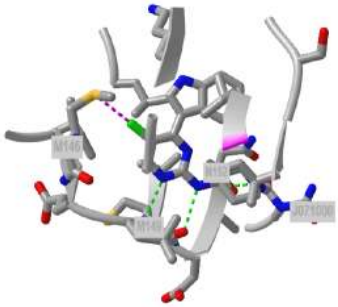


Figure 2.6.29: Haloaminopyrimidine inhibitor bound to its binding site on the c-Jun N-Terminal Kinase (JNK) protein (2P33). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?H4Wabj8my3VsYLSi6

Note that the sulfur of methionine is forming a halogen bond with the Cl atom. Although the electronegativity of sulfur is 2.58, close to that of carbon (2.55), sulfur is larger and more polarizable, so it also develops a slightly positive σ-hole distal to the carbon atom. Analysis of PDB files shows that S--O interactions are common in proteins and most likely impact protein stability.

Ultimately, all ensembles of molecules/ions reach a low, and often the lowest, energy state under a given set of conditions. Noncovalent attractions are maximized, and repulsions are minimized to achieve this state. Consider, for example, solid sodium chloride held together by ionic bonds. The ions are closest packed (face-centered cubic). They cannot get closer together (packing density of about 74%) as simple packing considerations, repulsive electrostatic forces, and collective van der Waals interactions would prevent it. Each Na^+ is surrounded by 6 Cl^- ions and vice versa.

When large molecules like proteins assume a low-energy state, they maximize attractive noncovalent interactions described in this section while minimizing repulsive interactions within a molecule (under given solvent conditions). Packing density reaches similar values to the closest packed spheres (NaCl, for example). Figure 2.6.30 shows a slice through a protein and the crystal lattice of NaCl. The gray circles on the protein show the faces of the sliced atoms. They are superimposed on the surface of the protein shown in colored spheres. If you took a series of cross-sectional slices throughout the protein, you would get a better picture of

packing density than a single slice alone. Collective van der Waals interactions are found among all atoms and ions in a protein, which accounts for the closest packing of most atoms, polar and nonpolar, within the packed protein structure.

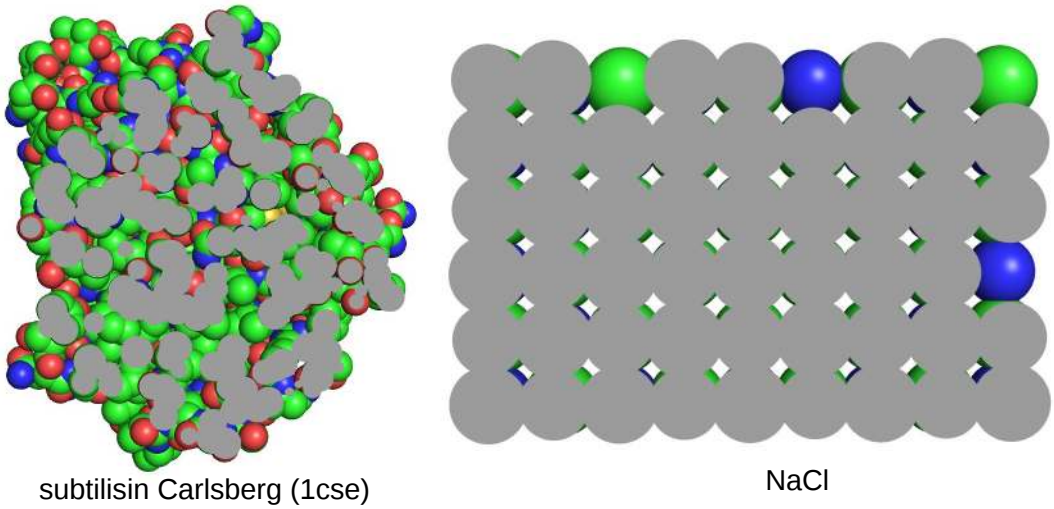


Figure **2.6.30**: A slice through a protein and through the crystal lattice of NaCl

Here is a link to a JSmol tutorial by David Marcy et al, An Introduction to Chemical Bonds and Protein Structure

2.6.11: Summary of Noncovalent Interactions in Biomolecules

It is difficult to understand noncovalent interactions among small molecules, let alone within solvated and densely packed proteins. To help quantify strong noncovalent interactions involving amino acid *side chains*, Xie et al. studied amino acids in the gaseous phase using quantum mechanics. Here are some general conclusions (Xie et al. PLoS ONE 10(9): e0137113. https://doi.org/10.1371/journal.pone.0137113. Creative Commons Attribution License)

- Ion-Ion (salt bridge) interactions between acidic amino acid side chains (Glu^- and Asp^-) and alkaline amino acid side chains (Arg^+, Lys^+, and His^+) are the strongest residue-residue interactions. However, this type of interaction may be weakened by solvation effects and broken by lower pH conditions.
- Cation-π interactions between protonated amino acid side chains (Arg^+, Lys^+, and His^+) and aromatic amino acid side chains (Phe, Tyr, Trp, and His) are 2.5 to 5-fold stronger than common hydrogen bond interactions. They are less affected by the solvation environment.
- Amide bridge interactions, shown in Figure **2.6.31** below, which contain two hydrogen bonds between the two amide-containing amino acid side chains (in the amino acids Asn and Gln), are three times stronger than hydrogen bond interactions, which are less influenced by the pH of the solution.

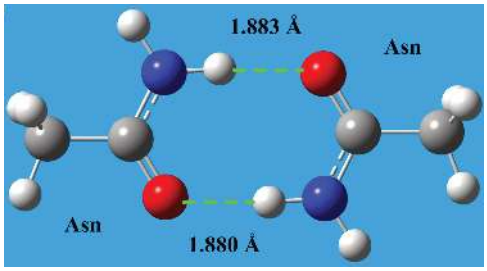


Figure **2.6.31**: An "amide" bridge (Xie et al., ibid)

- Ten of the twenty natural amino acids are involved in salt-bridge, cation, or amide-bridge interactions, often playing important roles in protein-protein, protein-peptide, protein-ligand, and protein-DNA interactions.

Secondary Interactions in Protein Structure and Function

Recent Updates (01/7/24). Another computational study categorized noncovalent interactions between proteins and small molecules (drugs, inhibitors) that bind to them. These small molecules are generically called **ligands**, a term used in the study of transition-metal complexes. They studied 11,016 unique structures in the Protein Data Bank of small-molecule ligands bound to proteins. A histogram displaying the number of each type of interaction between small ligands and proteins is shown in Figure **2.6.32** below.

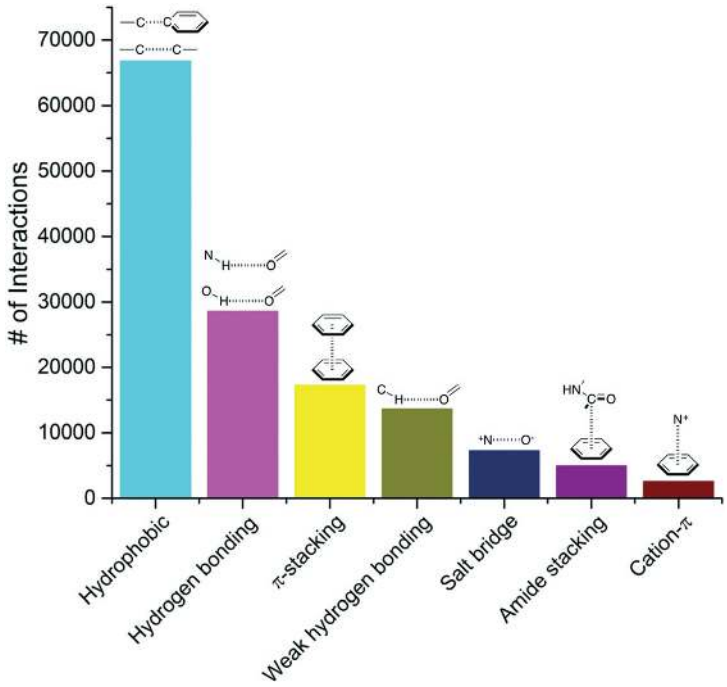


Figure **2.6.32**: Frequency distribution of the most common non-covalent interactions observed in protein–ligands extracted from the PDB. de Freitas and Schapira. *Med. Chem. Commun.*, 2017, **8**, 1970-1981. **de Freitas and Schapira.** DOI: 10.1039/C7MD00381A (Research Article)

Two new interactions are shown: the **weak,** better called **nontraditional, hydrogen bond** and amide stacking. Amide stacking is readily understandable as an interaction between the amide's slightly positive carbonyl carbon and the aromatic ring's electron-dense pi cloud. They are part of a group of interactions known as "**secondary interactions**" between a ligand and protein and within proteins and carbohydrates, which we will explore in a future chapter. They include main-chain interactions (nontraditional H-bonds, chalcogen bonds, and n•π* and n→π*interactions) and interactions with aromatic side chains. These secondary "interactions" are shown in Figure **2.6.33** below.

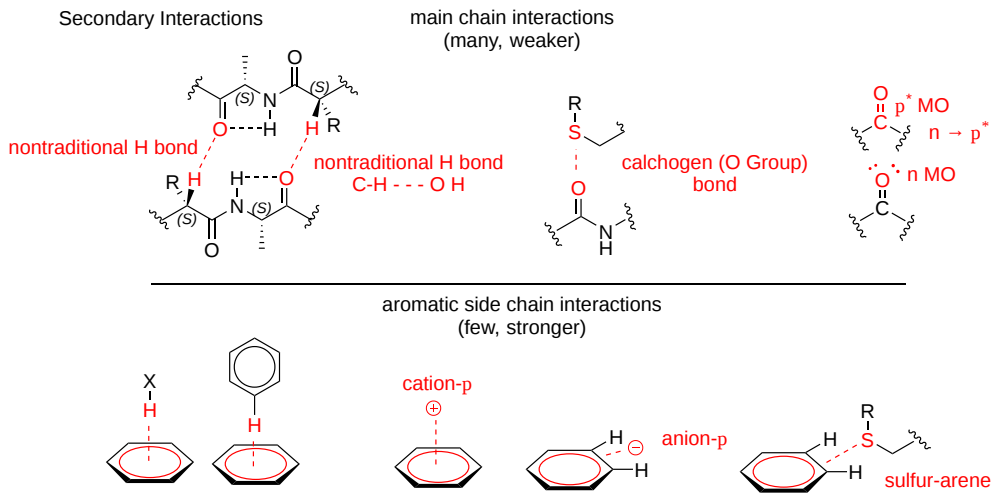


Figure 2.6.33: Secondary Interactions in protein (adapted from Newberry and Raines, *ACS Chem. Biol.* 2019, 14, 8, 1677–1686.

You probably find the nontraditional hydrogen bond more troubling. The hydrogen bond donor is a carbon atom attached to hydrogen, and the hydrogen bond acceptor is the carbonyl oxygen. We have stated that a C-H bond does **not** engage in a hydrogen bond. We assumed that a C-H bond is sufficiently nonpolar, so the carbon atom does not have a slight negative charge, which leaves hydrogen without a slight positive charge. Yet it appears that *there are many C–H⋯O weak hydrogen bonds between ligands and proteins and within proteins*.

C has an electronegativity of 2.5 and H 2.2, with a difference of 0.3, much smaller than that between N and H (3.04-2.2 = 0.8). If the electronic environment around the carbon enhances its slight negative charge, then you could imagine that a C-H could be a hydrogen bond donor. The median distance of the C–H⋯O nontraditional hydrogen bond was 3.4 Å, which is 0.4 Å longer than traditional hydrogen bonds (N–H⋯O, N–H⋯N, O–H⋯O), with an angle of around 130° (compared to an optimal of 109.5^0). The C_α–H⋯O=C interactions are about one-half the strength of an NH⋯O=C hydrogen bond. Hence, they are weak. In the rest of this book, we rarely see a nontraditional C-H bond as a candidate for a hydrogen bond. Nevertheless, it is important to mention it, given its prevalence.

The other main chain interaction, n → π^*, might look like a dipole-dipole interaction. It has its roots in the molecular (not atomic) orbitals of carbonyl groups in the backbone. Simple molecular orbital (MO) theory predicts the general energy levels for the molecular orbitals. The bonding orbitals (σ for the sigma bond and π for the pi bond) are occupied by electron pairs, which stabilize the molecule. The nonbonding electrons in the carbonyl O are in nonbonding orbitals n (neither stabilizing nor destabilizing). The antibonding σ* and π* orbitals on the adjacent carbonyl are unfilled and have higher energy. (Remember that if you bring two atomic orbitals (like 1s and 1s from 2 hydrogen atoms) to form an energy-stabilized covalent bond, you get 2 MOs (σ and σ*). The relative energy of these MOs for carbonyl is shown below in Figure 2.6.34:

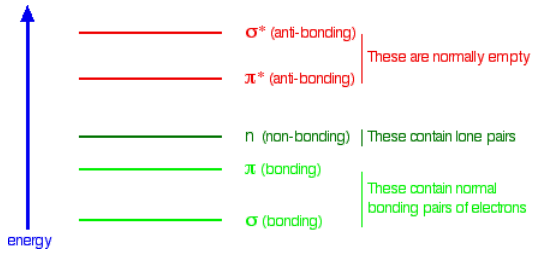


Figure 2.6.34: Relative energy levels of molecular orbitals (not to scale). https://chem.libretexts.org/Bookshel...c_Spectroscopy

Figure 33 shows some aromatic interactions that occur on the side of the ring, not on its top. These are understandable if you consider the 3D electrostatic potential map of benzene, as shown in Figure 2.6.35 below.

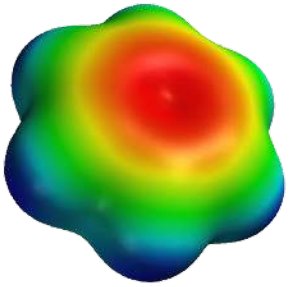


Figure 2.6.35: 3D electrostatic potential of benzene. https://en.wikipedia.org/wiki/File:B...-potential.png

Note that the sides are bluish, which represents a more positive electrostatic potential, which would lead to the attraction of an anion or S atoms with high electron density.

2.6.13: Summary

This chapter explores noncovalent interactions—often broadly labeled as intermolecular forces—that are central to understanding molecular behavior in biochemical systems. It begins by revisiting the concept of water as a solvent, emphasizing the “like dissolves like” principle and how the balance between solute-solvent, and solute–solute or solvent–solvent interactions, together with entropic contributions, determines solubility.

A detailed classification of noncovalent interactions follows, outlining key types such as ion–ion interactions (salt bridges), hydrogen bonds, dipole–dipole interactions, ion–dipole interactions, and induced dipole–induced dipole (hydrophobic) interactions. The chapter reinforces that all these interactions ultimately arise from the fundamental electrostatic forces described by Coulomb’s law, with each type exhibiting distinct distance and direction dependencies.

The text then delves into the energetic landscape of these interactions by comparing potential energy versus distance profiles. Graphical models, including the Morse and Lennard-Jones potentials, illustrate how covalent bonds and noncovalent attractions achieve optimal stabilization at specific interatomic distances. This discussion is further enriched by analogies—such as a ball on a hill—to link changes in potential energy with the forces acting between particles.

Subsequent sections focus on specific examples within biomolecules. For instance, ion–ion interactions (or salt bridges) are examined through the lens of protein structure, where charged side chains contribute to stability. The complexity of hydrogen bonding is unraveled by distinguishing between conventional, strong (low-barrier), and very strong hydrogen bonds, with attention to geometric criteria and their roles in molecular recognition and protein folding.

The chapter also highlights other interactions, such as dipole-dipole and ion-dipole forces, which are crucial for solvation and ligand binding, as well as aromatic interactions like pi stacking and cation–pi interactions, which underpin the stability of DNA and protein structures. A discussion on halogen bonds introduces the concept of the σ-hole, showcasing how anisotropic electron density distributions influence binding in drug design and molecular recognition.

Finally, the chapter touches on emerging secondary interactions—including nontraditional hydrogen bonds, amide stacking, and $n \rightarrow \pi^*$ interactions—that add further nuance to our understanding of protein structure and function. Together, these concepts illustrate how a network of noncovalent forces orchestrates the precise assembly, stability, and dynamic behavior of biomacromolecules in aqueous environments.

In summary, this chapter integrates theoretical foundations, quantitative models, and real-world examples to provide a comprehensive framework for understanding how noncovalent interactions govern molecular behavior in biochemistry.

2.6.14: References

IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). Online version (2019-) created by S. J. Chalk. ISBN 0-9678550-9-8. https://doi.org/10.1351/goldbook.

https://chem.libretexts.org/Bookshel...g_interactions

https://boisestate.pressbooks.pub/ch...ecular-forces/

A low-barrier hydrogen bond in the catalytic triad of serine proteases, PA Frey et al, *Science* 264, 1927-1930 (1994)
DOI: 10.1126/science.7661899

This page titled 2.6: Solubility in an aqueous world - noncovalent interactions in depth is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **2.4: Solubility in an aqueous world - noncovalent interactions in depth** by Henry Jakubowski and Patricia Flatt has no license indicated.

2.7: Solubility in an aqueous world - The Hydrophobic Effect

Learning Goals (ChaptGPT o1, 1/25/25)

Below is a series of targeted learning goals designed to help junior and senior biochemistry majors develop a deep understanding of the thermodynamics of solute–solvent interactions and the hydrophobic effect:

- **Relate "Like Dissolves Like" to Biomolecular Solubility:**
  - Explain why nonpolar molecules (e.g., triacylglycerols, cholesterol esters) are poorly soluble in water and how liquid-liquid extraction exploits these differences.
- **Analyze Thermodynamic Contributions to Solubility:**
  - Differentiate between enthalpic and entropic factors that contribute to the free energy change (ΔG) for the dissolution or partitioning of biomolecules.
  - Interpret experimental data that show how increasing nonpolar chain length affects ΔG, and discuss the significance of per-methylene free energy changes.
- **Apply Free Energy and Chemical Potential Concepts:**
  - Use the relationships $\Delta G = \Delta H - T\Delta S$ and $\Delta G° = -RT \ln K_{eq}$ to explain the spontaneity of solvation and partitioning reactions.
  - Define and relate chemical potential (μ) to free energy changes in a system at equilibrium.
- **Understand Partition Coefficients in Biphasic Systems:**
  - Describe how the partitioning of a solute between two immiscible phases (e.g., water and octanol) is quantified by the equilibrium partition coefficient (K_{part}) and how $\Delta G°$ can be calculated from experimental measurements.
- **Interpret the Hydrophobic Effect:**
  - Discuss the entropic penalty associated with forming an ordered water "cage" around a nonpolar molecule, and explain why this drives the aggregation of nonpolar molecules.
  - Analyze how releasing these structured water molecules during phase separation contributes to a favorable free-energy change.
- **Integrate Thermodynamic Data with Molecular Behavior:**
  - Evaluate experimental thermodynamic parameters ($\Delta\mu°$, $\Delta H°$, $\Delta S°$) for the transfer of small amphiphiles (e.g., aliphatic alcohols) from their pure liquid to water and rationalize why enthalpic contributions may be favorable while entropy disfavors solubilization.
- **Connect Thermodynamics to Biological Function:**
  - Relate the principles governing the solubility and partitioning of biomolecules to their roles in biological membranes, drug diffusion across membranes, and overall cellular homeostasis.

These learning goals encourage an integrated understanding of both the quantitative thermodynamics and the qualitative molecular interactions that underlie the hydrophobic effect and solubility in aqueous environments.

2.7.1: Introduction

Many biomolecules, such as triacylglycerols, cholesterol esters, and waxes, are nonpolar. Other biomolecules, such as proteins and many lipids, have both polar and nonpolar parts. We know from experience that oil floats on water, indicating that it is less dense than water and doesn't dissolve in it. You have also probably performed liquid-liquid extractions in chemistry labs, in which you used the solubility properties of nonpolar molecules to extract them from a water-based mixture and transfer them to a more nonpolar phase, such as octanol or chloroform. To understand the stability of biomolecules that contain nonpolar parts in aqueous solutions, we need to consider not only the noncovalent interactions of these molecules with water (which we explored in Chapter 2.4) but also the thermodynamics of their interactions in aqueous environments.

We have been taught and internalized the notion that "like dissolves like." We anthropomorphize molecules, saying that nonpolar molecules "like" to be in nonpolar environments. We can rationalize solubility properties by examining a molecule's noncovalent attractive and repulsive interactions in an aqueous solution. Still, when we do so, we usually focus on enthalpic contributions to stability. What about entropy? We should consider net changes in noncovalent solute-solute, solute-solvent, and solvent-solvent interactions, along with their thermodynamic contributions to overall stability. When considering the thermodynamics of solubility of molecules in water, we need to determine the ΔG (free energy change) for all processes involved.

2.7.2: The Change in Free Energy (G) and Chemical Potential (μ)

ΔG, the free energy change for a reaction, determines the spontaneity and extent of a chemical or physical reaction. The free energy of a system depends on three variables: temperature T, pressure P, and n, the number of moles of each substance. For the latter, think of solute X on two different sides of a permeable membrane. If the concentration of X is the same on each side, as shown in Figure 2.7.1, the system is in equilibrium.

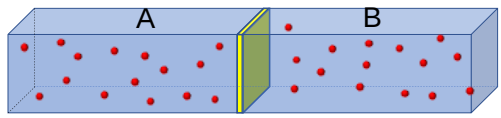


Figure 2.7.1: A system of a molecule in two compartments separated by a membrane

If the system is composed of two different parts, A and B, the system is at equilibrium (ΔG=0) if $T_A = T_B$, $P_A = P_B$, and the change in the absolute free energy per mole of A is $\Delta G_A/\Delta n = \Delta G_B/\Delta n$. More precisely, using simple calculus, we would discuss incremental changes in absolute free energy/mol, dG_A/dn for A, which is the chemical potential of A, μ_A) and dG_B/dn (μ_B) for B. At equilibrium, $dG_A/dn = dG_B/dn$. We will use the free energy G here but μ later in this section. G is the absolute free energy/mol (again chemical potential), where $G=G^o + RTln[A]$. The following equation can be written from the equations you used in introductory chemistry (which we reviewed in Chapter 1.3).

$$\begin{aligned} \Delta G &= \Delta G^0 + RTInQr \\ \Delta G &= \Delta H - T\Delta S \\ \Delta G^0 &= \Delta H^0 - T\Delta S^0 \\ \Delta G^0 &= -RTInKeq \end{aligned} \tag{2.7.1}$$

Now, let's apply this to the chemical equation for the solubility of a given solute in water. You eventually reach a saturation point if you add a sparingly soluble hydrocarbon (HC) or sodium chloride to water. The salt solution is saturated with dissolved NaCl, and no further increase in NaCl (aq) occurs. The solution reaches saturation for a sparingly soluble hydrocarbon, after which phase separation occurs.

Let's add a drop of a slightly soluble hydrocarbon liquid (HC_L) to water, as shown in the diagram below. At t=0, the system is not at equilibrium, and some of the HC will transfer from the pure liquid to water, so at time t=0, $\Delta G_{TOT} < 0$. This is illustrated in Figure 2.7.2.

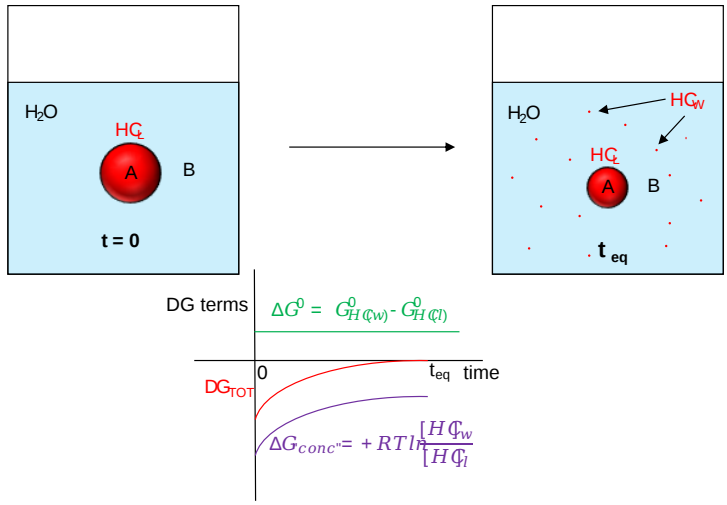


Figure 2.7.2: ΔG vs. time for interaction of a hydrocarbon with water

The following equations can be derived.

$$\Delta G_{TOT} = (G_{HC-W}) - (G_{HC-L}) = G^0_{HC-W} + RT\ln[HC]_W - (G^0_{HC-L} + RT\ln[HC]_L) =$$
$$\Delta G_{TOT} = (G^0_{HC-W} - G^0_{HC-L}) + RT\ln([HC]_W - \ln[HC]_L) = \tag{2.7.2}$$
$$\Delta G_{TOT} = \Delta G^0 + RT\ln\frac{[HC]_W}{[HC]_L}$$

Now, add a bit more complexity to the last example. Add a hydrocarbon x to a biphasic system of water and octanol as shown in Figure 2.7.3. Shake it vigorously. At equilibrium, x would have "partitioned" between the two mostly immiscible phases.

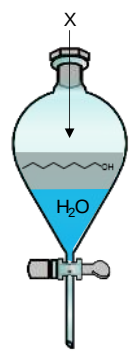


Figure 2.7.3: Use of a separatory funnel for separating immiscible liquid phases

A simple reaction can be written for this system: X aq ↔ X oct.

If X is a hydrocarbon, $\Delta G < 0$ for the reaction written above. Also, $\Delta G^o < 0$, since this term is independent of concentration and depends only on the intrinsic stability of X in water compared to that of octanol. This simple equation holds:

$$\Delta G_{TOT} = (G^0_{X-oct} - G^0_{X-w}) + RT\ln\frac{[X]_{oct}}{[X]_w} = \Delta G^0 + RT\ln\frac{[X]_{oct}}{[X]_w} \tag{2.7.3}$$

At equilibrium, ΔG=0 and the equation can be rewritten as:

$$\Delta G^0 = -RT\ln\frac{[X]_{oct}}{[X]_w} = -RT\ln K_{part} \tag{2.7.4}$$

K_{part} is the equilibrium partition coefficient for X in octanol and water. This value can readily be determined in the lab. Just shake a separatory flask with a biphasic system of octanol and water after injecting a bit of X. Then separate the layers and determine the concentration of X in each phase. Plug these numbers into the last equation. You should be able to predict the sign and relative magnitude of ΔG^o since it depends only on the intrinsic stability of the molecules in the different environments, not on

concentration. K_{part} values are often determined for drugs because they must diffuse across cell membranes to enter the cytoplasm, where they can act. Drugs, hence, must have a reasonable K_{part} to pass through the membrane but not so high that they are insoluble.

2.7.3: Introduction to the Hydrophobic Effect

Now let's ask this question: What are the enthalpic and entropic contributions to the **ΔG** for interacting a nonpolar hydrocarbon (HC) with water? For this section, we will replace ΔG with Δμ (the change in chemical potential, but we will use these terms interchangeably. Likewise, we will use this equation: $\Delta\mu^o = \Delta H^o - T\,\Delta S^o$.

Also, instead of framing the reaction as the dissolution of an organic molecule in water, we will frame it as the **transfer of a hydrocarbon X from an aqueous solution to the pure hydrocarbon liquid (HC) or**

$$X(aq) \leftrightarrow X(HC) \tag{2.7.5}$$

Figure 2.7.4 shows the standard free energies of transfer of a hydrocarbon X from an aqueous solution to a pure liquid hydrocarbon (HC), X (aq) ↔ X (HC). where

$$\Delta\mu^\circ = \mu^\circ x(HC) - \mu^\circ x(aq) \tag{2.7.6}$$

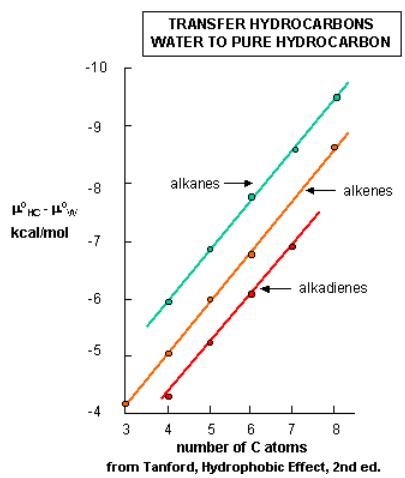


Figure 2.7.4: Standard free energies of transfer (μ^0) of a hydrocarbon X from aqueous solution to a pure liquid hydrocarbon (HC)

$\Delta\mu^o$ is negative because a transfer back to the pure HC is favored based on stability. In each graph, $\Delta\mu^o$ is less than 0, and the value of $\Delta\mu^o$ decreases (gets more negative as you go up the y-axis, which shows increasingly negative values of $\Delta\mu^o$) in a linear fashion with increasing numbers of carbon atoms in the alkyl chain. Notice the lines are unbelievably straight and parallel. Nature is speaking to us in these figures. By determining the surface area of the hydrocarbon molecules and the decrease in $\Delta\mu^o$ with each added CH_2 (methylene group), one can calculate that the $\Delta\mu^o$ decreases by 25 cal/Å^2 (105 J/Å^2), per methylene added.

We expected that $\Delta\mu^o$ for transferring X from a pure liquid HC would be negative. We could get more information if we could determine both the entropic and enthalpic contributions. Such data is presented in the table below, which shows the transfer of short, single-chain alcohol X (an amphiphile with a polar head and a longer nonpolar "tail") from the pure liquid alcohol (ROH) to water (the opposite of the previous figures.)

$$X(ROH) \leftrightarrow X(W) \tag{2.7.7}$$

Thermodynamic Parameters for Transfer of Aliphatic Alcohol X from the **Pure Liquid to Water** at 25 °C (enthalpy determined by calorimetry)

alcohol X	μ_w^0-μ_{ROH}^0 kcal/mol (kJ/mol)	H_w^0-H_{ROH}^0 kcal/mol (kJ/mol)	S_w^0-S_{ROH}^0 cal/deg mol (J/deg mol)	$(Cp)_w^0$-$(Cp)_{ROH}^0$ **cal**/deg mol (J/deg mol)
ethanol	0.760 (3.18)	-2.43 (-10.2)	-10.7 (-44.8)	39 (163)
n-propanol	1.58 (6.61)	-2.42 (-10.2)	-13.4 (-56.1)	56 (234)
n-butanol	2.4 (10)	-2.25 (-9.41)	-15.6 (-65.3)	72 (301)
n-pentanol (solubility 22g/L H_2O)	3.22 (13.5)	-1.87 (-7.82)	-17.1 (-71.5)	84 (352)

We expect $\Delta\mu^o$ to become increasingly positive as the chain length increases and water solubility becomes increasingly disfavored. What is perplexing about this data is not that the transfer of these ROHs to water is disfavored, but that the transfer is **enthalpically** favored (negative ΔH^0). This seems counterintuitive since it goes against the adage that "like dissolves like," as discussed earlier. From an enthalpic point of view, the amphiphiles prefer (albeit marginally) to be in water. What makes this reaction disfavored is entropy. The data show that the nonpolar molecule would prefer not to be in the water because it is disfavored entropically.

At first glance, you might guess that the entropy should favor the movement of ROHs into the water since they could access a larger volume and have greater freedom of motion. Hence, there are more possible microstates for the ROH in water. However, this is only part of the process. What we haven't considered is the water's entropy. A literal cavity must be created to accommodate a hydrocarbon in water. The creation of this more ordered cavity must be entropically disfavored (again because the process proceeds to a state with fewer microstates and lower positional entropy).

In the reverse process, transferring the hydrocarbon from water to the pure liquid dissipates the cavity, thereby increasing the number of available microstates for the released solvent, bulk water. This entropic contribution favors the movement of a hydrocarbon from water to the pure hydrocarbon lipid. This "**hydrophobic effect**" is the main thermodynamic driving force for moving organic molecules out of water.

Imagine this scenario. When you place a hydrocarbon group into water, water seeks (admittedly an anthropomorphic term) to maintain its hydrogen bonding. Hence, it is forced into a more ordered structure around the HC to maintain its H-bonding, characterized by fewer microstates. We will explore the hydrophobic effect in greater detail in a future chapter.

How can we explain the favorable enthalpic contribution of placing a nonpolar molecule into water? Again, this goes against our adage: "like dissolves like." The negative ΔH suggests interactions among all the participants are more favorable when the nonpolar group is in water. One source of such interactions could be the highly structured water in the "cage" surrounding the nonpolar molecule. If it were more structured than bulk water —hence more "ice-like" —then the formation of these extra H-bonds would contribute to the negative enthalpy change. When the nonpolar molecule is removed from the water, which proceeds with a positive ΔH, the "ice-like" water cage would "melt", which, like ice melting, is not favored enthalpically, as heat must be added. Heat energy must be supplied to break the H-bonds as ice melts into liquid water. This molecular model to understand the thermodynamic data might be simplistic, but for now, let's use it.

2.7.4: Summary

This chapter examines the thermodynamics underlying the solubility of biomolecules in aqueous environments, focusing on understanding how nonpolar and amphiphilic molecules behave in water. It begins by revisiting the principle of "like dissolves like," illustrated by everyday observations such as oil floating on water and the practice of liquid-liquid extraction. These examples show that nonpolar substances (e.g., triacylglycerols, cholesterol esters, waxes) do not readily mix with water, whereas many biomolecules possess both polar and nonpolar regions.

The discussion then shifts to a detailed exploration of the thermodynamic factors that govern solubility. Key concepts such as free energy (ΔG), chemical potential (μ), and their dependence on temperature, pressure, and concentration are introduced. The chapter

emphasizes that favorable enthalpic interactions and entropic contributions determine the spontaneity of solvation processes. Using the example of transferring a hydrocarbon from water to a pure hydrocarbon liquid, the text derives equations that relate the free energy change (ΔG) to the equilibrium partition coefficient (K_part) through the relation $\Delta G° = –RT \ln K_{part}$.

A significant portion of the chapter is dedicated to the hydrophobic effect. It explains that although the transfer of nonpolar molecules into water might be enthalpically favorable—owing to the formation of structured, “ice-like” water cages around the nonpolar moieties—the overall process is disfavored due to a substantial entropic penalty. This entropy loss arises from ordering water molecules that must occur to accommodate the nonpolar solute. Conversely, when nonpolar molecules aggregate or partition into a nonpolar phase, water is released from this ordered state, resulting in an overall favorable free energy change that drives many biological processes, such as membrane formation and protein folding.

By integrating theoretical thermodynamics with experimental data, this chapter provides a comprehensive framework for understanding how enthalpic and entropic forces interplay to determine the solubility and partitioning behavior of biomolecules, a foundational concept in biochemistry.

This page titled 2.7: Solubility in an aqueous world - The Hydrophobic Effect is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

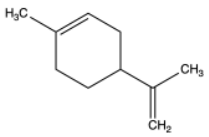
- **2.5: Solubility in an aqueous world - The Hydrophobic Effect** by Henry Jakubowski and Patricia Flatt has no license indicated.

2.8: Chapter 2 Questions

Section 1 Questions:

Q1) Based on the interactive figure 2.1.1 in the default view, do you hypothesize that water could enter the core of the micelle? Now, let's test your hypothesis! Open the interactive figure, and click Style → Surface Type → Solvent Accessibility. Based on the space-filled solvent accessibility map, do you see any openings for water molecules to enter through? Explain your answer.
A1) Hypothesis can range from no water is accessible, or water could freely occupy the "empty" space in the default view of the micelle. However, once the solvent accessibility filter is applied, it is clear that the micelle will exclude water from entering the hydrophobic core. Some areas of the micelle (the polar head groups) are able to for interactions with the water and therefore show up as red/yellow in this view, but it is key to note there are no pores or solvent-accessible spaces on the surface of the micelle.

Q2) The structure for Limonene, the compound that gives citrus fruits their classic smell is shown below. Examine the structure and answer the following questions.



a) Will Limonene form any associations with water molecules? Explain.
b) For our brain to register the citrus aroma, Limonene needs to bind to a surface protein receptor, what type of interactions could be taking place? Explain.
c) When Limonene binds to its receptor are the interactions from b) stabilized by enthalpy ΔH or entropy ΔS?

A2)
a) No, there are no dipoles or polar functional groups on limonene.
b) The methyl and ethyl functional groups would likely form hydrophobic bonds with nonpolar amino acids in the protein. Because there are no polar/dipole residues, Limonene cannot make hydrogen or ionic bonds.
c) Binding of Limonene to its surface receptor will likely be stabilized by entropy. When the nonpolar Limonene binds to a nonpolar region of the surface receptor, order water will be released, thus creating a +ΔS.

Q3) Consider the following reaction of a ***polar*** substance and water at room temperature (22°C):
$\text{Glucose}_{(s)} + H_2O_{(l)} \rightarrow \text{Glucose}_{(aq)}$
a) Estimate the enthalpy ΔH and entropy ΔS (+,-,≈0) for the reactants and product. Consider the order/disorder of the over all reaction as well as the net charge of the bond enthalpy.
b) Based on the Gibbs free energy equation ΔG = ΔH - TΔS, will this reaction be spontaneous, non-spontaneous, or at equilibrium?

A3)
a) $\text{Glucose}_{(s)} + H_2O_{(aq)} \rightarrow \text{Glucose}_{(aq)}$
ΔH: H-bonds H-bonds H-bonds ΔH ≈ 0
ΔS: Solid Liquid Solution ΔS = +

b) ΔG = ΔH - TΔS
= 0 - (+)
= -
This dissolution of polar substances into water is spontaneous.

Q4) Consider the following reaction of a ***nonpolar*** substance and water at room temperature (22°C):
$\text{Methane } (CH_4)_{(g)} + H_2O_{(l)} \rightarrow \text{Methane } (CH_4)_{(aq)}$
a) Estimate the enthalpy ΔH and entropy ΔS (+,-,≈0) for the reactants and product. Consider the order/disorder of the overall reaction as well as the net charge of the bond enthalpy.
b) Based on the Gibbs free energy equation ΔG = ΔH - TΔS, will this reaction be spontaneous, non-spontaneous, or at equilibrium?

A4)
a) Methane $(CH_4)_{(g)}$ + $H_2O_{(l)}$ → Methane $(CH_4)_{(aq)}$
ΔH: None H-bonds H-bonds ΔH ≈ 0 or slightly -
ΔS: Gas Liquid Solution ΔS = -

b) ΔG = ΔH - TΔS
= 0/- - (-)
= +
This dissolution of nonpolar substances into water is non-spontaneous. This is primarily due to water forming an ordered cage around the nonpolar methane gas. In terms of enthalpy, the water cage creates a more favorable environment for methane than the pure gas in liquid water, thus the enthalpy can be considered slightly negative. However, the cage structure the water molecules form around the methane gas makes the system ordered. Therefore the overall ΔG for nonpolar substances solubilized with water is positive and therefore non-spontaneous.

Need to add in some reactions? (last chunk of section 1)

Section 2 Questions

Q1) Tyrosine is commonly found in the active sites of enzymes, as the unique structure of its R-group can act as either an acid or base.

a) The enzyme DcpS, an mRNA capping enzyme, utilizes a tyrosine residue as an active site acid, and measurements show that the tyrosine is 75% ionized. What must the local pH need to be for this to occur?
b) Another way for tyrosine to be used as an acid is to lower its pKa by creating weak interactions with the oxygen of its R-group using neighboring amino acids. What type of interaction(s) could achieve this goal?
c) One enzyme utilizing the strategy in b) is Glutathione-S-Transferase. Measurements show that the active site pH is 7.5, and the tyrosine residue is 95% ionized. What must the pKa of tyrosine be for this to occur?

A1)
a)

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$pH = 10.1 + log\frac{[75]}{[25]}$$

pH = 10.1 + (0.477)
pH = 10.58

b) Hydrogen bonding can create a change in the net dipole of the oxygen in tyrosine by making it more acidic and decreasing the pKa.
c)

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$7.5 = pK_a + log\frac{[95]}{[5]}$$

7.5 = pK_a + 1.28

6.22 = pK_a

Section 3 Questions

Q1) Your lab wants to study an enzyme that catalyzes a reaction inside the chloroplast stroma, which has a pH of 8.0 due to the proton gradient that pumps H^+ from the stroma to the thylakoid lumen.
a) Which of the buffers below would be the best choice to study this enzyme, in vitro? Explain your choice. Are there any other buffers that could work?
-Tricine: pKa 8.05
-TAPS: pKa 8.40
-MES: pKa 6.1
-Citrate: pKa 6.40
-HEPES: pKa 7.48
b) How many moles of the conjugate base form of HEPES would there be in 2.5 L of a 175 mM solution at pH 8.0?

A1)
a) Tricine would be the best choice as buffers work best at a pH closest to their pKa. TAPS or HEPES could work in a pinch if Tricine was not available, but remember, pH and pKa are log scales! So, while the pKa values might not seem that far from the intended pH of 8.0, on the log scale that is quite a difference in H^+ concentration.
b)

$$2.5L \times \frac{175mmoles}{L} \times \frac{1mole}{1000mmoles} = 0.4375mmoles$$

Now we want to determine what fraction of those moles are in the conjugate base form:

$$[A^-] + [HA] = 0.4375$$

$$[A^-] = 0.4375 - [HA]$$

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$8.0 = 7.48 + log\frac{[A^-]}{[HA]}$$

$$0.52 = log\frac{[A^-]}{[HA]}$$

$$3.31 = \frac{0.4375 - [HA]}{[HA]}$$

At pH 8.0, [HA] = 0.101 mmoles, which is 23% of the total moles of HEPES in the solution.

Q2) As discussed in Section 2.3, when CO_2 is inhaled, it reacts with water to form the weak acid carbonic acid, acidifying the blood. The reaction is given below for reference.
$CO_2 + H_2O \rightarrow H^+ = HCO_3^-$

Currently, the air you breathe contains about 0.04% CO_2. This number has risen from 0.03% in the 1960s and is projected to increase to 0.08% by 2100 if fossil fuel consumption remains at its current rate. (P_{CO2} =0.0003 atm, Keeling, 1960, P_{CO2} =0.0008 atm MIT System Dynamics Group, 2015).

Using the Ideal Gas Laws, there were 0.44 μM CO_2 in the 1960s, 0.60 μM CO_2 today, and as much as 1.2 μM CO_2 in 2100. The pKa of carbonic acid is 6.35 and the pH of your blood is 7.60.

a) What is the change in blood pH due to the increase in atmospheric CO_2 from the 1960s to today?
b) If you were to do the calculations for what the pH of the blood would rise to in 2100, you'd find it to be pH 7.3. Do you think the body would be able to compensate for this? Use the pKa of histidine to explain your answer.

A2)
a)

$$pH = pK_a + log\frac{[A^-]}{[HA]}$$

$$7.60 = 6.35 + log\frac{[A^-]}{0.6}$$

$$1.25 = log\frac{[A^-]}{0.6}$$

$$17.8 = \frac{[A^-]}{0.6}$$

$[A^-]$ = 10.7 μM

$$pH_1960 = 6.35 + log\frac{10.7}{0.44}$$

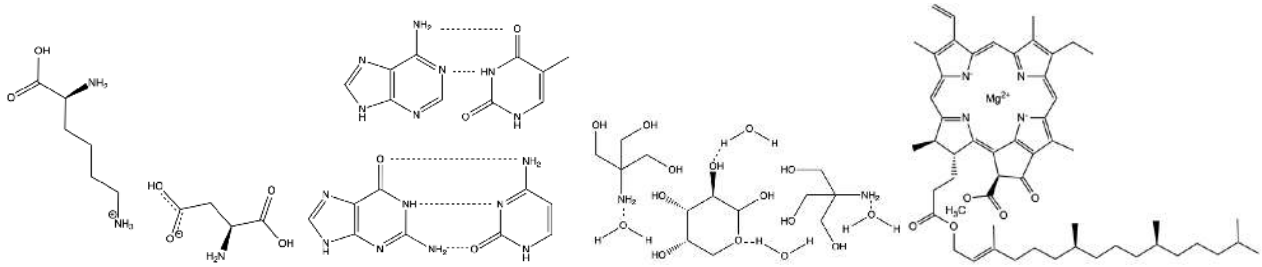
pH_{1960} = 6.35 + 1.39
pH_{1960} = 7.74

Now knowing the pH of the blood from 1960, the change is 7.74 - 7.6 = 0.14 pH units.

b) Yes the body would compensate by increasing the amount of carbonic acid in the blood to buffer the increase in free H^+. If the body did not do this, the protonation state of histidine would change and affect every histidine-containing protein in the blood.

Section 4 Questions

Q1) Categorize the following bonds as ion-ion, ion-dipole, dipole-dipole, H-bond, or hydrophobic, (some maybe be used more than once, or not at all).



A1)

Q2) You discover a new enzyme, "biochemase" and decide to crystalize the protein to determine if you can determine the role from its structure. You hypothesize the protein forms a homodimer when two identical subunits of a protein come together to form one functional structure. You identify two regions that you believe could be the dimer interface. The first region looks to contain several polar amino acids, whose R-groups are less than 5 Å apart, while the second section contains a large cluster of hydrophobic amino acids.

a) You perform some biochemical experiments and discover that when the **polar region** is removed, the enzyme is not able to self-assemble, but maintains two **intact separate structures**. What can you conclude about this region and the type of interactions that are stabilizing the dimer interface?

b) When you remove the **nonpolar** region, you notice the protein aggregates or crashes out of the solution due to an inability to fold correctly. What role can you hypothesize the nonpolar amino acids have in keeping the protein properly folded?

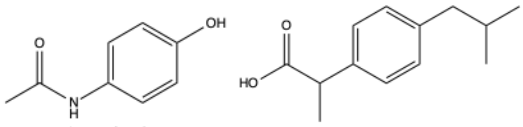
A2)

a) When the polar region is removed, the protein is still able to fold into two stable subunits, but unable to form its homodimer. Therefore we can conclude that the polar region is necessary for the dimer interface, but the protein can still form subunits without it present.

b Now, when the nonpolar region is removed the protein cannot form subunits or dimers. So, we can conclude that the nonpolar region is necessary for protein stability, and without this region present, the protein cannot fold properly. Most likely, the hydrophobic region forces the protein to fold in the correct way by forcing the nonpolar amino acids to the core, and the polar to the surface. This would then allow for a proper dimer interface to form.

<u>Section 5 Questions</u>

Q1) Acetaminophen (Tylenol) and Ibuprofen (Aspirin) are both common pain-relieving/fever-reducing drugs. However, their chemical properties differ, making acetaminophen more suitable for relieving headaches and fever, while ibuprofen can more effectively reduce pain. The pK_{part} for acetaminophen is 0.91, while the pK_{part} is 3.97. (DrugBank) The chemical structures for each are given below:



Acetaminophen Ibuprofen

a) Using equation 2.5.4, determine the ΔG° for each compound, and hypothesize which is more soluble based on your answer and explain. Assume a normal body temperature of T = 37°C

b) With your answer from a) and the information reviews in chapter 1 on functional groups, and types of non-polar interactions in chapter 2, identity the regions of both acetaminophen and ibuprofen that can facilitate interactions.

c) Using your ΔG° calculations and functional group analysis, predict which compound is more likely to stay in the blood and bound to red blood cells, and which can rapidly diffuse across the cell membrane.

Q2) You are tasked with creating a solution of nonpolar and polar solvents to use for thin layer chromatography (a technique we

will cover later), to separate lipids. Consider the following enthalpic and entropic values for these nonpolar and polar solvents. Using the information discussed in section 5, which of the following combinations of solvents will yield a homo

---

This page titled 2.8: Chapter 2 Questions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **2.6: Chapter 2 Questions** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

3: Amino Acids and Peptides

3.1: Amino Acids and Peptides
3.2: Amino Acid Charges
3.3: Cysteine Chemistry
3.4: 3.1-Structure and Property of the Naturally-Occurring Amino Acids
3.5: From Amino Acids to Peptides

3: Amino Acids and Peptides is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.1: Amino Acids and Peptides

Learning Goals (ChatGPT o1, 1/25/25)

- **Understand the Role and Diversity of Proteins:**
  - Describe the various functions of proteins (structural, regulatory, enzymatic, transport, etc.) and explain why proteins are considered the most functionally diverse macromolecules.
  - Recognize that proteins are linear polymers of alpha-amino acids linked by peptide bonds.
- **Comprehend Alpha Amino Acid Structure:**
  - Identify the common structural features of an alpha amino acid (α-amino group, α-carboxylic acid, hydrogen, and a variable R-group).
  - Memorize the 20 naturally occurring amino acids along with their three-letter and one-letter abbreviations.
  - Explain how variations in the R-group confer different chemical properties (nonpolar, polar uncharged, acidic, and basic) to amino acids.
- **Master Peptide Bond Formation and Protein Primary Structure:**
  - Describe the mechanism of peptide bond formation via nucleophilic attack and the release of water.
  - Illustrate the concept of primary protein structure as the linear sequence of amino acids and appreciate the enormous diversity of possible sequences.
- **Analyze Amino Acid Classification and Side Chain Characteristics:**
  - Categorize amino acids based on their side chain properties (e.g., aliphatic vs. aromatic nonpolar, polar uncharged, acidic, and basic).
  - Discuss how side chain properties influence protein folding, stability, and function.
  - Interpret hydrophobicity (or hydropathy) scales (e.g., Kyte-Doolittle, Hopp-Woods) and relate these to protein topology (buried versus surface residues).
- **Grasp Stereochemistry and Chirality in Amino Acids:**
  - Explain the concept of chirality and why all naturally occurring proteinogenic amino acids (except glycine) are L isomers.
  - Differentiate between the D/L nomenclature and the R/S system, and understand why the D/L system is preferred in biochemistry.
- **Apply Acid-Base Chemistry to Proteins:**
  - Utilize the Henderson–Hasselbalch equation to predict the ionization state of amino acid side chains and terminal groups.
  - Interpret titration curves for amino acids and proteins and calculate the isoelectric point (pI) of proteins based on the ionizable groups present.
- **Examine Chemical Reactivity of Amino Acid Side Chains:**
  - Identify which side chains serve as hydrogen bond donors and acceptors, and explain their roles in catalysis and substrate binding.
  - Compare the nucleophilicity and basicity of key amino acid side chains (e.g., lysine, cysteine, histidine) and discuss factors (like electronegativity and steric effects) that influence reactivity.
  - Describe specific chemical reactions involving amino acid side chains, including acylation, Schiff base formation, and nucleophilic substitution reactions used in protein modification.
- **Explore Post-Translational Modifications (PTMs) and Their Biological Significance:**
  - List common PTMs (e.g., phosphorylation, acetylation, glycosylation, oxidation) and explain how these modifications alter protein structure and function.

- Understand how aberrant or deleterious PTMs (e.g., glycation, carbonylation) can lead to altered protein activity and contribute to disease processes.

These learning goals are intended to help students integrate structural, chemical, and functional perspectives on proteins, preparing them for more advanced topics in enzymology, protein engineering, and molecular biology.

3.1.1: Introduction

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective. They may serve in transport, storage, or membranes. They may be toxins or enzymes. Each cell in a living system may contain thousands of proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, polymers of alpha amino acids arranged in a linear sequence and connected by covalent bonds.

3.1.2: Alpha Amino Acid Structure

The major building blocks of proteins are called ***alpha (α) amino acids***. As their name implies, they contain a carboxylic acid and an amine functional group. The alpha designation indicates that a single carbon atom separates these two functional groups. In addition to the amine and carboxylic acid, the alpha carbon is also attached to a hydrogen atom and to one additional group that can vary in size and length. In the diagram below, this group is designated as an R-group. Within living organisms, 20 common amino acids are used as protein building blocks. They differ only in the R-group position. The fully protonated structure of an amino acid (at low pH) is shown in Figure **3.1.1**.

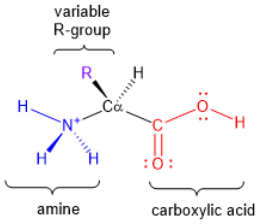


Figure **3.1.1**: Generic Structure of an Amino Acid

The twenty common naturally occurring amino acids contain an alpha-carbon, an amino, a carboxylic acid, and an R group (or side chain). The R group side chains may be either nonpolar, polar uncharged, or charged, depending on the functional group, the pH, and the pKa of any ionizable group in the side chain.

Two other amino acids occasionally appear in proteins. One is selenocysteine, found in Archaea, eubacteria, and animals. Another is pyrrolysine, found in Archaea. Bacteria have been modified to incorporate two new amino acids, O-methyl-tyrosine and p-aminophenylalanine. The yeast strain *Saccharomyces cerevisiae* has been engineered to incorporate five new unnatural amino acids (using the TAG nonsense codon and new, modified tRNA and tRNA synthetases) with keto groups that allow chemical modifications to the protein. We will concentrate only on the 20 abundant, naturally occurring amino acids.

Proteins are polymers of monomeric amino acids with an amide link (also called a peptide bond) between the α-carboxylic group of one amino acid and the α-amine of the next one. Figure **3.1.2** shows the twenty naturally occurring α-amino acids as they would appear internally within a protein sequence. The squiggles show the connecting amide/peptide bond between adjacent amino acids. Students often assume that the α-amino and α-carboxylic acid groups within a protein sequence are free and not part of the peptide bond. This figure should help in resolving that misconception. The three-letter and one-letter abbreviations of each amino acid are shown, as well as the typical pKa values of side chains (R groups). It is important to memorize the three-letter and one-letter codes for the amino acids.

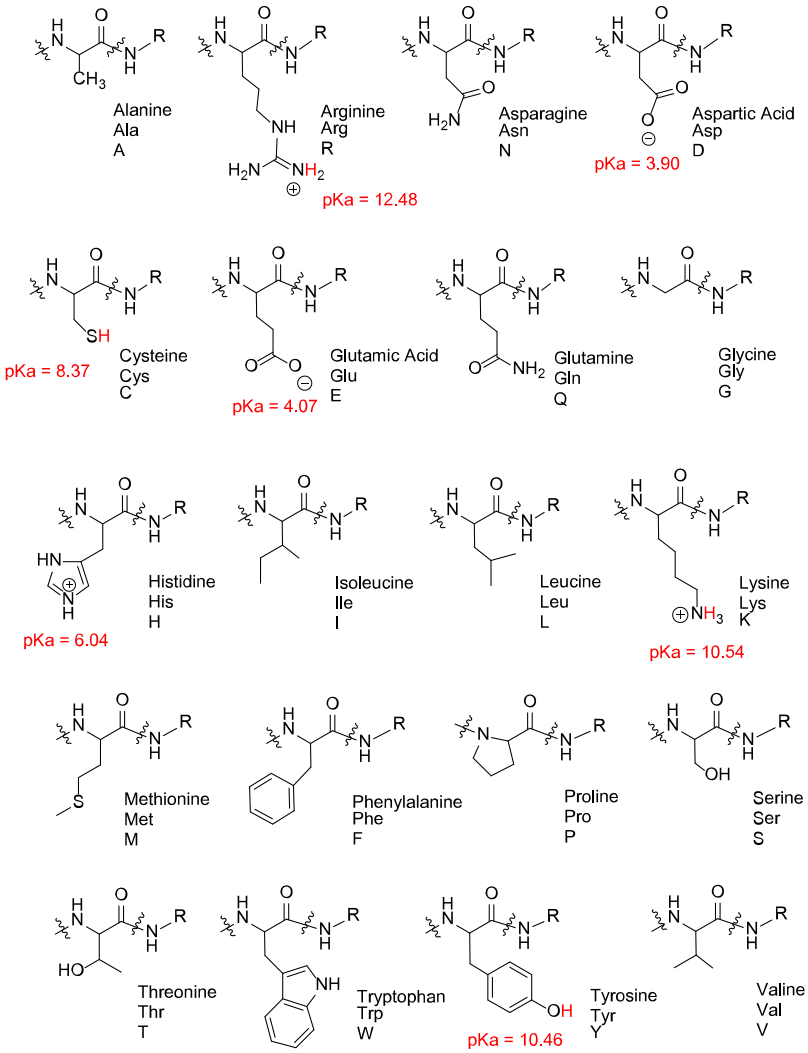


Figure **3.1.2**: Side chains of naturally occurring amino acids embedded in a protein

Amino acids form polymers through a nucleophilic attack by the amino group of an amino acid at the electrophilic carbonyl carbon of the carboxyl group of another amino acid. The carboxyl group of the amino acid must first be activated to provide a better leaving group than OH-. The resulting link between the amino acids is an amide link, which biochemists call a peptide bond. In this reaction, water is released. In a reverse reaction, the peptide bond can be cleaved by water (hydrolysis). This is illustrated in Figure **3.1.3**.

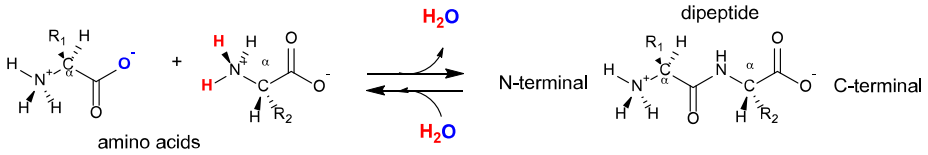


Figure **3.1.3**: Amino Acids React to Form a Dipeptide

Proteins are polymers of twenty naturally occurring amino acids. In contrast, nucleic acids are polymers of just four different monomeric nucleotides. The sequence of a protein and its total length differentiate one protein from another. Just for an octapeptide, there are over 25 billion different possible arrangements of amino acids. Compare this to just 65536 oligonucleotides (4 different monomeric deoxynucleotides) of 8 monomeric units, an 8mer. Hence, the diversity of possible proteins is enormous.

The resulting structure is called a dipeptide when an amide bond links two amino acids. Likewise, we can have tripeptides, tetrapeptides, and other polypeptides. At some point, when the structure is long enough, it is called a protein. The average molecular weight of yeast proteins is about 50,000, with about 450 amino acids. The largest human protein might be titin, with a

molecular weight of about 3 million (about 30,0000 amino acids). A new class of very small proteins (30 or fewer amino acids, and perhaps better named polypeptides) called smORFs (small open reading frames) has recently been discovered to have significant biological activity. These are encoded directly in the genome and are produced by the same processes that produce regular proteins (DNA transcription and RNA translation). They are not the result of the selective cleavage of a larger protein into smaller peptide fragments.

Figure **3.1.4** shows several ways to represent the structure of a polypeptide or protein, each showing differing amounts of information. Note that the atoms in the side chains are denoted alpha, beta, gamma, delta, epsilon ...

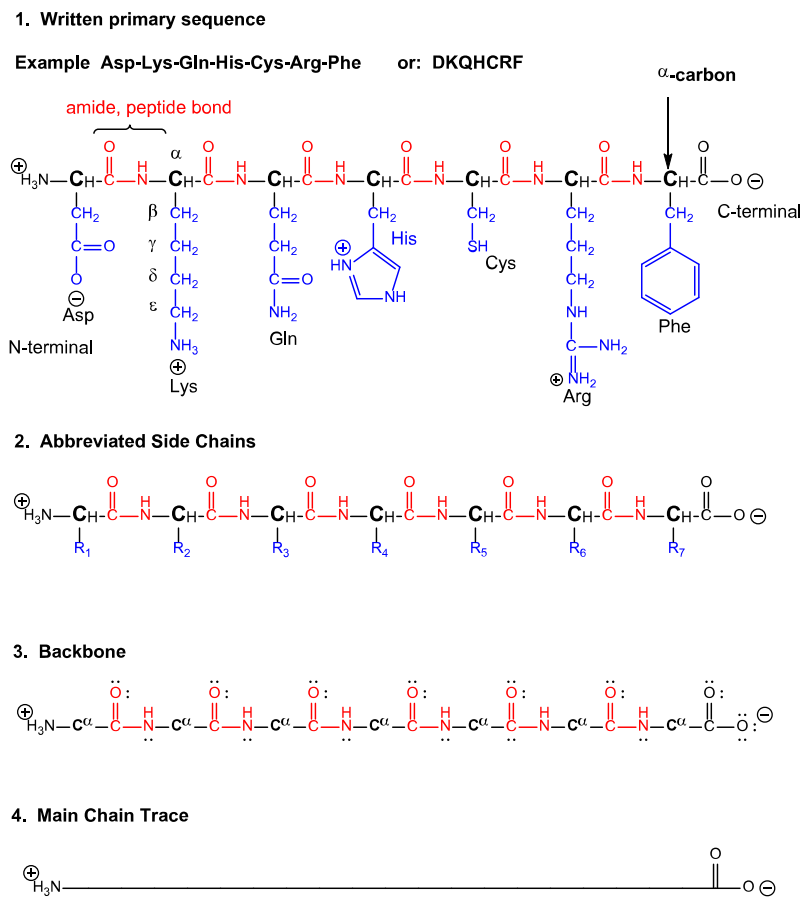


Figure **3.1.4**: Different ways to represent the structure of a peptide/protein sequence.

3.1.3: Characteristics of Amino Acids

The different R-groups exhibit distinct characteristics depending on the nature of the atoms incorporated into the functional groups. Some R-groups predominantly contain carbon and hydrogen and are very nonpolar or hydrophobic. Others contain polar uncharged functional groups such as alcohols, amides, and thiols. A few amino acids are basic (containing amine functional groups) or acidic (containing carboxylic acid functional groups). These amino acid side chains can carry full charges and engage in ionic interactions. Each amino acid can be abbreviated using a three-letter and a one-letter code. Figure **3.1.5** shows groupings of the amino acids based on their side chain properties.

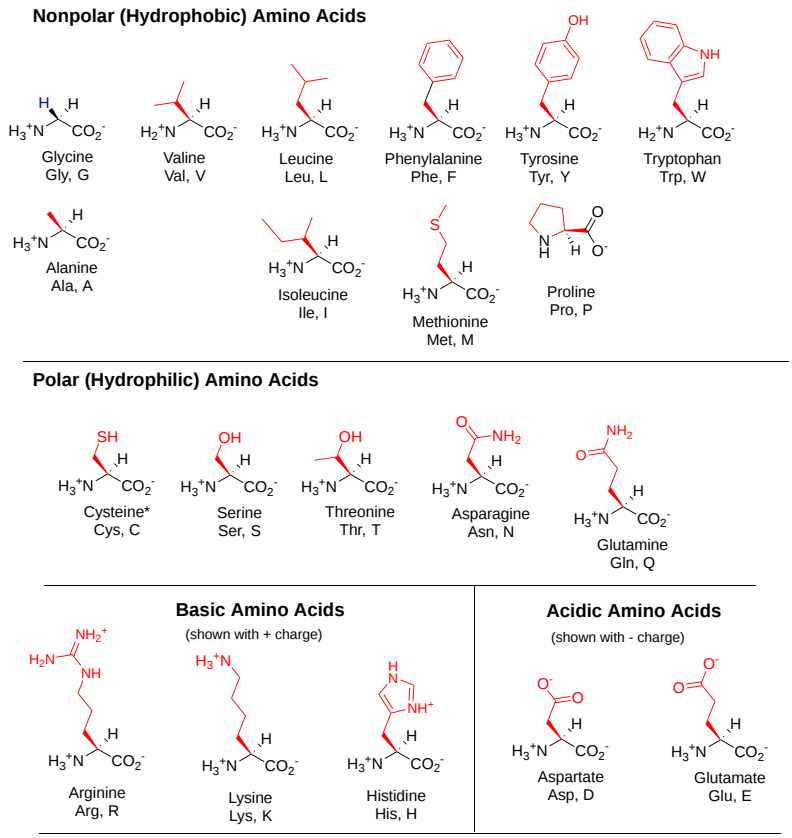


Figure **3.1.5**: Structure of the 20 Alpha Amino Acids used in Protein Synthesis. R-groups are shown in **red**, connected to the alpha-carbon by a **red** wedge bond.

The above classification for a few amino acids is somewhat arbitrary, as described below. For example, cysteine is often buried in protein in a more hydrophobic environment.

3.1.4: Nonpolar (Hydrophobic) Amino Acids

The nonpolar amino acids can be subdivided into two more specific classes: ***aliphatic*** and ***aromatic***. The ***aliphatic amino acids (glycine, alanine, valine, leucine, isoleucine, and proline)*** typically contain straight or branched hydrocarbon chains, with two exceptions. Glycine, the simplest amino acid, has a single hydrogen atoms as the "side" chain. Proline is also classified as an aliphatic amino acid, but the hydrocarbon side chain has cyclized with the terminal amine, creating a unique 5-membered ring, non-aromatic ring structure. As we will see in the next section on primary structure, proline can significantly alter the 3-dimensional structure of the protein due to the rigidity of its ring when incorporated into the polypeptide chain, and it is commonly found in regions of the protein where folds or turns occur.

As their names imply, the ***aromatic amino acids (phenylalanine, tyrosine, and tryptophan)*** contain an aromatic functional group, making them largely nonpolar and hydrophobic due to their high carbon-to-hydrogen ratio. However, it should be noted that hydrophobicity and hydrophilicity represent a sliding scale, and each amino acid can have different physical and chemical properties depending on its structure. For example, the hydroxyl group in tyrosine increases its reactivity and solubility compared to phenylalanine.

Methionine, one of the sulfur-containing amino acids, is usually classified as a nonpolar, hydrophobic amino acid. The terminal methyl group is a thioether, which generally cannot form a permanent dipole within the molecule and retains low solubility.

3.1.5: Polar (Hydrophilic) Amino Acids

The polar, hydrophilic amino acids can be subdivided into three major classes: the polar uncharged, acidic, and basic functional groups. Within the ***polar uncharged class***, the side chains contain heteroatoms (O, S, or N) capable of forming permanent dipoles within the R-group. These include **the hydroxyl- and sulfhydryl-containing amino acids, serine, threonine, and cysteine**, and

the amide-containing amino acids, **glutamine and asparagine.** Two amino acids, **glutamic acid (glutamate),** and **aspartic acid (aspartate)** constitute **the acidic amino acids** and contain side chains with carboxylic acid functional groups capable of fully ionizing in solution. The **basic amino acids, lysine, arginine, and histidine,** contain amine functional groups that can be protonated to carry a full charge.

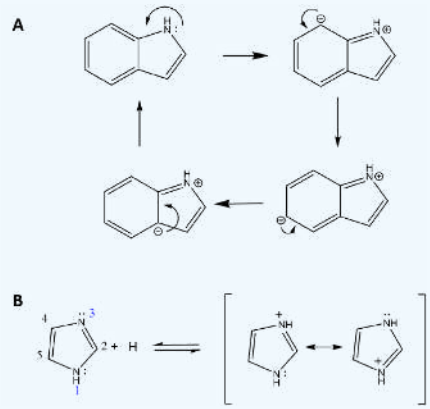
Many amino acids with hydrophilic R groups can participate in the ***active sites*** of enzymes. An ***active site*** is the part of an enzyme that directly binds to a substrate and carries out a reaction. Protein-derived enzymes contain ***catalytic groups*** consisting of amino acid R-groups that promote the formation and degradation of bonds. The amino acids that play a significant role in the binding specificity of the active site are usually not adjacent to each other in the primary structure but form the active site as a result of folding in creating the tertiary structure, as you will see later in the chapter.

Example 3.1.1

Tryptophan contains an amine functional group. Why isn't tryptophan basic?

Answer

Tryptophan contains an indole ring structure with the amine functional group. However, due to the proximity of and electron-withdrawing nature of the aromatic ring structure, the lone pair of electrons on the nitrogen is unavailable to accept a proton. Instead, they form *pi*-bonds within several of the different resonance structures possible for the indole ring. Figure 2.3A shows four of the possible resonance structures for indole. Conversely, within the imidazole ring structure found in histidine, there are two nitrogen atoms, one of which is involved in the formation of resonance structures (Nitrogen #1 in Figure 2.3B) and cannot accept a proton, and the other (Nitrogen #3) that has a lone pair of electrons that is available to accept a proton.



Comparison of the Structural Availability of Lone Pair of Electrons on Nitrogen to Accept a Proton in the Indole and Imidazole Ring Structures. (A) Four resonance structures of the indole ring structure show that the lone pair of electrons on the nitrogen is involved in forming pi-bonds. (B) The imidazole ring structure has one nitrogen (1) that is involved in resonance structures (not shown) and is not available to accept a proton. In contrast, the second nitrogen (3) has a lone pair of electrons available to accept a proton, as shown.

Exercise 3.1.1

Given the example above, describe, using a chemical diagram, why the amide nitrogen atoms found in asparagine and glutamine are not basic.

Answer

The lone pair is delocalized into the peptide bond (via resonance), so it is unavailable for sharing.

Recent Updates: 5/5/24

3.1.6: Quantitative measures of amino acid polarity and hydrophobicity

There are quantitative methods to determine the relative polarity of an amino acid side chain, so it's not just a matter of visual inspection or guesswork. Let's consider hydrophobicity or hydropathy scales. Most are based on the standard free energy

of transfer of a side chain from water to a nonpolar solvent. Each amino acid side chain is assigned a number ranging from negative to positive. The Kyte-Doolittle and the Hopp-Woods hydropathy scales are the two commonly used scales, as shown in Table **3.1.1** below.

Table **3.1.1**: Kyte-Doolittle and Hopp-Woods hydrophobicity values

Amino Acid	Kyte-Doolittle	Hopp-Woods
Alanine	1.8	-0.5
Arginine	-4.5	3.0
Asparagine	-3.5	0.2
Aspartic acid	-3.5	3.0
Cysteine	2.5	-1.0
Glutamine	-3.5	0.2
Glutamic acid	-3.5	3.0
Glycine	-0.4	0.0
Histidine	-3.2	-0.5
Isoleucine	4.5	-1.8
Leucine	3.8	-1.8
Lysine	-3.9	3.0
Methionine	1.9	-1.3
Phenylalanine	2.8	-2.5
Proline	-1.6	0.0
Serine	-0.8	0.3
Threonine	-0.7	-0.4
Tryptophan	-0.9	-3.4
Tyrosine	-1.3	-2.3
Valine	4.2	-1.5

Note that the Hopp-Woods scale is more like a hydrophilicity scale, since more polar residues have higher positive values. It was developed to identify likely antibody or other protein interaction sites on protein surfaces that display more hydrophilic side chains.

Some discrepancies exist in which amino acid side chains are nonpolar between the Kyte-Doolittle values and Figure **3.1.5**. The Kyte-Doolittle scale shows that glycine, the two large aromatic tyrosine and tryptophan, and proline are more polar than nonpolar, and that cysteine is quite nonpolar.

As we will see in subsequent sections, a continuous stretch of amino acids found to have a high average hydrophobicity (low hydrophilicity) is probably buried in the interior of a protein away from the aqueous environment. Conversely, a continuous stretch

with low hydrophobicity (high hydrophilicity) is likely buried in a protein or a membrane bilayer. Consider the example of the water-soluble bovine alpha-chymotrypsinogen, a 245 amino acid protein, whose sequence is shown below in single-letter code.

1 CGVPAIQPVLSGLSRIVNGEEAVPGSWPWQVSLQDKTGFHFCGGSLINENWVVTAAHCGV
61 TTSDVVVAGEFDQGSSSEKIQKLKIAKVFKNSKYNSLTINNDITLLKLSTAASFSQTVSA
121 VCLPSASDDFAAGTTCVTTGWGLTRYTNANTPDRLQQASLPLLSNTNCKKYWGTKIKDAM
181 ICAGASGVSSCMGDSGGPLVCKKNGAWTLVGIVSWGSSTCSTSTPGVYARVTALVNWVQQ
241 TLAAN

Let's use the ExPasy Prot Scale server to produce hydrophobicity plots of the protein bovine α-chymotrypsinogen. Input the Uniprot number **P00766** for the protein or the sequence into the appropriate boxes. Select a length of continuous amino acids (called a window) of 7, and the program will calculate an average hydrophobicity for the "window." The window slides down the linear sequence, and a new value is calculated to determine a series of values for the entire sequence. Hydropathy plots (average score for the midpoint amino acid in the window) for chymotrypsinogen (window of seven consecutive residues) are shown in Figure **3.1.6** below. The Kyte-Doolittle scale (+ is hydrophobic) shows many stretches with high average values. The amino acids at those positions are likely buried in the protein's interior. The Hoop-Woods value (+ represents hydrophilic) indicates stretches with high average values. These amino acids are likely exposed to water. The two plots are complementary.

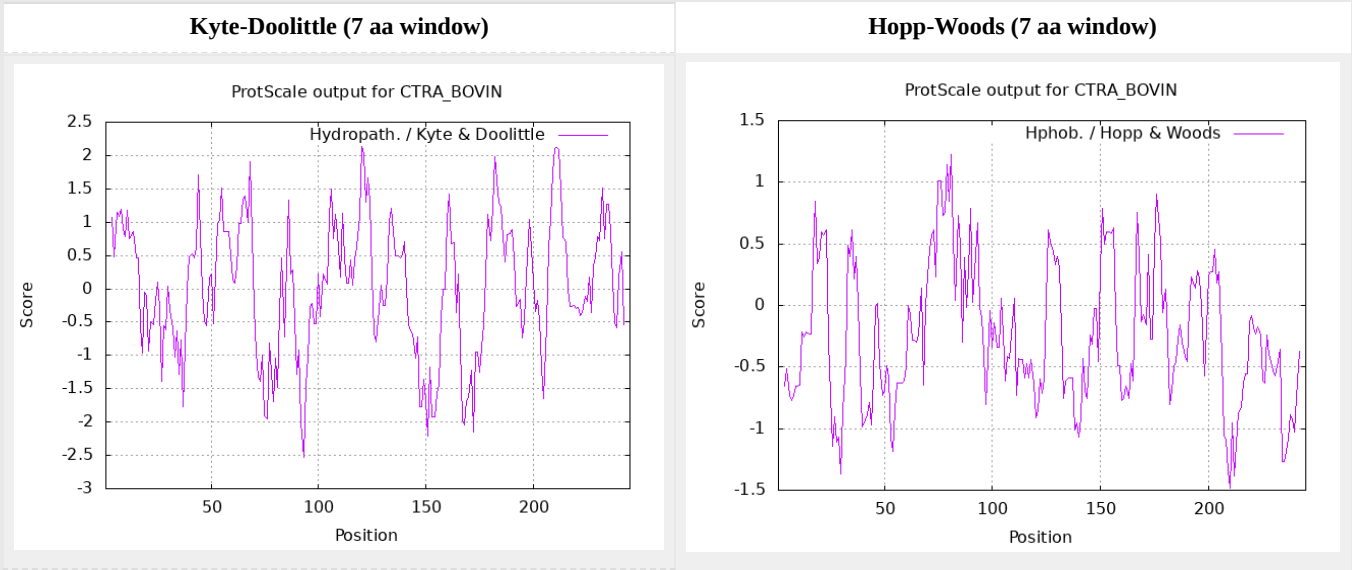


Figure **3.1.6**: Kyte-Doolittle and Hopp-Woods plots for bovine α-chymotrypsinogen

3.1.7: Amino Acid Stereochemistry

The amino acids are all ***chiral*** except glycine, whose side chain is H. A ***chiral molecule*** is not superimposable with its mirror image. Like left and right hands that have a thumb and fingers in the same order but are mirror images and not the same, chiral molecules have the same things attached in the same order but are mirror images and not the same. The mirror image versions of chiral molecules have nearly identical physical properties, making it very difficult to tell them apart or separate them. Because of this, they are given a special stereoisomer name called ***enantiomers,*** and the compounds themselves are given the same name! These molecules differ in how they rotate plane-polarized light and in their interactions with biological molecules. Molecules that rotate light in the right-handed direction are called dextrorotary and are given a small "d" letter designation. Molecules that rotate light to the left are called levorotary and are designated with a small "l" to distinguish one enantiomer from the other. Biochemists also use the older nomenclature of the large "L" and "D" forms to characterize the 3D stereochemistry of amino acids. All naturally occurring proteins from all living organisms consist of L-amino acids based on their structural similarities to L-glyceraldehyde.

Again, the d- and l-designations are specific terms for how a molecule rotates plane-polarized light. It does **not** denote the absolute stereo configuration of a molecule. An **absolute configuration** refers to the spatial arrangement of the atoms of a chiral molecular entity (or group) and its modern stereochemical description, e.g., **R** or **S**, referring to **Rectus** or **Sinister**, respectively. Absolute

configurations for a chiral molecule (in pure form) are often determined by X-ray crystallography. Alternative techniques are optical rotatory dispersion, vibrational circular dichroism, chiral shift reagents in proton NMR, and Coulomb explosion imaging. When the absolute configuration is known, the assignment of *R* or *S* is based on the **Cahn–Ingold–Prelog priority rules**. The absolute stereochemistry is related to L-glyceraldehyde, as shown below in Figure **3.1.6**.

All naturally occurring amino acids in proteins are L, which corresponds to the S isomer, except cysteine. As shown in the bottom left of Figure **3.1.7** below, the absolute configuration of the amino acids can be shown with the H pointed to the rear, the COOH groups pointing out to the left, the R group to the right, and the NH_3 group upwards. You can remember this with the mnemonic "CORN".

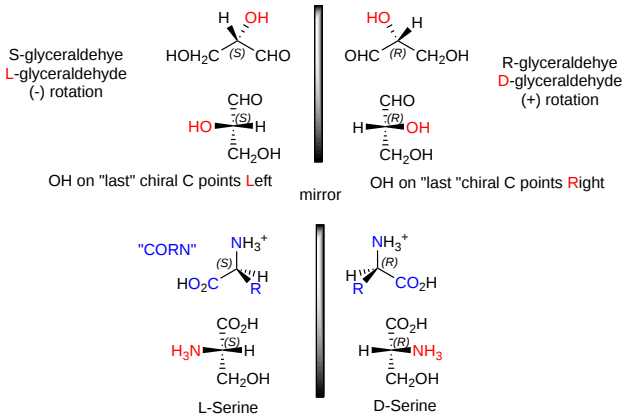


Figure **3.1.7**: Stereochemistry of amino acids

Why does Biochemistry still use D and L for sugars and amino acids? This explanation (taken from a website that may no longer be available, so no reference is available) seems reasonable.

"In addition, however, chemists often need to define a configuration unambiguously in the absence of any reference compound, and for this purpose, the alternative (R,S) system is ideal, as it uses priority rules to specify configurations. These rules sometimes produce absurd results when applied to biochemical molecules. For example, as we have seen, all common amino acids are L because they all have the same structure, including the position of the R group (if we just write the R group as R). However, they do not all have the same configuration in the (R, S) system: L-cysteine is also (R)-cysteine. Still, all the other L-amino acids are (S), but this reflects the human decision to give a sulfur atom a higher priority than a carbon atom and does not reflect a real difference in configuration. Worse problems can sometimes arise in substitution reactions: sometimes, inversion of configuration can result in no change in the (R) or (S) prefix, and sometimes, retention of configuration can result in a change of prefix.

It follows that it is not just conservatism or a failure to understand the (R, S) system that causes biochemists to continue using D and L; it is simply that the DL system meets their needs much better. As mentioned, chemists also use D and L when appropriate. The “explanation” given above of why the (R, S) system is little used in biochemistry is thus almost the exact opposite of reality. This system is the only practical way to unambiguously represent the stereochemistry of complex molecules with several asymmetric centers. Still, it is inconvenient with a regular series of molecules like amino acids and simple sugars."

If you are told to draw the correct stereochemistry of a molecule with one chiral C (S isomer, for example) and are given the substituents, you could do so easily following the R, S priority rules. However, how would you draw the correct isomer for the L isomer of the amino acid alanine? You couldn't do it without prior knowledge of the absolute configuration of the related molecule, L glyceraldehyde, or unless you remembered the anagram CORN. However, this disadvantage is more than made up for, as different L-amino acids with the same absolute stereochemistry might be labeled R or S, making this nomenclature unappealing to biochemists.

3.1.8: Amino Acid Charges

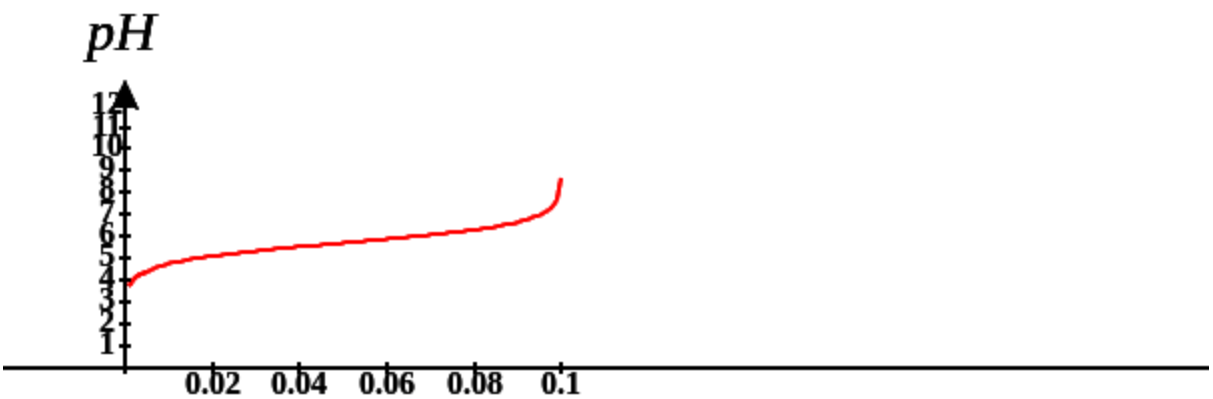
Monomeric amino acids have an alpha-amino group and a carboxyl group, both of which may be protonated or deprotonated, and an R group, some of which may be protonated or deprotonated. When protonated, the amino group carries a +1 charge, and the

carboxyl group carries a zero charge. When deprotonated, the amino group is neutral, while the carboxyl group carries a -1 charge. The R groups that can be protonated/deprotonated include Lys, Arg, and His, which have a + 1 charge when protonated, and Glu and Asp (carboxylic acids), Tyr and Ser (alcohols), and Cys (thiol), which have zero charges when protonated. Of course, when the amino acids are linked by peptide bonds (amide link), the alpha N and the carboxyl C are in an amide link and are not charged.

However, the amino group of the N-terminal amino acid and the carboxyl group of the C-terminal amino acid of a protein may be charged. The Henderson-Hasselbalch equation allows us to determine the charge state of any ionizable group given its pKa. Write each functional group capable of being deprotonated as an acid, HA, and the deprotonated form as A. The charge of HA and A can be determined for the functional groups using the Henderson-Hasselbalch equation from Chapter 2.2.

$$\mathrm{pH} = \mathrm{p}K_{\mathrm{a}} + \log \frac{[\mathrm{A}^-]}{[\mathrm{HA}]} \tag{3.1.1}$$

The titration curve for a single ionizable acid with different pKa values is shown below.



At the curve's inflection point, pH = pKa, the system is most resistant to changes in pH when either acid or base is added. At this pH, [HA]=[A^-].

The properties of a protein will be determined partly by whether the side chain functional groups, the N-terminal, and the C-terminal are charged or not. The HH equation tells us that this will depend on the pH and the pKa of the functional group.

- If the pH is 2 units below the pKa, the HH equation becomes -2 = log A/HA, or .01 = A/HA. This means the functional group will be about 99% protonated (with either 0 or +1 charge, depending on the functional group).
- If the pH is 2 units above the pKa, the HH equation becomes 2 = log A/HA, or 100 = A/HA. Therefore, the functional group will be 99% deprotonated.

- If the pH = pKa, the HH equation becomes 0 = log A/HA or 1 = A/HA. Therefore, the functional group will be 50% deprotonated.

From these simple examples, we have derived the +2 rule. This rule is used to quickly determine protonation, and hence charge state, and is extremely important to know (and easy to derive). Titration curves for Gly (no ionizable side chain), Glu (carboxylic acid side chain), and Lys (amine side chain) are shown in Figure **3.1.8**. You should be able to associate various sections of these curves with the titration of specific ionizable groups in the amino acids.

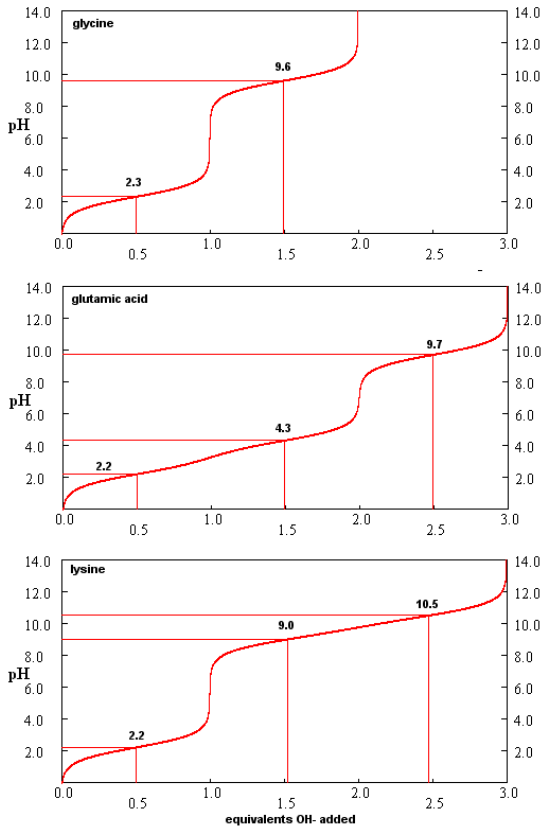


Figure **3.1.8**: Titration curves for Gly, Glu, and Lys

New 5/16/23: Download this Excel spreadsheet for Titration Curves for a Triprotic Acid. It has adjustable scroll bars to change pK_a values.

3.1.9: Buffer Review

The Henderson-Hasselbalch equation is also useful in calculating the composition of buffer solutions. Remember that buffer solutions are composed of a weak acid and its conjugate base. Consider the equilibrium for a weak acid, like acetic acid, and its conjugate base, acetate:

$$CH_3CO_2H + H_2O \rightleftharpoons H_3O^+ + CH_3CO_2^-$$

If the buffer solution contains equal concentrations of acetic acid and acetate, the pH of the solution is:

or pH = pKa + log [A]/[HA] = 4.7 + log 1 = 4.7

A look at the titration curve for the carboxyl group of Gly (see above) shows that when the pH = pKa, the slope of the curve (i.e. the change in pH on addition of base or acid) is at a minimum. As a general rule of thumb, buffer solutions can be made for a weak acid or base within ±1 pH units of its pKa. At pH = pKa, the buffer solution best resists the addition of either an acid or a base and has its greatest buffering capacity. The weak acid can react with the added strong base to form the weak conjugate base, and the conjugate base can react with the added strong acid to form the weak acid (as shown below), so pH changes with the addition of strong acid and base are minimized.

- addition of a strong base produces a weak conjugate base: $CH_3CO_2H + OH^- \leftrightarrow CH_3CO_2^- + H_2O$
- addition of strong acid produces weak acid: $H_3O^+ + CH_3CO_2 \rightarrow CH_3CO_2H + H_2O$

There are two simple ways to make a buffered solution. Consider an acetic acid/acetate buffer solution.

- Make equal molar solutions of acetic acid and sodium acetate and mix them, monitoring pH with a pH meter, until the desired pH is reached (+/- 1 unit from the pKa).
- Add acetic acid solution to NaOH in stoichiometric amounts until the desired pH is reached (+/- 1 unit from the pKa). In this method, you are forming the conjugate base, acetate, with the addition of NaOH:

$$CH_3CO_2H + OH^- \rightarrow CH_3CO_2^- + H_2O$$

- Buffers for pH control: Recipes based on pKas for acids, temperature, and ionic strength

3.1.10: Isoelectric Point

What happens if you have many ionizable groups in a single molecule, as is the case with a polypeptide or a protein? Consider a protein. At a pH of 2, all ionizable groups would be protonated, and the overall charge of the protein would be positive. (Remember, when carboxylic acid side chains are protonated, their net charge is 0.) As pH increases, the most acidic groups will begin to deprotonate, and the net charge will become less positive. At high pH, all the ionizable groups will become deprotonated in the strong base, and the overall charge of the protein will be negative. At some pH, then, the net charge will be 0. This pH is called the isoelectric point (pI). The pI can be determined by averaging the pKa values of the two groups closest to and on either side of the pI. One of the online problems will address this in more detail.

Remember that pKa is a measure of the equilibrium constant for the reaction. And, of course, you remember that $\Delta G^o = -RT \ln Keq$. Therefore, pKa is independent of concentration and depends only on the intrinsic stability of the reactants relative to the products. This is true only AT A GIVEN SET OF CONDITIONS, SUCH AS T, P, AND SOLVENT CONDITIONS.

Consider, for example, acetic acid, which in aqueous solution has a pKa of about 4.7. It is a weak acid, which dissociates only slightly to form H+ (in water, the hydronium ion, H_3O^+, is formed) and acetate (Ac^-). These ions are moderately stable in water but reassociate readily to form the starting product. The pKa of acetic acid in 80% ethanol is 6.87. This can be accounted for by the decrease in stability of the charged products, which are less shielded from each other by the less polar ethanol. Ethanol has a lower dielectric constant than water. The pKa increases to 10.32 in 100% ethanol and a whopping 130 in air!

Because amino acids are zwitterions and several also contain ionizable R-groups, their charge state *in vivo* —and thus their reactivity —can vary with pH, temperature, and the local microenvironment's solvation status. Table **3.1.2** shows the standard pK_a values for the amino acids and can be used to predict their ionization/charge status and their resulting peptides/proteins.

pK$_a$ Values for Common Alpha Amino Acids				
Amino Acid Type	Amino Acid	α-COOH	α-NH$_3^+$	RH or RH$^+$
Hydrophobic: Aliphatic	Glycine	2.34	9.60	
	Alanine	2.34	9.69	
	Valine	2.32	9.62	
	Leucine	2.36	9.68	
	Isoleucine	2.36	9.68	
	Proline	1.99	10.6	
	Methionine	2.28	9.21	
Hydrophobic: Aromatic	Phenylalanine	1.83	9.13	
	Tyrosine	2.2	9.11	10.07
	Tryptophan	2.38	9.39	
Hydrophilic: Polar Uncharged	Serine	2.21	9.15	
	Threonine	2.63	10.43	
	Cysteine	1.71	10.78	8.33
	Asparagine	2.02	8.8	
	Glutamine	2.17	9.13	
Hydrophilic: Acidic	Aspartic Acid	2.09	9.82	3.86
	Glutamic Acid	2.19	9.67	4.25
Hydrophilic: Basic	Arginine	2.17	9.04	12.48
	Histidine	1.82	9.17	6.00
	Lysine	2.18	8.95	10.53

Table **3.1.2**: Summary of pKas of amino acids

However, it should be noted that the solvation status in the microenvironment of an amino acid can alter the relative pK_a values of these functional groups and provide unique reactive properties within the active sites of enzymes. A more in-depth discussion of the effects of desolvation will be given in Chapter 6, discussing enzyme reaction mechanisms.

- *Printable Version of pKa Values*
- Expasy pI and molecular weight calculator for any protein sequence

3.1.11: Introduction to Amino Acid Reactivity

You should be able to identify which side chains contain H-bond donors and acceptors. Likewise, some are acids and bases. You should know the approximate pKas of the side chains and the N and C-terminal groups. Three of the amino acid side chains (Trp, Tyr, and Phe) contribute significantly to the UV absorption of a protein at 280 nm. This section will primarily address the chemical reactivity of the side chains, which is important for understanding protein properties. Many of the side chains are nucleophiles. Nucleophilicity measures how rapidly molecules with lone pairs of electrons can react in nucleophilic substitution reactions. It correlates with basicity, which measures the extent to which a molecule with lone pairs can react with an acid (Bronsted or Lewis). The properties of the atom that holds the lone pair are important in determining both nucleophilicity and basicity. In both cases, the atom must be willing to share its unbonded electron pair. If the atoms holding the nonbonded pair are more electronegative, they will be less likely to share electrons, and that molecule will be a poorer nucleophile (nu:) and weaker base. Using these ideas, it should be clear that RNH_2 is a better nucleophile than ROH, OH^- is better than H_2O, and RSH is better than H_2O. In the latter case, S is larger, and its electron cloud is more polarizable—hence, it is more reactive. The important side chain nucleophiles (in order from most to least nucleophilic) are Cys (RSH, pKa 8.5-9.5), His (pKa 6-7), Lys (pKa 10.5), and Ser (ROH, pKa 13). The side chain of serine is generally no more reactive than ethanol. It is a potent nucleophile in a certain class of proteins (such as proteases) when deprotonated. The amino group of lysine is a potent nucleophile only when deprotonated.

An understanding of the chemical reactivity of the various R group side chains of the amino acids in a protein is important since chemical reagents that react specifically with a given amino acid side chain can be used to:

- Identify the presence of the amino acids in unknown proteins or

- Determine if a given amino acid is critical for the structure or function of the protein. For example, suppose a reagent that covalently interacts with only Lys is found to inhibit the protein's function. Lysine might be considered important for the protein's catalytic activity in that case.

Figure **3.1.9** summarizes nucleophilic addition and substitution at carbonyl carbons.

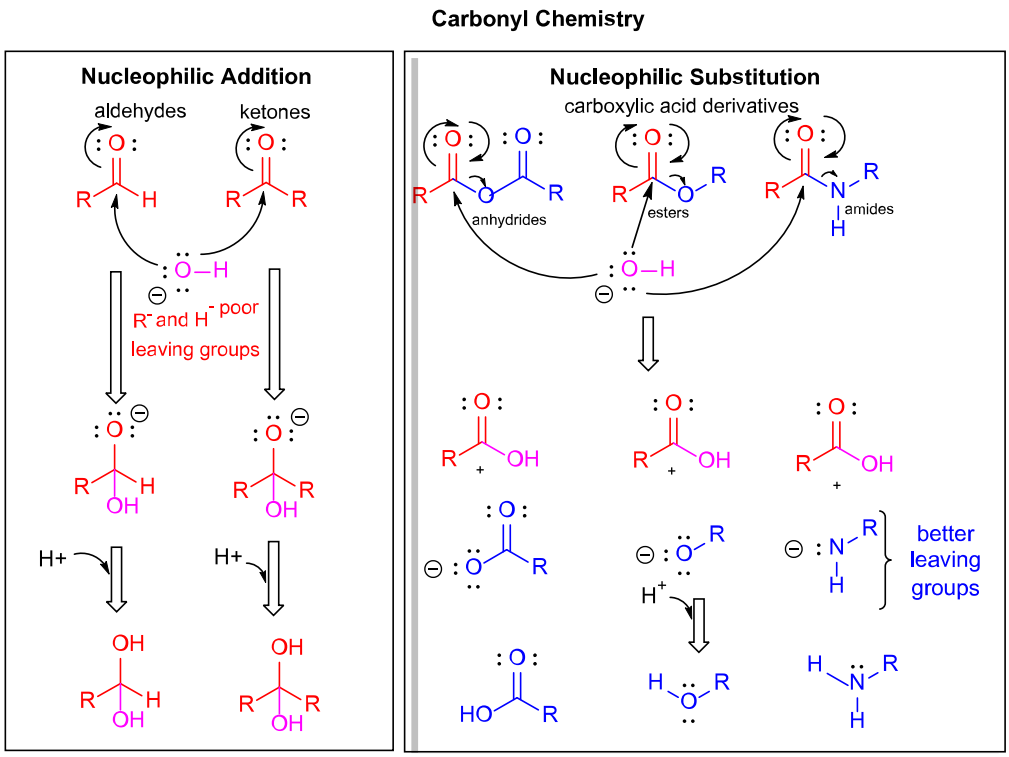


Figure **3.1.9**: A review summary of the chemistry of aldehydes, ketones, and carboxylic acid derivatives

The rest of the section will summarize the chemistry of the side chains of reactive amino acids. Historically, the function of a given amino acid in a protein has been studied by reacting it with side chain-specific chemical modifying agents. In addition, some side chains are covalently modified after synthesis in vivo (post-translational modifications—see below).

3.1.12: Reactions of Lysine

Figure **3.1.10** shows the reaction of lysine with anhydrides and ethylacetimidate.

- reacts with anhydrides in a nucleophilic substitution reaction (acylation).
- reacts reversibly with methylmaleic anhydride (also called citraconic anhydride) in a nucleophilic substitution reaction.
- reacts with high specificity with ethylacetimidate in a nucleophilic substitution reaction (ethylacetimidate is like ethylacetate, only with an imido group replacing the carbonyl oxygen). Ethanol leaves as the amidino group forms. (has two N -i.e., din - attached to the C)

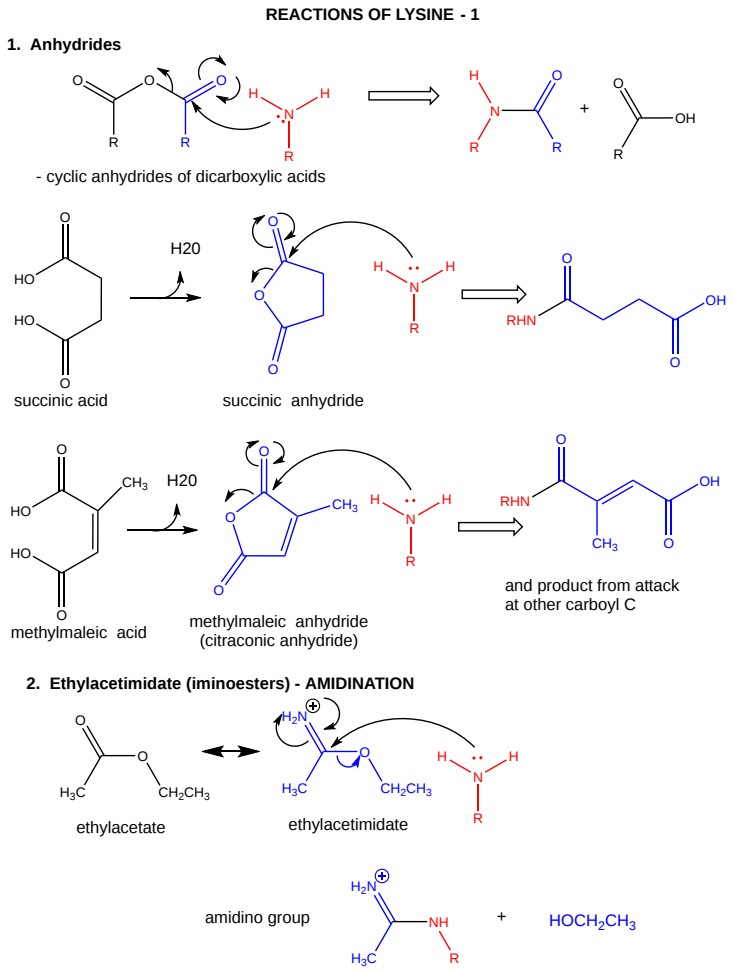


Figure 3.1.10: Reaction of lysine with anhydrides and ethylacetimidate.

Figure 3.1.11 shows a second set of common reactions of lysine, including those used to attach a chromophore or a fluorescent label to the side chain.

- reacts with O-methylisourea in a nucleophilic substitution reaction with the expulsion of methanol to form a guanidino group (has 3 N attached to C, nidi)
- reacts with fluorodinitrobenzene (FDNB or Sanger's reagent) or trinitrobenzenesulfonate (TNBS, as we saw with the reaction with phosphatidylethanolamine) in a nucleophilic aromatic substitution reaction to form 2,4-DNP-lysine or TNB-lysine.
- reacts with dimethylaminonapthelenesulfonylchloride (Dansyl Chloride) in a nucleophilic substitution reaction.

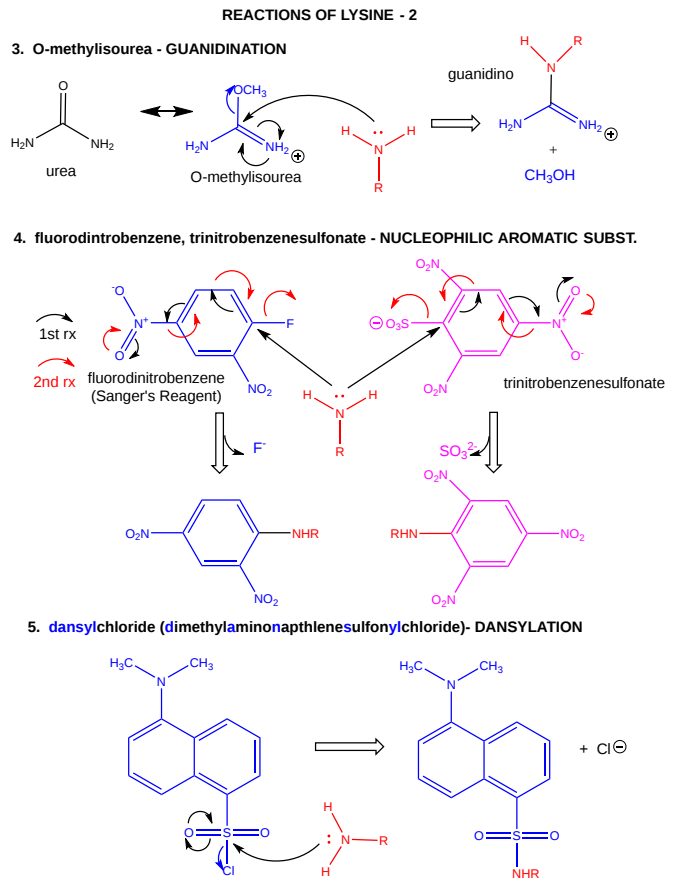


Figure 3.1.11: Reaction of lysine with O-methylisourea, chromophores, and fluorophores

Figure 3.1.12 shows a final common reaction we will encounter: the formation of an imine or Schiff base on the reaction of lysine with an aldehyde or ketone.

- reacts with high specificity toward aldehydes to form imines (Schiff bases), which can be reduced with sodium borohydride or cyanoborohydride to form a secondary amine.

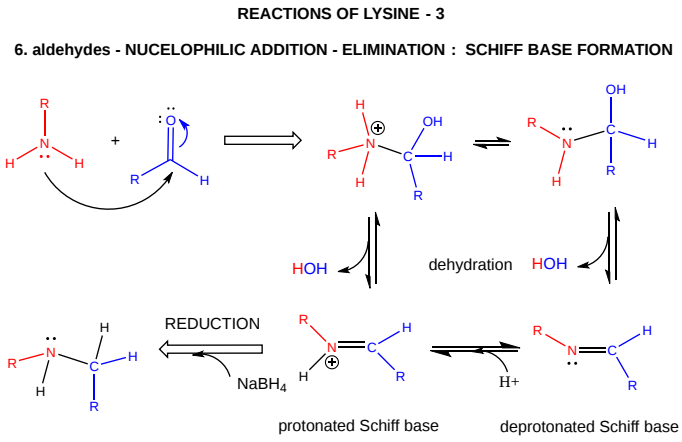


Figure 3.1.12: Reaction of lysine with an aldehyde or ketone to form a Schiff base

3.1.13: Reactions of Cysteine

Cysteine is a potent nucleophile that often forms a covalent disulfide bond with another Cys.

Figure 3.1.13 shows common reagents used in the lab to label free Cys side chains. These reagents are used to alter Cys side chains to determine if they have functional significance in a protein (such as an active nucleophile in an enzyme-catalyzed reaction).

- reacts with iodoacetic acid in an S_N2 reaction, adding a carboxymethyl group to the S.
- reacts with iodoacetamide in an S_N2 reaction, adding a carboxyamidomethyl group to S.
- reacts with N-ethylmaleimide in an addition reaction to the double bond

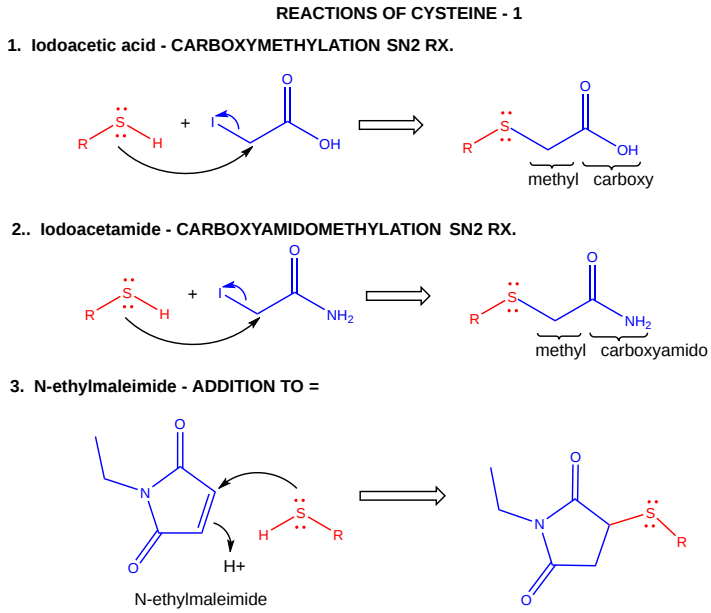


Figure 3.1.13: Common labeling reactions of cysteine

Sulfur is directly below oxygen in the periodic table, and, in analogy to water, sulfur-containing amino acids are found in different redox states, as illustrated in Figure 3.1.14.

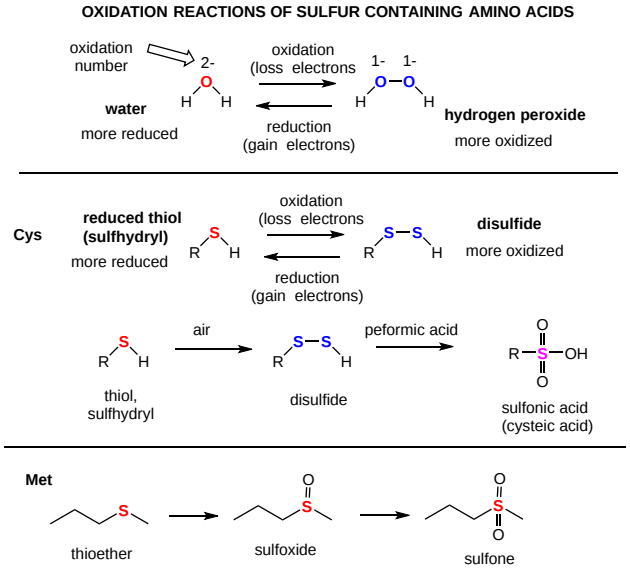


Figure 3.1.14: Oxidation states of sulfur

3.1.14: Cystine Chemistry

Two cysteine side chains can covalently interact in a protein to form a disulfide bond (RS-SR), called cystine. Just as HOOH (hydrogen peroxide) is more oxidized than HOH (O in H_2O_2 has an oxidation number of 1- while the O in H_2O has an oxidation number of -2), RSSR is the oxidized form (S oxidation number -1), and RSH is the reduced form (S oxidation number -2) of thiols. Their oxidation numbers are analogous since O and S are both in Group 6 of the periodic table and are more electronegative than C.

Cystine can react with a free sulfhydryl (RSH) in a thermodynamically non-challenging disulfide **exchange** reaction, which, when conducted with excess free sulfhydryls, results in the reduction of cystine in the protein, as shown in Figure 3.1.15.

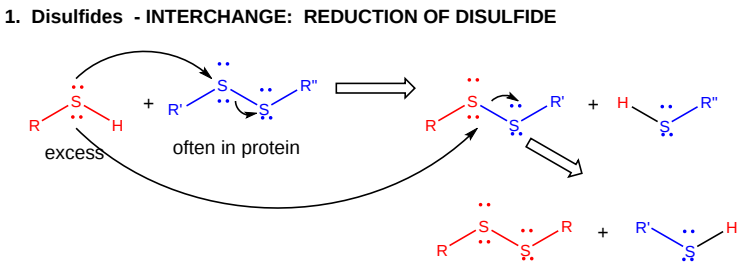


Figure 3.1.15: Disulfide interchange and reduction of protein disulfides

This reaction is often used in the lab to quantify the amount of free cysteine side chains in a protein using Ellman's reagent, as shown in Figure 3.1.16.

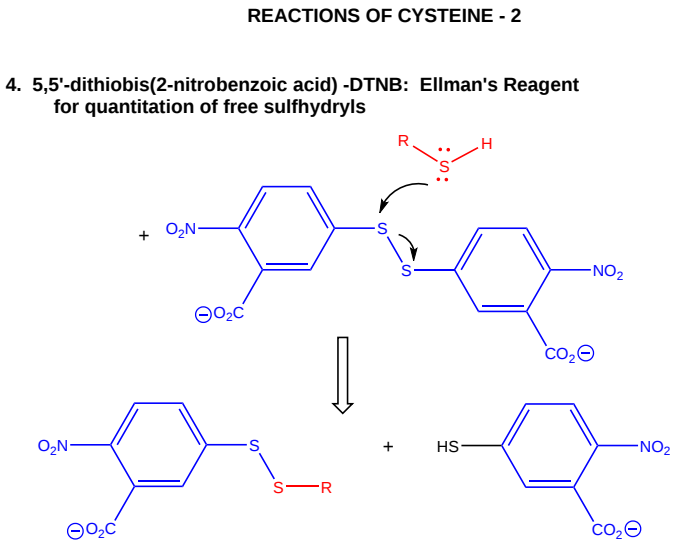


Figure 3.1.16: Reaction of free cysteine with Ellman's reagent.

The 2-nitro-5-thiobennzoate anion leaving group absorbs at 412 nm, making quantitation easy. However, unless the protein is unfolded to expose all the cysteines, only surface-exposed, not buried, free cysteines will be labeled.

When a protein folds, two Cys side chains might approach each other and form an intra-chain disulfide bond. Likewise, two Cys side chains on separate proteins might approach each other and form an inter-chain disulfide. For protein structure analysis, disulfides are typically cleaved, and the chains are separated. The disulfides can be cleaved by reducing agents such as beta-mercaptoethanol, dithiothreitol, or tris(2-carboxyethyl)phosphine (TCEP), or by oxidizing agents such as performic acid, which further oxidizes the disulfide to separate cysteic acids. Three common reagents used in disulfide cleavage reactions in the lab are shown in Figure 3.1.17.

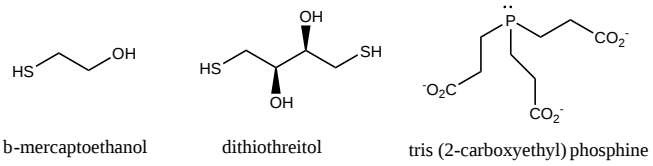


Figure 3.1.17: Three common disulfide-cleaving (reducing) agents used in the lab

The reaction for beta-mercaptoethanol (BME) and performic acid is shown in Figure 3.1.18 below.

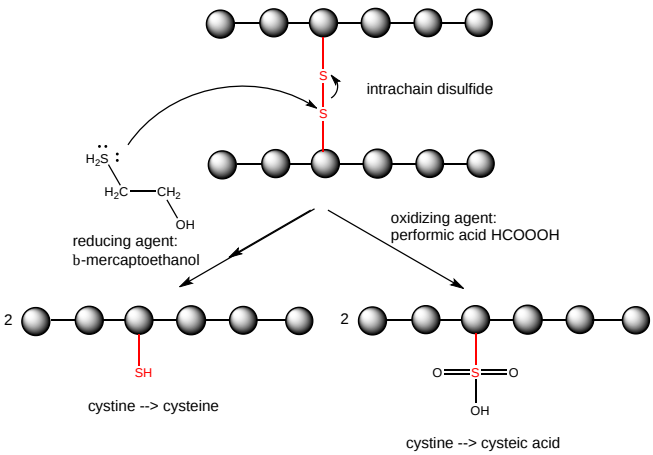


Figure 3.1.18: Cleavage of intrachain cystine disulfide bonds in proteins by beta-mercaptoethanol and performic acid

Figure 3.1.19 shows the reaction for dithiothreitol (DTT). Note that it forms a stable cyclohexane-like ring, which thermodynamically favors this reaction. It does not require as much excess DTT as the reaction with BME.

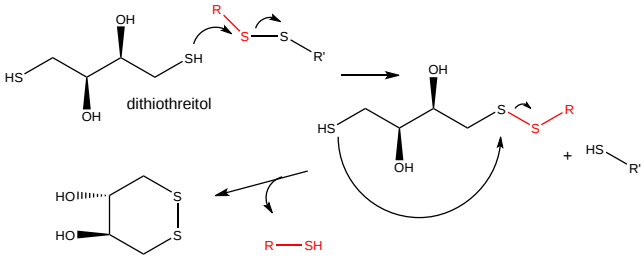


Figure 3.1.19: Cleave of disulfides with dithiothreitol

The reaction with tris (2-carboxyethyl) phosphine (TCEP) is not a disulfide interchange reaction, as is shown in Figure 3.1.20.

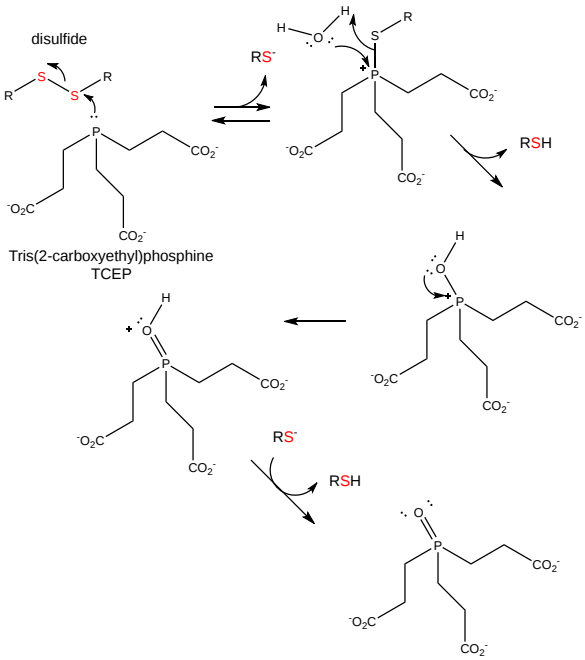


Figure 3.1.20: Reaction of TCEP with disulfides

Cells maintain a reducing environment using many "reducing" agents, such as the tripeptide gamma-Glu-Cys-Gly (glutathione). Hence, intracellular proteins usually do **not** contain disulfides, which are abundant in extracellular proteins (such as those found in blood) or in certain organelles, such as the endoplasmic reticulum and mitochondrial intermembrane space, where disulfides can be introduced.

Sulfur redox chemistry is very important biologically. As described above, the sulfur in cysteine is redox-active. Hence, it can exist in various states, depending on the local redox environment and the presence of oxidizing and reducing agents. A potent oxidizing agent produced in cells is hydrogen peroxide, which can lead to more drastic, irreversible chemical modifications of Cys side chains. Suppose a reactive Cys is important to protein function. In that case, the function of the protein can be modulated (sometimes reversibly, sometimes irreversibly) with various oxidizing agents, as shown in Figure **3.1.21**.

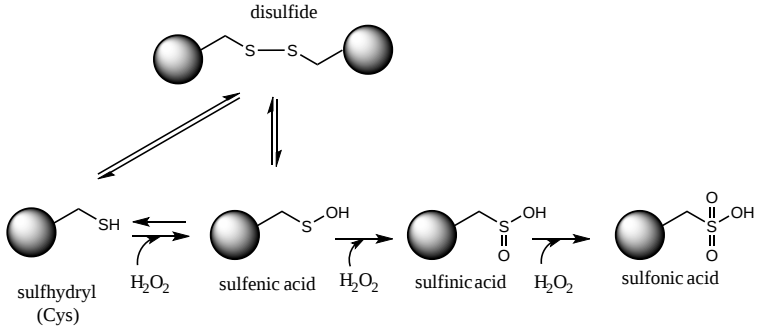


Figure **3.1.20**: Reaction of Cysteine with H_2O_2

3.1.15: Reactions of Histidine

Histidine is one of the most important bases at physiological pH. Remember from introductory chemistry that for any acid/conjugate base pair, the pKa and pKb of the acid and base are related by this expression:

$$pKa + pKb = 14$$

Table **3.1.3** below shows the pKa and pKb of three amino acids.

Table **3.1.3**: pKa and pKb values for three amino acid side chains

amino acid	pKa	pKb
Histidine	6.5	7.5
Lysine	10.5	3.5
Arginine	12.5	1.5

The deprotonated forms of Lys and Arg with lower pKbs are much stronger bases than the deprotonated form of His, so at physiological pH, they would always be protonated (unless their local environment lowers their pKa and pKb values). In contrast, His exists in both protonated and deprotonated states at physiological pH, so it can readily gain a proton and act as a general base in reactions.

Histidine can exist as two tautomers, as shown in Figure **3.1.22**. NMR studies show that in model peptides, the proton is predominantly on the ε_2, N_3, or tele N in the imidazole ring, as it has a pK_a 0.6 units higher than δ_1, N_1, or pro N.

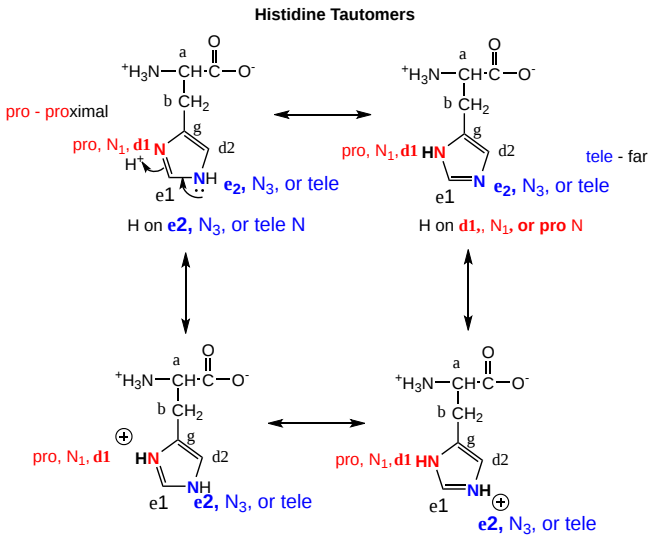


Figure 3.1.22: Histidine tautomers

The nitrogen atom in a secondary amine might be expected to be a stronger nucleophile than a primary amine through electron release to that N in a secondary amine. Opposing this effect is the steric hindrance by the two attached Cs of the N on attachment to an electrophile. However, in His, this steric effect is minimized since the ring restrains the 2Cs. With a pKa of about 6.5, this amino acid is one of the strongest available bases at physiological pH (7.0). Hence, it can often cross-react with many reagents used to modify Lys side chains. Histidine reacts with reasonably high selectivity with diethylpyrocarbonate as shown in Figure 3.1.23.

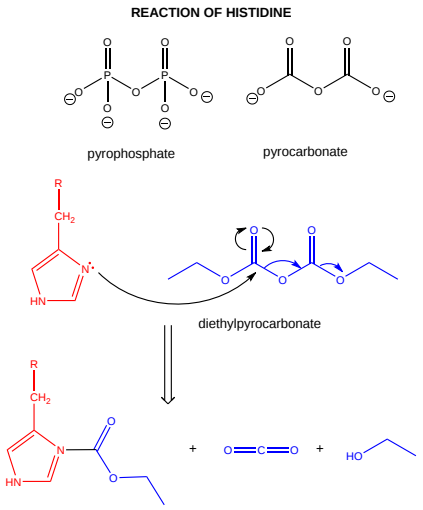


Figure 3.1.23: Reaction of histidine with diethylpyrocarbonate

3.1.16: In vivo Post-Translational Modification of Amino Acids

Amino acids in naturally occurring proteins are also subjected to chemical modifications within cells. These modifications alter the properties of the amino acid, thereby affecting the protein's structure and function. Most chemical modifications of proteins within cells occur after they are synthesized through a process called translation. The resulting chemical changes are termed post-translational modifications. Several are shown in Figure 3.1.24. Simple acid/base reactions are included, but these are not considered examples of post-translational modifications.

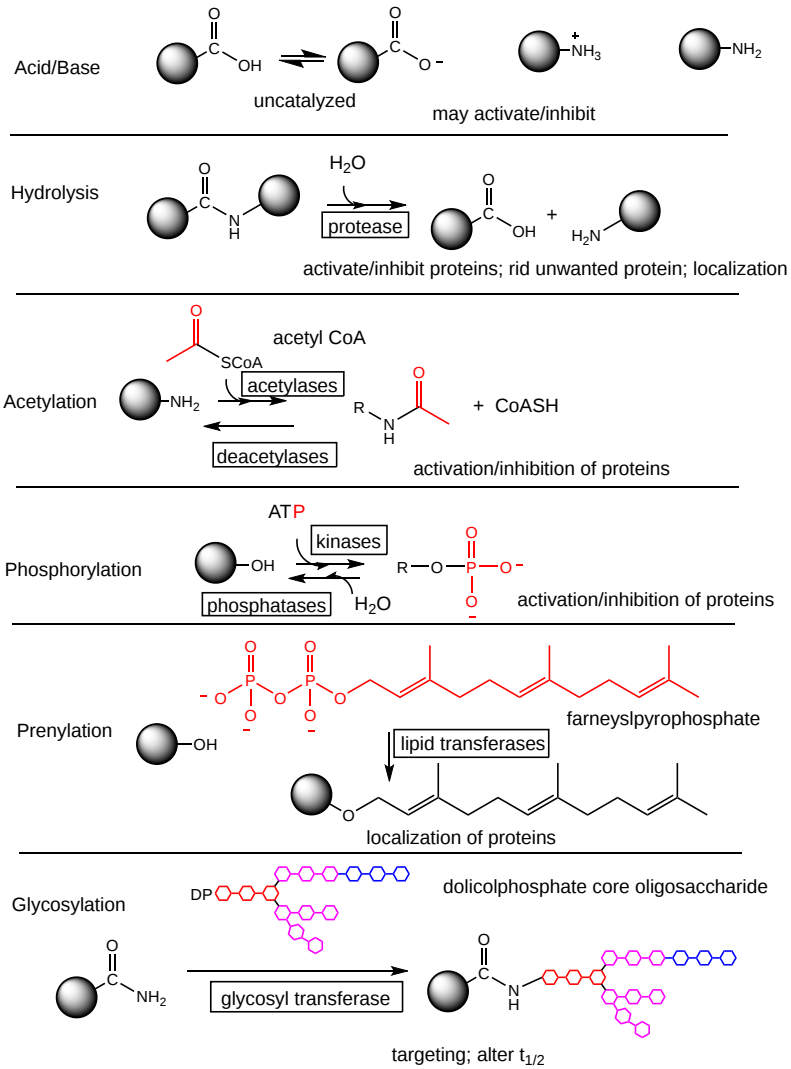


Figure 3.1.24: Common post-translational modifications of protein

There are 100s of PTMs, and many are part of an elaborate system within a cell to respond to both external (hormones, neurotransmitters, nutrients, metabolites) and internal chemical signals. PTMs (such as phosphorylation and acetylation) and their removal by enzymes are part of an elaborate cell signaling system that we will explore in great detail in Chapter 28. However, not all PTMs are benign. Examples include glycation, oxidation, citrullination, and carbonylation of protein side chains. These are often increased during periods of inflammatory stress (both acute and chronic). These modified proteins are degraded within the cell to short peptides that retain the chemical modification. Unfortunately, these can be recognized by the immune system as foreign, triggering an immune response against the body's own tissues and leading to autoimmune disease. One potentially deleterious PTM is the carboxyethylation of cysteine, catalyzed by the enzyme cystathionine β-synthase as shown in Figure 3.1.25 below.

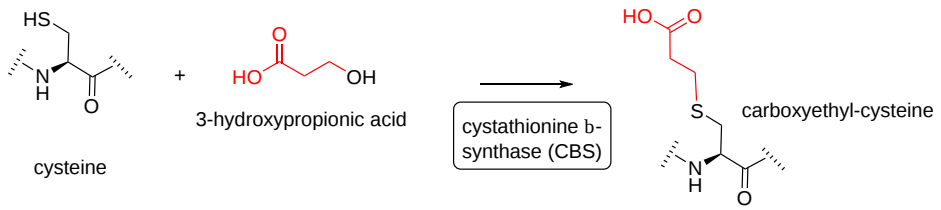


Figure 3.1.25: Carboxyethylation of cysteine

The product is very similar to the carboxymethylation of cysteine shown in Figure 13 above. The modifying reagent, 3-hydroxypropionic acid, is a metabolite produced by gut microbes. This modification has been shown to produce an autoimmune response in the disease ankylosing spondylitis.

3.1.17: Summary

This chapter provides an extensive overview of proteins, their building blocks, and the chemical principles that govern their structure and function. It begins by emphasizing that proteins are not only among the most abundant macromolecules in living systems, but also the most functionally diverse. Proteins serve many roles—including structural support, catalysis, regulation, transport, and defense—and their diversity arises from the linear polymerization of alpha-amino acids linked via peptide bonds.

Key topics include:

- **Alpha-Amino Acid Structure and Protein Primary Structure:**
  Every protein is composed of 20 naturally occurring alpha-amino acids, each sharing a common backbone (an α-amino group, an α-carboxyl group, and a hydrogen atom) and a unique R-group that defines its chemical properties. The chapter clarifies the process of peptide bond formation (and hydrolysis) through nucleophilic attack, and it dispels common misconceptions about the free nature of the terminal groups in a protein chain.
- **Amino Acid Side Chain Properties:**
  The diverse chemical nature of proteins is largely due to the variability of the amino acid side chains, which can be classified as nonpolar (hydrophobic), polar uncharged, acidic, or basic. These classifications are critical for understanding protein folding, stability, and function. The chapter introduces hydropathy scales, such as the Kyte-Doolittle and Hopp-Woods scales, to quantitatively assess the relative hydrophobicity or hydrophilicity of side chains, providing insight into protein topology.
- **Stereochemistry and Chirality:**
  Except for glycine, all amino acids are chiral and naturally occur as L isomers in proteins. The chapter contrasts the D/L nomenclature with the R/S system and explains why biochemists favor the former, particularly when describing the stereochemistry of amino acids in the context of protein structure.
- **Acid-Base Chemistry and Ionization:**
  The ionizable groups on amino acids, including the α-amino, α-carboxyl, and various side chains, determine the overall charge and reactivity of proteins. The Henderson–Hasselbalch equation is introduced as a tool to predict the protonation state of these groups under varying pH conditions, leading to the concept of the isoelectric point (pI) — the pH at which a protein has no net charge.
- **Chemical Reactivity of Side Chains:**
  Beyond structural roles, the chemical reactivity of amino acid side chains is central to protein function. The chapter details how key side chains—such as those of lysine, cysteine, and histidine—participate in nucleophilic reactions, form covalent modifications (e.g., Schiff bases), and are targeted by specific chemical reagents. This reactivity is foundational for both enzyme catalysis and experimental methods used to probe protein structure and function.
- **Post-Translational Modifications (PTMs):**
  The chapter concludes by discussing PTMs, chemical modifications that occur after protein synthesis. These modifications, including phosphorylation, acetylation, glycosylation, and oxidation, can dramatically alter a protein's activity, stability, and interactions. While many PTMs are critical for normal cellular signaling, aberrant modifications may contribute to disease.

Overall, the chapter integrates structural biology with chemical reactivity, providing a comprehensive framework that links the primary structure of proteins to their three-dimensional architecture and dynamic functions in the cell. This foundation is essential for understanding advanced topics in enzymology, protein engineering, and molecular regulation.

This page titled 3.1: Amino Acids and Peptides is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

3.2: Amino Acid Charges

Monomeric amino acids have an alpha amino group and a carboxyl group, both of which may be protonated or deprotonated, and a R group, some of which may be protonated or deprotonated. When protonated, the amino group has a +1 charge, and the carboxyl group a 0 charge. When deprotonated the amino group has no charge, while the carboxyl group has a -1 charge. The R groups which can be protonated/deprotonated include Lys, Arg and His, which have a + 1 charge when protonated, and Glu and Asp (carboxylic acids), Tyr and Ser (alcohols) and Cys (thiol), which have 0 charge when protonated. Of course, when the amino acids are linked by peptide bonds (amide link), the alpha N and the carboxyl C are in an amide link, and are not charged. However, the amino group of the N -terminal amino acid and the carboxyl group of the C-terminal amino acid of a protein may be charged. The Henderson Hasselbach equation gives us a way to determine the charge state of any ionizable group knowing the pKa of the group. Write each functional group capable of being deprotonated as an acid, HA, and the deprotonated form as A. The charge of HA and A will be determined by the functional group. The Ka for the reaction is:

$$K_a = \frac{[H_3O^+][A]}{[HA]} \tag{3.2.1}$$

or

$$[H_3O^+] = K_a \frac{[HA]}{[A]} \tag{3.2.2}$$

$$-\log[H_3O^+] = -\log K_a + \log \frac{[A]}{[HA]} \tag{3.2.3}$$

or

$$pH = pK_a + \log \frac{[A]}{[HA]} \tag{3.2.4}$$

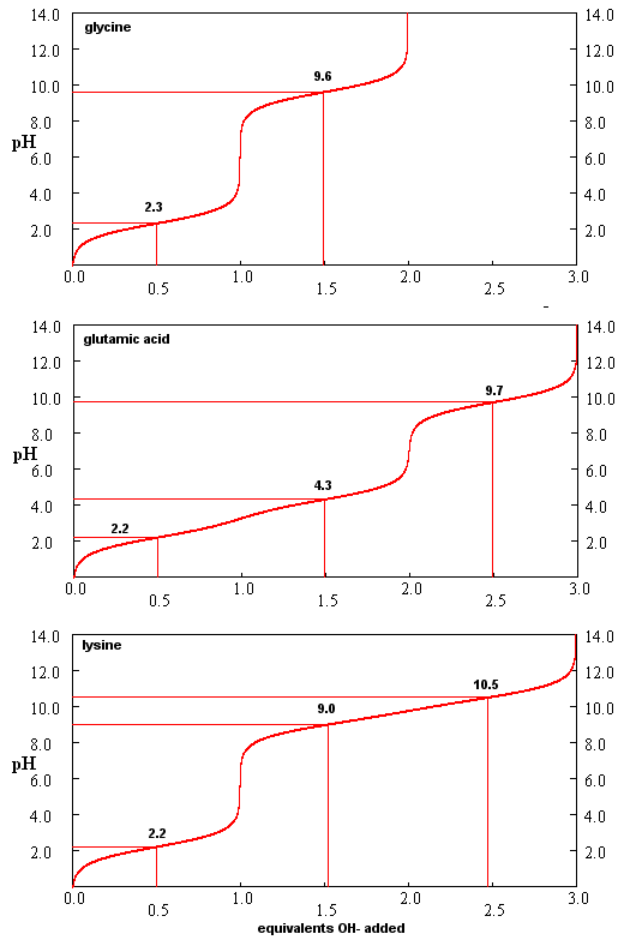
This is the (in)famous Henderson-Hasselbach (HH) equation.

The properties of a protein will be determined partly by whether the side chain functional groups, the N terminal, and the C terminal are charged or not. The HH equation tells us that this will depend on the pH and the pKa of the functional group.

- If the pH is 2 units below the pKa, the HH equation becomes, -2 = log A/HA, or .01 = A/HA. This means that the functional group will be about 99% protonated (with either 0 or +1 charge, depending of the functional group).
- If the pH is 2 units above the pKa, the HH equation becomes 2 = log A/HA, or 100 = A/HA. Therefore the functional group will be 99% deprotonated.
- If the pH = pka, the HH equation becomes 0 = log A/HA, or 1 = A/HA. Therefore the functional group will be 50% deprotonated

From these simple examples, we have derived the +2 rule. This rule is used to quickly determine protonation, and hence charge state, and is extremely important to know (and easy to derive). Titration curves for Gly (no ionizable) side chain, Glu (carboxlic acid side chain) and Lys (amine side chain) are shown below. You should be able to associate various sections of these curves with titration of specific ionizable groups in the amino acids.

Figure: Titration curves for Gly, Glu, and Lys



Buffer Review

The Henderson-Hasselbach equation is also useful in calculating the composition of buffer solutions. Remember that buffer solutions are composed of a weak acid and its conjugate base. Consider the equilibrium for a weak acid, like acetic acid, and its conjugate base, acetate:

$$CH_3CO_2H + H_2O \Leftrightarrow H_3O^+ + \text{msubsup} \tag{3.2.5}$$

If the buffer solution contains equal concentrations of acetic acid and acetate, the pH of the solution is:

$$pH = pK_a + \log\frac{[A]}{[HA]} = 4.7 + \log 1 = 4.7 \tag{3.2.6}$$

A look at the titration curve for the carboxyl group of Gly (see above) shows that when the pH = pKa, the slope of the curve (i.e. the change in pH with addition of base or acid) is at a minimum. As a general rule of thumb, buffer solution can be made for a weak acid/base in the range of +/- 1 pH unit from the pKa of the weak acids. At the pH = pKa, the buffer solution best resists addition of either acid and base, and hence has its greatest buffering ability. The weak acid can react with added strong base to form the weak conjugate base, and the conjugate base can react with added strong acid to form the weak acid (as shown below) so pH changes on addition of strong acid and base are minimized.

- addition of strong base produces weak conjugate base: $CH_3CO_2H + OH^- \rightarrow$ msubsup $+H_2O$
- addition of strong acid produces weak acid: $H_3O^+ +$ msubsup $\rightarrow CH_3CO_2H + H_2O$

There are two simples ways to make a buffered solution. Consider an acetic acid/acetate buffer solution.

- make equal molar solution of acetic acid and sodium acetate, and mix them, monitoring pH with a pH meter, until the desired pH is reached (+/- 1 unit from the pKa).

- take a solution of acetic acid and add NaOH at substoichiometric amounts until the desired pH is reached (+/- 1 unit from the pKa). In this method you are forming the conjugate base, acetate, on addition of the weak base:

$$CH_3CO_2H + OH^- \rightarrow \text{msubsup} + H_2O \tag{3.2.7}$$

- EXTERNAL Buffers for pH control: Recipes based on pKas for acids, temperature, and ionic strength

Isoelectric Point

What happens if you have many ionizable groups in a single molecule, as is the case with a polypeptide or protein. Consider a protein. At a pH of 2, all ionizable groups would be protonated, and the overall charge of the protein would be positive. (Remember, when carboxylic acid side chains are protonated, their net charge is 0.) As the pH is increased, the most acidic groups will start to deprotonate and the net charge will become less positive. At high pH, all the ionizable groups will become deprotonated in the strong base, and the overall charge of the protein will be negative. At some pH, then, the net charge will be 0. This pH is called the isoelectric point (pI). The pI can be determined by averaging the pKa values of the two groups which are closest to and straddle the pI. One of the online problems will address this in more detail

- LIst of pI and MW for proteins derived from 2D gels

Remember that pKa is really a measure of the equilibrium constant for the reaction. And of course, you remember that DGo = -RT ln Keq. Therefore, pKa is independent of concentration, and depends only on the intrinsic stability of reactants with respect to the products. This is true only AT A GIVEN SET OF CONDITIONS, SUCH AS T, P, AND SOLVENT CONDITIONS.

Consider, for example acetic acid, which in aqueous solution has a pKa of about 4.7. It is a weak acid, which dissociates only slightly to form H+ (in water the hydronium ion, H3O+, is formed) and acetate (Ac-). These ions are moderately stable in water, but reassociate readily to form the starting product. The pKa of acetic acid in 80% ethanol is 6.87. This can be accounted for by the decrease in stability of the charged products which are less shielded from each other by the less polar ethanol. Ethanol has a lower dielectric constant than does water. The pKa increases to 10.32 in 100% ethanol, and to a whopping 130 in air!

- A great interactive web site: Amino Acid Acid/Base Titration Curves
- pI calculator for any protein sequence
- Amino Acid Repository: Properties of Amino Acids

Contributors and Attributions

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

This page titled 3.2: Amino Acid Charges is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

- **A3. Amino Acid Charges** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.
- **3.2: Amino Acid Charges** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.

3.3: Cysteine Chemistry

Two cysteine side chains can covalently interact in a protein to produce a disulfide. Just as HOOH (hydrogen peroxide) is more oxidized than HOH, RSSR is the oxidized form and RSH is the reduced form of thiols.

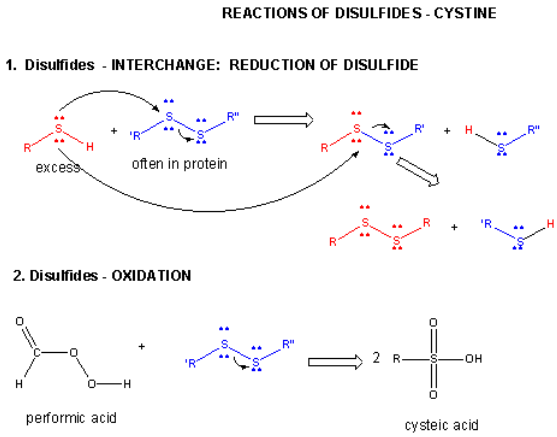


Figure: DISULFIDE - CYSTINE - REACTIONS

When a protein folds, two Cys side chains might approach each other, and form an intrachain disulfide bond. Likewise, two Cys side chains on separate proteins might approach each other and form an interchain disulfide. Such disulfides must be cleaved, and the chains separated before analyzing the sequence of the protein. The disulfide in protein can be cleaved by reducing agents such as beta-mercaptoethanol or dithiothreitol (DTT).

The inside of cells are maintained in a reduced environment by the presence of many "reducing" agents, such as the tripeptide g-glu-cys-gly (glutathione). Hence intracellular proteins usually do not contain disulfides, which are abundant in extracellular proteins (such as those found in blood) or in certain organelles such as the endoplasmic reticulum and mitochondrial intermembrane space where disulfides can be introduced.

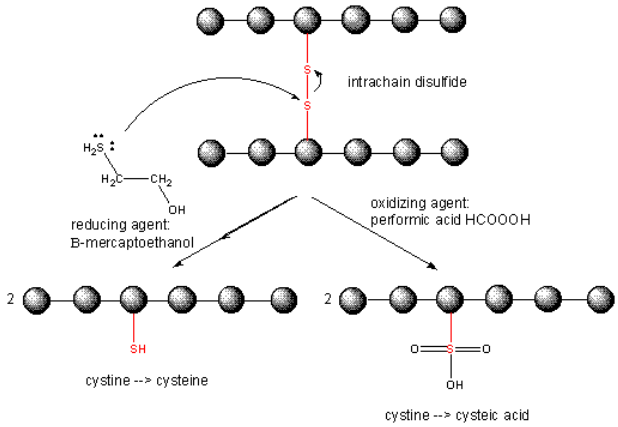


Figure: Cleaving Disulfide Bonds in Proteins

Cysteine Redox Chemistry

The sulfur in cysteine is redox-active and hence can exist in a wide variety of states, depending on the local redox environment and the presence of oxidizing and reducing agents. A potent oxidizing agent that can be made in cells is hydrogen peroxide, which can lead to more drastic and irreversible chemical modifications to the Cys side chains. If a reactive Cys is important to protein function, then the function of the protein can be modulated (sometimes reversibly, sometimes irreversibly) with various oxidizing agents, as shown in the figure below.

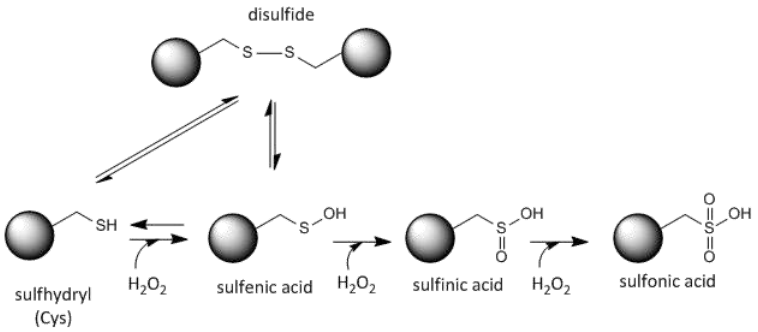


Figure: Redox state of Cysteine

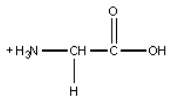
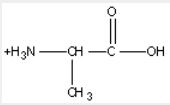
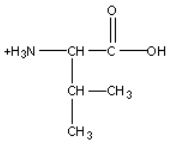
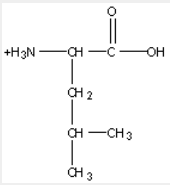
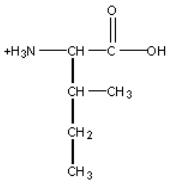
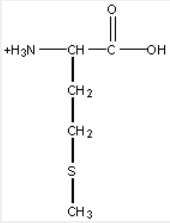
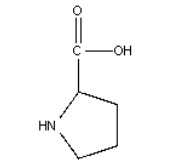
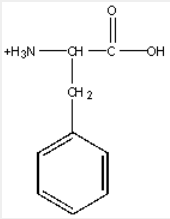
Contributors and Attributions

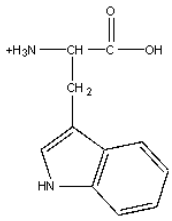
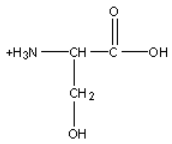
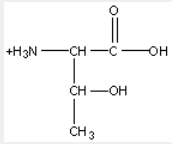
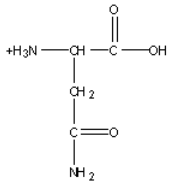
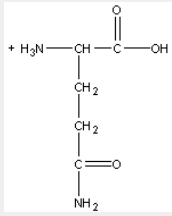
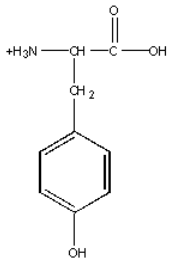
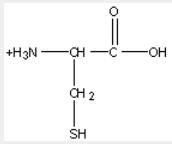
- - Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

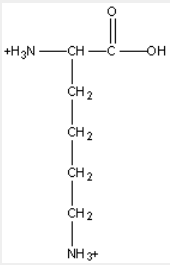
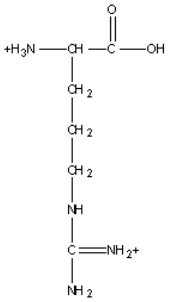
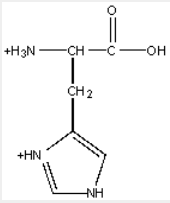
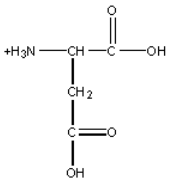
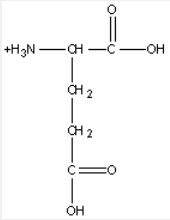
This page titled 3.3: Cysteine Chemistry is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

- **A7. Cysteine Chemistry** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.
- **3.3: Cysteine Chemistry** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.

3.4: 3.1-Structure and Property of the Naturally-Occurring Amino Acids

Amino Acid	3 Letter Code	1 Letter Code	Structure	pK_{a1} (α-carboxyl)	pk_{a2} (α-amino)	pk_{a3} (side chain)
Nonpolar Side Chains						
Glycine	Gly	G		2.35	9.78	.
Alanine	Ala	A		2.35	9.87	.
Valine	Val	V		2.29	9.74	.
Leucine	Leu	L		2.33	9.74	.
Isoleucine	Ile	I		2.32	9.76	.
Methionine	Met	M		2.13	9.28	.
Proline	Pro	P		1.95	10.64	.
Phenylalanine	Phe	F		2.20	9.31	.

Tryptophan	Trp	W		2.46	9.41	.
Uncharged Polar Side Chains						
Serine	Ser	S		2.19	9.21	.
Threonine	Thr	T		2.09	9.10	.
Asparagine	Asn	N		2.14	8.72	.
Glutamine	Gln	Q		2.17	9.13	.
Tyrosine	Tyr	Y		2.20	9.21	10.46
Cysteine	Cys	C		1.92	10.70	8.37
Charged Polar Side Chains10.54						

Lysine	Lys	K		2.16	9.06	10.54
Arginine	Arg	R		1.82	8.99	12.48
Histidine	His	H		1.80	9.33	6.04
Aspartic Acid	Asp	D		1.99	9.90	3.90
Glutamic Acid	Glu	E		2.10	9.47	4.07

Contributors and Attributions

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

This page titled 3.4: 3.1-Structure and Property of the Naturally-Occurring Amino Acids is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

- **Structure and Property of the Naturally-Occurring Amino Acids** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.

3.5: From Amino Acids to Peptides

Proteins are polymers of a bifunctional monomer, the amino acid. The twenty common naturally-occurring amino acids each contain an a-carbon, an a-amino group, an a-carboxylic acid group, and an a-side chain or side group. These side chains (or R groups) may be either nonpolar, polar and uncharged, or charged, depending on the pH and pKa of the ionizable group. Two other amino acids occasionally appear in proteins. One is selenocysteine, which is found in Arachea, eubacteria, and animals. Another just recently found is pyrrolysine, found in Arachea. Shultz et al. have gone one step further. They have engineered bacterial to incorporate two new amino acids, O-methyl-tyrosine and p-aminophenylalanine. More recently, they (Chin et al.) have engineered the yeast strain Saccharomyces cerevisiae to incorporate five new unnatural amino acid (using the TAG nonsense codon and new, modified tRNA and tRNA synthetases) with keto groups that allow chemical modifications to the protein. We will concentrate only on the 20 abundant, naturally-occurring amino acids.

- Structure and Property of the Naturally-Occurring Amino Acids (Too large to include in text: print separately)
- EXTERNAL Learning Amino Acids Structure: YouTube - Part 1 | Part 2

Amino acids form polymers through a nucleophilic attack by the amino group of an amino acid at the electrophilic carbonyl carbon of the carboxyl group of another amino acid. The carboxyl group of the amino acid must first be activated to provide a better leaving group than OH-. (We will discuss this activation by ATP latter in the course.) The resulting link between the amino acids is an amide link which biochemists call a peptide bond. In this reaction, water is released. In a reverse reaction, the peptide bond can be cleaved by water (hydrolysis).

When two amino acids link together to form an amide link, the resulting structure is called a dipeptide. Likewise, we can have tripeptides, tetrapeptides, and other polypeptides. At some point, when the structure is long enough, it is called a protein. The average molecular weight of proteins in yeast is about 50,000 with about 450 amino acids. The large protein might be titin with molecular weight of about 3 million (about 27,0000 amino acids). A new class of very small proteins (30 or fewer amino acids and perhaps better named as polypeptides) called smORFs (small open reading frames) have recently been discovered to have significant biological activity (Science, doi:10.1126/science.1238802, 2013). These are encoded directly in the genome and are produced by the same processes that produce regular proteins (DNA transcription and RNA translation). They are not simply the result of selective cleavage of a larger protein into smaller peptide fragments.

There are many different ways to represent the structure of a polypeptide or protein. each showing differing amounts of information.

Figure: Different Representations of a Polypeptide (heptapeptide)

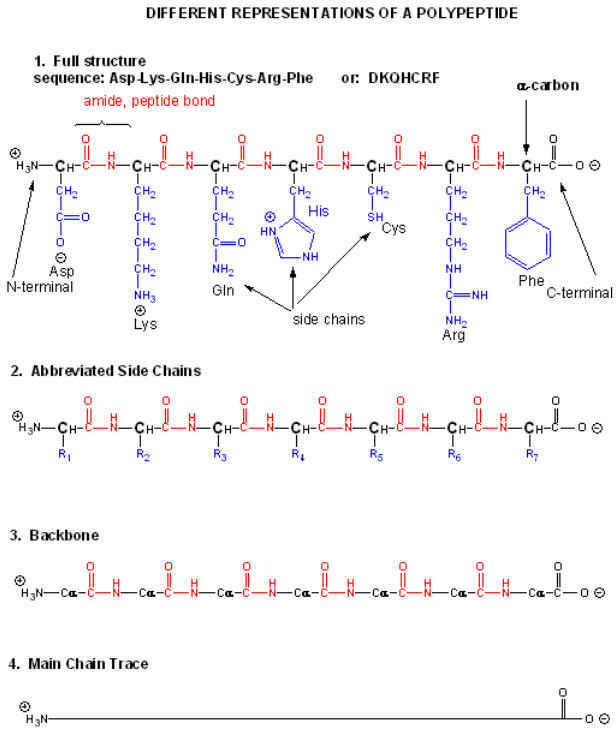
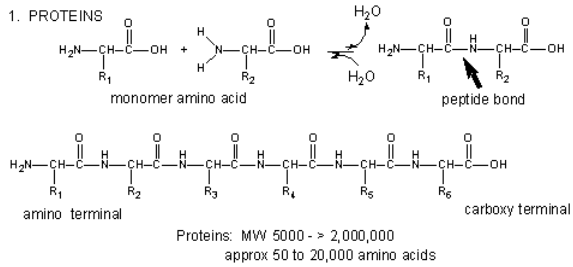


Figure: Amino Acids React to Form Proteins



(Note: above picture represents the amino acid in an unlikely protonation state with the weak acid protonated and the weak base deprotonated for simplicity in showing removal of water on peptide bond formation and the hydrolysis reaction.) Proteins are polymers of twenty naturally occurring amino acids. In contrast, nucleic acids are polymers of just 4 different monomeric nucleotides. Both the sequence of a protein and it's total length differentiate one protein from another. Just for an octapeptide, there are over 25 billion different possible arrangement of amino acids (820). Compare this to just 65536 different oligonucleotides (4 different monomeric deoxynucleotides) of 8 monomeric units, an 8mer (84). Hence the diversity of possible proteins is enormous.

Please consult the Jmol site below dealing with amino acids. Please learn the 3 letter code for the amino acids.

EXTERNAL Jsmol: Amino Acids from Charles S. Gasser, UC Davis Jmol: Amino Acids

Contributors and Attributions

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

This page titled 3.5: From Amino Acids to Peptides is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

- **A1. Amino Acid Structure** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.
- **4.1: From Amino Acids to Peptides** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.

CHAPTER OVERVIEW

4: Proteins- Structure and Folding

4.1: The Structure of Proteins- An Overview
4.2: Levels of Protein Structure
4.3: Secondary Structure
4.4: Protein folding
4.5: Analyses of Protein Structure
4.6: Protein function, domains and cooperativity
4.7: The Three-Dimensional Structure of Proteins
4.7.1: Main Chain Conformations
4.7.2: Secondary Structure and Loops
4.7.3: Tertiary, Quaternary, and Symmetrical Structures
4.7.4: Secondary Structural Motifs and Domains
4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure
4.7.6: Intrinsically Disordered Proteins
4.7.7: Fibrillar Proteins
4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics
4.7.9: Protein Stability - Thermodynamics
4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules
4.7.11: Biomolecular Visualization - Conceptions and Misconceptions
4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation
4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools
4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24)
4.7.15: Chapter 4 Questions
4.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs

4: Proteins- Structure and Folding is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.1: The Structure of Proteins- An Overview

Learning Goals (ChatGPT o1, 1/25/25)

- **Understand the Mechanism of Peptide Bond Formation:**
  - Explain the process of dehydration synthesis that links amino acids, including the role of water loss in forming the amide (peptide) bond.
  - Describe the chemical activation of the carboxylic acid group (via ATP and formation of a high-energy intermediate) that facilitates nucleophilic attack by the amino group.
  - Recognize common misconceptions about protonation states during peptide bond formation and clarify the proper chemical representation.
- **Comprehend the Role of the Ribosome and Genetic Information:**
  - Describe how the ribosome, as a ribonucleoprotein nanoparticle, catalyzes peptide bond formation during translation.
  - Connect the sequence information encoded in mRNA (and ultimately genomic DNA) to the primary sequence of a protein.
  - Understand the directionality of protein synthesis (N-terminus to C-terminus) and its implications for protein structure.
- **Define Protein Primary Structure:**
  - Identify the primary structure of a protein as its unique linear sequence of amino acids.
  - Explain how the order of amino acids determines the eventual folding and function of the protein.
  - Differentiate between peptides and proteins based on size, synthesis, and structural criteria.
- **Explore the Diversity of Protein Sequences and Structures:**
  - Quantify the vast number of potential protein sequences, even for short peptides, and appreciate the diversity generated by 20 amino acids.
  - Discuss the hierarchical organization of protein structure (primary, secondary, tertiary, and quaternary) and the role of primary structure as the foundation for higher-order folding.
  - Use examples (e.g., hydroxynitrile lyase) to illustrate how distinct regions of the primary sequence can adopt specific secondary structures (α-helices, β-sheets, coils) and contribute to overall tertiary and quaternary arrangements.
- **Connect Chemical Structure to Protein Function:**
  - Analyze how the primary sequence and covalent peptide bonds establish the framework for protein folding and function.
  - Discuss how post-translational modifications and disulfide bond formation (both intra- and inter-chain) further stabilize protein structure and mediate functional interactions.
  - Relate the principles of peptide bond formation and sequence variability to the extraordinary range of protein functions in biological systems.

These learning goals will guide students in integrating chemical, structural, and genetic principles that underlie peptide synthesis and protein structure, setting the stage for deeper exploration of protein folding, function, and regulation in later chapters.

4.1.1: Peptide Bond Formation and Primary Protein Structure

Proteins are polymers of amino acids that fold into shapes that confer function on the proteins. In biological systems, the amino acids are linked by a large ribonucleic acid/protein nanoparticle called the ***ribosome***. Thus, as amino acids are linked to form a specific protein, they are arranged in an order dictated by the genetic information encoded in messenger RNA (mRNA). The mRNA sequences are encoded in the genomic DNA sequence. The specific ordering of amino acids is the protein's ***primary sequence***. The translation mechanism used by the ribosome to synthesize proteins will be discussed in detail in Chapter 26.

The amino acids are linked together using ***dehydration synthesis*** (loss of water) reaction that connects the carboxylic acid of the upstream amino acid with the amine functional group of the downstream amino acid to form an amide linkage (Figure 2.10). You will remember from other chemistry courses that forming an amide from a carboxylic acid (thermodynamically stable) and an amine requires activating the carboxylic acid end to form a derivative with a better leaving group. This carbonyl of the modified end serves as an electrophile in the attack of the amine nitrogen, a nucleophile, in a nucleophilic substitution reaction. The activation reaction, which we will discuss in subsequent chapters, involves the transfer of a phosphate from a phosphoanhydride, ATP, to the carboxylic acid group, forming a mixed anhydride with the phosphate serving as a leaving group. Note that the reverse reaction, ***hydrolysis,*** requires the incorporation of a water molecule to separate two amino acids and break the amide bond. Notably, the ***ribosome*** serves as the enzyme that catalyzes the dehydration synthesis reactions required to build protein molecules, whereas a class of enzymes called ***proteases*** catalyzes protein hydrolysis.

A **peptide bond** is the amide linkage between amino acids in a protein. Subsequent amino acids will be added to the carboxylic acid terminal of the growing structure. Proteins are always synthesized directionally, starting with the amino terminus and ending with the carboxy terminus. New amino acids are always added onto the carboxylic acid tail, never onto the amine of the first amino acid in the chain. The ribosome dictates the directionality of protein synthesis. Figure **4.1.1** below shows an overly simplistic version of the reaction that produces the amide bond.

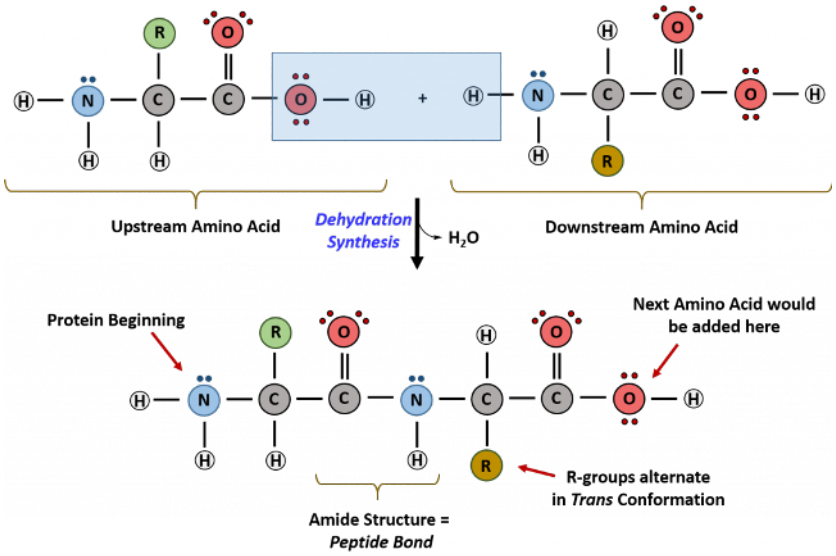


Figure **4.1.1**: Formation of the Peptide Bond. The addition of two amino acids to form a peptide requires dehydration synthesis. (Copyright; author via source)

Please note two features of the reaction as shown in the diagram:

1. The activation step (phosphorylation of the carboxylic acid end of the amino acid by ATP) is not shown.
2. The reaction is shown with an unlikely protonation state. If the carboxyl group is protonated (which occurs at low pH), the amine would also be protonated and should be correctly shown as RNH_3^+. The protonation state in the above figure was chosen to emphasize the loss of H_2O (dehydration) in the reaction. Many textbooks that lack a rigorous foundation in chemistry show unlikely protonation states for this reaction. By discussing this now, we hope to highlight common mistakes and misconceptions found in many resources.

Figure **4.1.2** shows a generic structure from a longer peptide or protein.

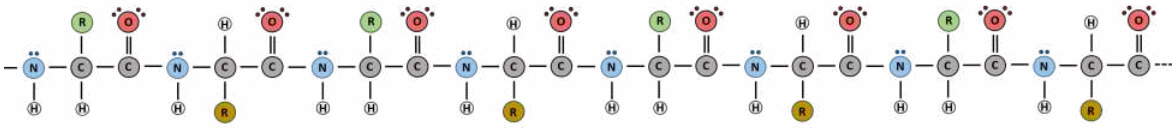


Figure **4.1.2**: Peptides and Proteins are macromolecules built from long chains of amino acids joined together through amide linkages. The order and nature of amino acids in the primary sequence of a protein determine the folding pattern of the protein based on the surrounding environment of the protein.

Proteins range in size from around 50 amino acids in length to rare proteins with molecular weights in the millions. One example is the human protein titin (also called connectin), a muscle protein. The human version has over 34,000 amino acids and a molecular weight of over 3.9 million! A polyketide synthase (PKS or a PKZILLAs) has 45,212 amino acids with a molecular weight of 4.7 million. "Giant" proteins, up to 85,804 amino acids, are encoded by some bacteria in the phylum Omnitrophota found in wastewater and hot pools. However, only pieces of the protein are found.

Some consider structures with fewer than 50 amino acids as **peptides** (Figure 2.13). Others suggest that 40-50 amino acid structures should be considered **small proteins**. One way to differentiate them is by how they are synthesized in vivo. Storz et al consider polypeptides to be small proteins if they are encoded in the genome by a continuous stretch of DNA base in an "open reading frame" (doi: 10.1146/annurev-biochem-070611-102400). Peptides, on the other hand, could be structures that are:

- "intrinsically disordered" with no definite fold,
- derived from proteins by proteolysis and/or
- not synthesized by ribosomes

As genomes were sequenced and annotated, some arbitrary parameters were set. In the yeast genome, annotated proteins were defined as those encoded by an open reading frame (ORF) in the DNA sequence, leading to a polypeptide of 100 amino acids (which, on average, has a molecular weight of 11,000). If no cutoff were used, the number of proteins encoded by the genome would be huge. Submissions of DNA sequences to the NIH GenBank must encode proteins of at least about 66 amino acids (MW about 7250). Even this ignores small proteins isolated and characterized from cells. So the cutoff of 50 amino acids (MW about 5500) derived from open reading frames appears to be the best arbitrary cutoff in the transition from peptides to proteins.

The definition of a protein as encoded in an "open reading frame" in the DNA/RNA is contingent on how you define an open reading frame. Evidence suggests that many proteins are made from atypical or noncanonical open reading frames (ncORFs). This study found that at least 25% of 7,264 ncORFs were transcribed and translated into over 3,000 peptides/proteins. Most DNA/RNA sequences were less than 100 codons (or the equivalent number of amino acids). Most were also within long noncoding RNAs (see RNA Chapter) or untranslated regions of messenger RNAs. Their functional significance is, in most cases, uncertain, yet their presence would expand the number of protein-encoding genes in the human genome. These sequences are available in two databases: GENCODE, which identifies and classifies gene features in human and mouse genomes based on biological evidence, and PeptideAtlas, which houses peptide structures determined by mass spectrometry. It's difficult to differentiate peptides transcribed from ncORFs from the 1000s of peptides produced in cells from proteolysis and other means. The new sequences, found in what has been called the "dark proteome," add 1000s of proteins to the previous total of around 20,000 human protein-coding genes.

Due to the large pool of amino acids that can be incorporated at each position within the protein, billions of different possible protein combinations can be used to create novel protein structures! For example, think about a tripeptide made from this amino acid pool. At each position, 20 different options can be incorporated. Thus, the total number of resulting tripeptides possible would be 20 X 20 X 20 or 20^3, or 8,000 different tripeptide sequences! Now, think about how many options there would be for a small peptide containing 40 amino acids. There would be 20^{40} options or a mind-boggling 1.09×10^{52} potential sequence options! Each of these options would vary in overall protein shape, as the nature of amino acid side chains helps determine the protein's interactions with other residues and its surrounding environment.

Nearly 200,000 3D structures of biomacromolecules are known, and over a million have been determined using artificial intelligence computer programs. How can we simplify our understanding of the diversity of protein structures? Is each structure unique? What do they have in common?

To simplify and clarify our understanding of biological diversity, we group organisms (from domains and kingdoms to genera and species) based on common characteristics. Likewise, proteins are divided into a hierarchy of structures with increasing information content. This overview describes the four basic levels of protein structure: primary (10), secondary (20), tertiary (30), and quaternary (40). Each will be probed in greater detail in the next chapter. These structural classes will be illustrated below using the protein **hydroxynitrile lyase** (5Y02). (This protein has been simplified to demonstrate key structural features, which will be described at the end.)

4.1.2: Primary (1°) Structure: the amino acid sequence of a protein.

A protein's primary (10) structure is simply the amino acid sequence of a protein written from N- to C-terminal. It does not require visualization to describe it. Consider two different short, continuous sequences from the hydroxynitrile lyase:

- Gln-Lys-Gln-Ile-Asp-Gln-Ile or in single letter code QKQIDQI. This is the sequence for amino acids 20-26 in the protein. This stretch of the 1° structure has multiple repeated amino acids.
- Asp-Leu-Gly-Pro-Ala-Val or in single letter code DLGPAV. This is the sequence for amino acids 48-53 in the protein. This stretch of the 1° structure does not contain repetitive amino acids.

A 2-D line drawing of the sequence offers more information but does not provide information about the conformation of these sections of 1° structures within a given protein. These can be shown in Figure **4.1.3**, in which the protein's overall structure is shown in grey sticks with short stretches of primary structure shown in colored spacefill and 2D line drawings.

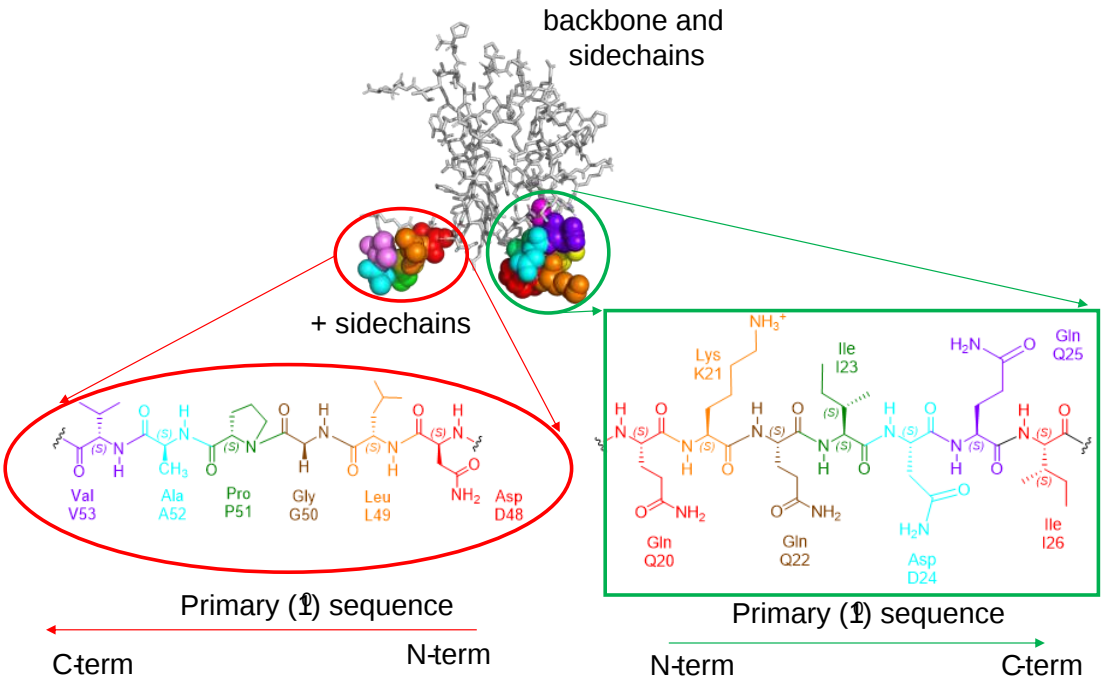


Figure **4.1.3**: Alternative renderings of a "primary" sequence within a protein

4.1.3: Secondary (2^0) and Tertiary (3^0) Structures

Secondary (2^0) structures are repetitive structures within a protein held together by hydrogen bonds between amide Hs and carbonyl Os in the backbone main chain atoms. It's most easily examined through the specific rendering of the overall **tertiary (3^0)** or 3-D protein structure. Five different renderings showing the 3D (the **tertiary**) structure of the protein are shown in Figure **4.1.4**.

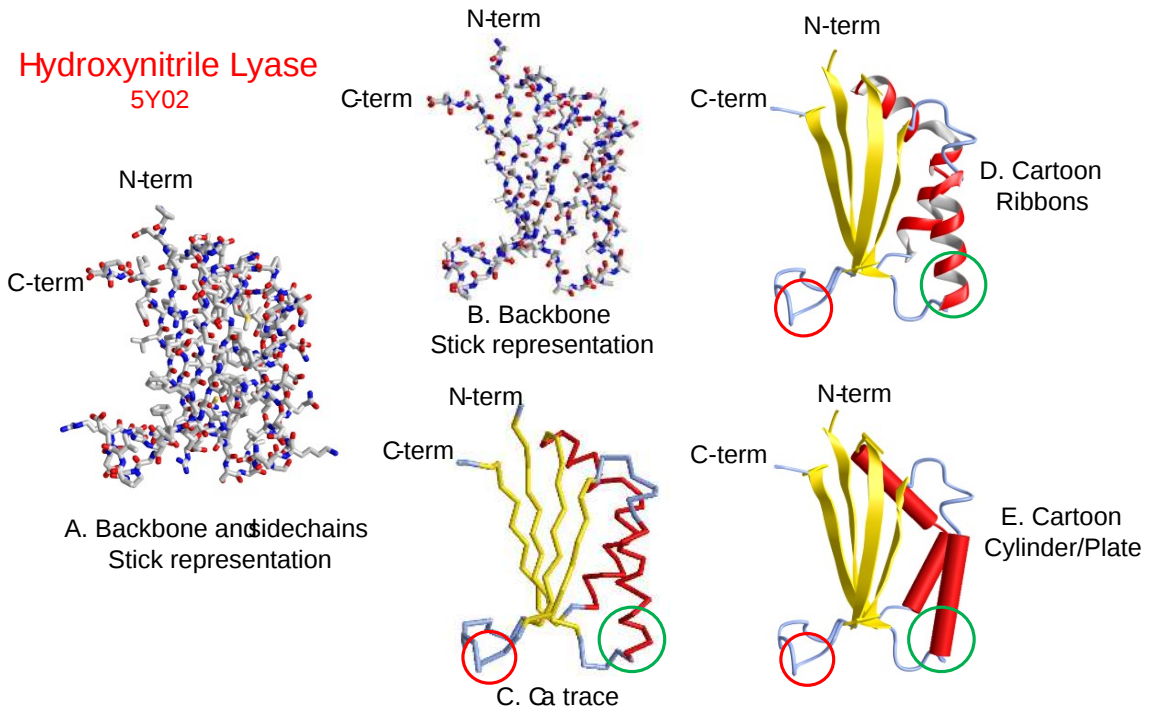


Figure **4.1.4**: Alternative renderings of a protein

Representation **A** shows a stick drawing of the protein, with red indicating bonds to oxygen and blue bonds to nitrogen. No covalent bonds to hydrogen are shown, as hydrogen atoms are too small to be detected by common techniques used to determine the structures of such large molecules. It looks like a complicated tangle of bonds, making it difficult to understand the unique features of the protein's 3D structure. Representation **B** shows only the backbone of the protein—the outline of how it twists and turns in space becomes more evident. The N- and C-terminal ends are more clearly seen.

Representation **C** shows just the bonds connecting the alpha C atoms of each amino acid. The protein's overall topology is now clearly evident. If you follow the chain from the N- to the C-terminal ends, it should be evident that there are regularities in the conformations of the protein chain. The individual yellow zig-zags are called **beta** strands. These strands appear elongated and aligned with other beta strands, forming a larger **beta sheet**. The sheet is held together by hydrogen bonds between backbone amide Hs and carbonyl O atoms on adjacent strands. Beta strands are a type of **secondary structure**.

The red zig-zag lines represent another type of secondary structure called the **alpha helix**. Hydrogen bonds hold the helix together between amide Hs and carbonyl Os within a single continuous strand. The backbone of the alpha helix appears less elongated than in a beta-strand, as it is wound into a coil (the alpha helix) along a central axis. If you took tweezers (using atomic force microscopy) and pulled on the helix, it could stretch and become more elongated, like the beta strands.

The remaining protein alpha carbon chain, shown in blue, is less regular. However, it is still ordered as it propagates through space in a **random coil**. It mainly adopts a fixed conformation but is more flexible than alpha helices and beta sheets. The alpha helices and beta strands (sheets) are examples of secondary structures.

Representations **D** and **E** are cartoon drawings clearly showing alpha helices (red), beta strands, and sheets (yellow). It would be challenging to discern alpha or beta secondary structures from stick representations that show all the bonds in a protein. Some atoms must be removed visually (not literally) to see the protein backbone's repetitive propagation through the overall structure. A cartoon view alone would not be helpful if you aimed to understand the disposition of side chains in a small part of a protein. Modeling programs allow mixed rendering of a protein, including cartoon and stick representations.

Secondary structures, held together by hydrogen bonds between backbone atoms, are characterized by repetitive changes in the chain propagation angle between connected amino acids in an alpha helix and a beta-strand. In a given beta-strand, the relative change in the propagation angle is nearly 0^0 compared to a much larger angular change required to bend the amino acid backbone

into an alpha helix. Here is the IUPAC definition of secondary structure. We added the word "repetitive" to indicate that random coils are not examples of secondary structures.

Definition: Secondary Structure (from the IUPAC Gold Book)

The [repetitive] conformational arrangement (α-helix, β-pleated sheet, etc.) of the backbone segments of a macromolecule, such as a polypeptide chain of a protein, without regard to the conformation of the side chains or the relationship to other segments.

4.1.4: Quaternary Structure

Separate protein chains often interact through noncovalent interactions and, sometimes, through disulfide bond formation between free cysteine side chains on different chains, forming dimers, trimers, tetramers, octamers, etc. Dimers can be homodimers (if the two chains are identical) or heterodimers (if they are different). The example in the section, hydroxynitrile lyase, forms a homodimer, as shown in Figure 4.1.5. The left image shows a cartoon version, with one monomer in orange and the other identical monomer in green. The right image shows a translucent surface representation of the dimer, with the cartoon image underneath.

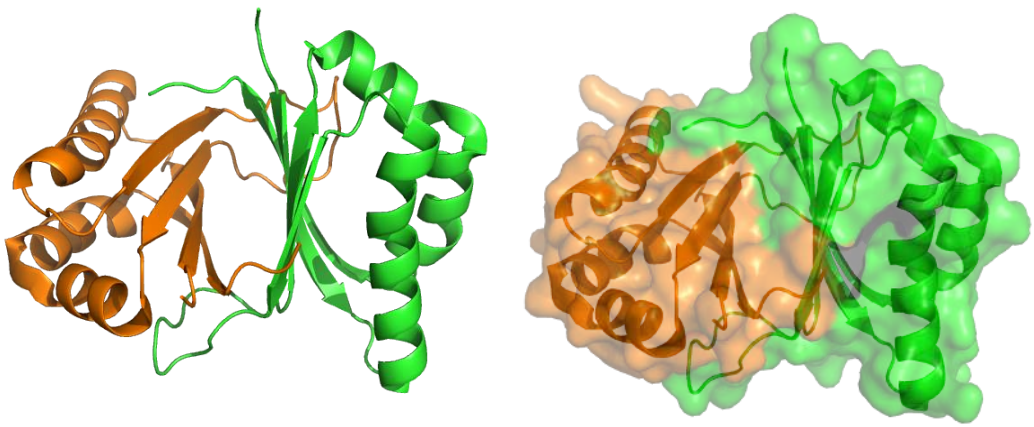


Figure 4.1.5: Dimer of hydroxynitrile lyase

The mixed-rendered image on the right shows a translucent surface image of each monomer, and underneath the cartoon image

- primary structure: the linear amino acid sequence of a protein
- secondary structure: regular repeating structures arising when hydrogen bonds between the peptide backbone amide hydrogens and carbonyl oxygens occur at regular intervals within a given linear sequence (strand) of a protein or between two adjacent strands

Disulfide bonds within individual chains and between them stabilize both tertiary and quaternary structures of both peptides and proteins. These are illustrated in Figure 4.1.6.

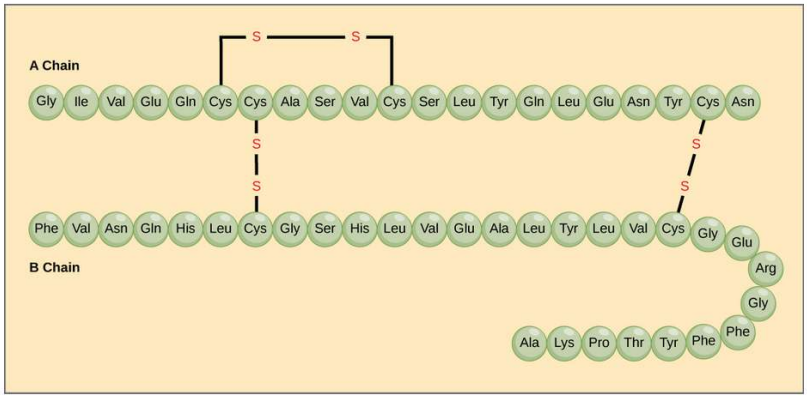


Figure **4.1.6**: **Disulfide Bonds.** Disulfide bonds are formed between two cysteine residues within a peptide or protein sequence or between different peptide or protein chains. The two peptide chains that form the hormone insulin are depicted in the example above. Disulfide bridges between the two chains are required for this hormone to regulate blood glucose levels properly. Image by: CNX OpenStax via Wikimedia Commons

4.1.5: Summary

Chapter Summary

This chapter delves into the foundation of protein structure by examining peptide bond formation and the concept of primary structure. It begins by outlining how proteins, as polymers of amino acids, are synthesized by the ribosome. This complex ribonucleoprotein machine translates the genetic code carried by mRNA into a specific linear sequence of amino acids. This sequence, known as the protein's primary structure, is determined by the genomic DNA and dictates the protein's ultimate three-dimensional structure and function.

Key topics include:

- **Peptide Bond Formation:**
  The chapter explains that amino acids are linked via dehydration synthesis—a process that forms an amide (peptide) bond by losing a water molecule. It discusses the necessity of activating the carboxylic acid group (typically via phosphorylation with ATP) to create a better leaving group, thereby enabling nucleophilic attack by the amine group of the incoming amino acid. The role of the ribosome in catalyzing this reaction and common misconceptions regarding protonation states during the reaction is clarified.
- **Primary Protein Structure:**
  Emphasis is placed on the importance of the linear sequence of amino acids. The chapter details how the unique ordering of amino acids in a protein—ranging from short peptides to giant proteins like titin—forms the basis for its folding into higher-order structures (secondary, tertiary, and quaternary). This primary structure not only defines the protein's chemical properties but also underlies its functional diversity in biological systems.
- **Diversity and Definition of Proteins vs. Peptides:**
  The discussion highlights the enormous diversity possible from 20 amino acids and how even a short peptide can have millions of potential sequences. It also addresses the criteria used to distinguish between peptides and proteins, including size and the synthesis mechanism (continuous open reading frames versus proteolytic products).
- **Hierarchy of Protein Structure:**
  Finally, the chapter introduces the concept of hierarchical protein organization—starting with the primary structure, moving through regular secondary structures (such as α-helices and β-sheets stabilized by hydrogen bonds), and culminating in complex tertiary and quaternary structures that often involve additional stabilizing elements like disulfide bonds.

Overall, this chapter lays the groundwork for understanding how the precise amino acid sequence is essential for protein folding and function, setting the stage for more detailed studies of protein structure and dynamics in subsequent chapters.

4.1.6: References

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

This page titled 4.1: The Structure of Proteins- An Overview is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

4.2: Levels of Protein Structure

Source: BiochemFFA_2_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Proteins are the workhorses of the cell. Virtually everything that goes on inside of cells happens as a result of the actions of proteins. Among other things, protein enzymes catalyze the vast majority of cellular reactions, mediate signaling, give structure both to cells and to multicellular organisms, and exert control over the expression of genes. Life, as we know it, would not exist if there were no proteins. The versatility of proteins arises because of their varied structures.

Proteins are made by linking together amino acids, with each protein having a characteristic and unique amino acid sequence. To get a sense for the diversity of proteins that can be made using 20 different amino acids, consider that the number of different combinations possible with 20 amino acids is 20^n, where n=the number of amino acids in the chain. It becomes apparent that even a dipeptide made of just two amino acids joined together gives us 20^2 = 400 different combinations. If we do the calculation for a short peptide of 10 amino acids, we arrive at an enormous 10,240,000,000,000 combinations. Most proteins are much larger than this, making the possible number of proteins with unique amino acid sequences unimaginably huge.

Levels of Structure

The significance of the unique sequence, or order, of amino acids, known as the protein's primary structure, is that it dictates the 3-D conformation the folded protein will have. This conformation, in turn, will determine the function of the protein. We shall examine protein structure at four distinct levels (Figure 2.17) - 1) how sequence of the amino acids in a protein (primary structure) gives identity and characteristics to a protein (Figure 2.18); 2) how local interactions between one part of the polypeptide backbone and another affect protein shape (secondary structure); 3) how the polypeptide chain of a protein can fold to allow amino acids to interact with each other that are not close in primary structure (tertiary structure); and 4) how different polypeptide chains interact with each other within a multi-subunit protein (quaternary structure).

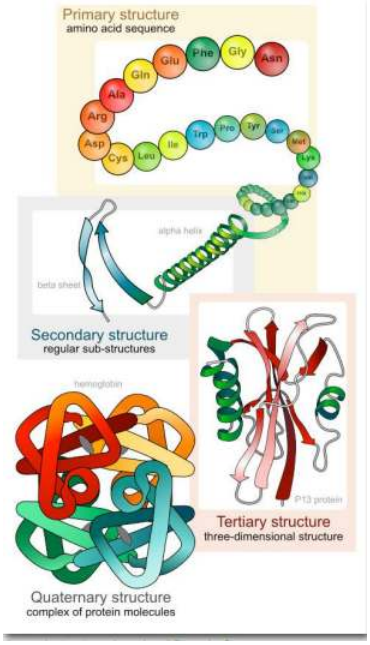


Figure 2.17 - Four Levels of Protein Structure

At this point, we should provide a couple of definitions. We use the term polypeptide to refer to a single polymer of amino acids. It may or may not have folded into its final, functional form. The term protein is sometimes used interchangeably with polypeptide, as in "protein synthesis". It is generally used, however, to refer to a folded, functional molecule that may have one or more subunits (made up of individual polypeptides). Thus, when we use the term protein, we are usually referring to a functional, folded polypeptide or peptides. Structure is essential for function. If you alter the structure, you alter the function - usually, but not always, this means you lose all function. For many proteins, it is not difficult to alter the structure.

Proteins are flexible, not rigidly fixed in structure. As we shall see, it is the flexibility of proteins that allows them to be amazing catalysts and allows them to adapt to, respond to, and pass on signals upon binding of other molecules or proteins. However, proteins are not infinitely flexible. There are constraints on the conformations that proteins can adopt and these constraints govern the conformations that proteins display.

Subtle changes

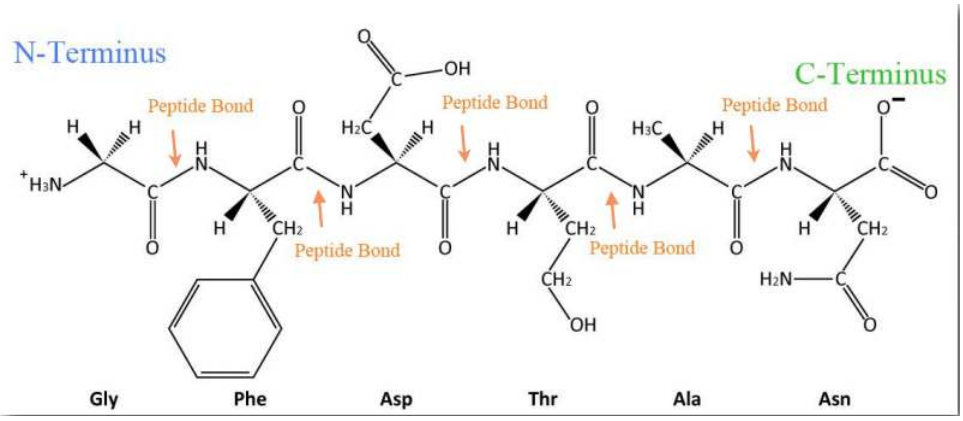


Figure 2.18 - Sequence of a simple polypeptide Wikipedia

Even very tiny, subtle changes in protein structure can give rise to big changes in the behavior of proteins. Hemoglobin, for example, undergoes an incredibly small structural change upon binding of one oxygen molecule, and that simple change causes the remainder of the protein to gain a considerably greater affinity for oxygen that the protein didn't have before the structural change.

Sequence, structure and function

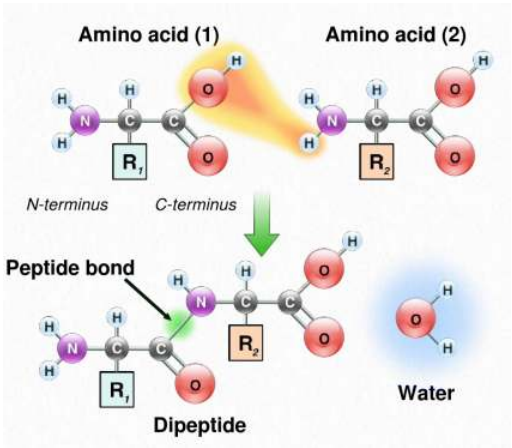


Figure 2.19 Linking of amino acids through peptide bond formation

As discussed earlier, the number of different amino acid sequences possible, even for short peptides, is very large. No two proteins with different amino acid sequences (primary structure) have identical overall structure. The unique amino acid sequence of a protein is reflected in its unique folded structure. This structure, in turn, determines the protein's function. This is why mutations that alter amino acid sequence can affect the function of a protein.

Protein Synthesis

Synthesis of proteins occurs in the ribosomes and proceeds by joining the carboxyl terminus of the first amino acid to the amino terminus of the next one (Figure 2.19). The end of the protein that has the free α-amino group is referred to as the amino terminus or N-terminus. The other end is called the carboxyl terminus or C-terminus , since it contains the only free α-carboxyl group. All of the other α-amino groups and α-carboxyl groups are tied up in forming peptide Figure 2.19 Linking of amino acids through peptide bond formation bonds that join adjacent amino acids together. Proteins are synthesized starting with the amino terminus and ending at the carboxyl terminus.

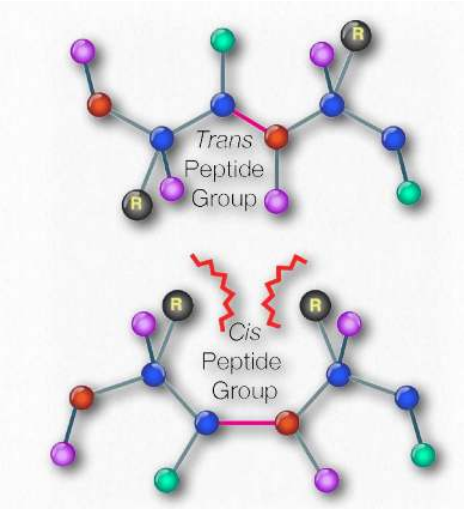


Figure 2.20 - Cis vs trans orientation of R-groups around peptide bond Image by Aleia Kim

Schematically, in Figure 2.18, we can see how sequential R-groups of a protein are arranged in an alternating orientation on either side of the polypeptide chain. Organization of R-groups in this fashion is not random. Steric hindrance can occur when consecutive R-groups are oriented on the same side of a peptide backbone (Figure 2.20)

Primary Structure

Primary structure is the ultimate determinant of the overall conformation of a protein. The primary structure of any protein arrived at its current state as a result of mutation and selection over evolutionary time. Primary structure of proteins is mandated by the sequence of DNA coding for it in the genome. Regions of DNA specifying proteins are known as coding regions (or genes).

The base sequences of these regions directly specify the sequence of amino acids in proteins, with a one-to-one correspondence between the codons (groups of three consecutive bases) in the DNA and the amino acids in the encoded protein. The sequence of codons in DNA, copied into messenger RNA, specifies a sequence of amino acids in a protein. (Figure 2.21).

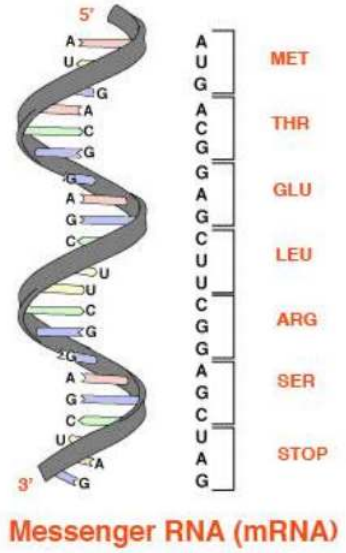


Figure 2.21 - From RNA to amino acids - the genetic code Wikipedia

The order in which the amino acids are joined together in protein synthesis starts defining a set of interactions between amino acids even as the synthesis is occurring. That is, a polypeptide can fold even as it is being made. The order of the R-group structures and resulting interactions are very important because early interactions affect later interactions. This is because interactions start establishing structures - secondary and tertiary. If a helical structure (secondary structure), for example, starts to form, the

possibilities for interaction of a particular amino acid Rgroup may be different than if the helix had not formed (Figure 2.22). R-group interactions can also cause bends in a polypeptide sequence (tertiary structure) and these bends can create (in some cases) opportunities for interactions that wouldn't have been possible without the bend or prevent (in other cases) similar interaction possibilities.

Secondary Structure

As protein synthesis progresses, interactions between amino acids close to each other begin to occur, giving rise to local patterns called secondary structure. These secondary structures include the well known α- helix and β-strands. Both were predicted by Linus Pauling, Robert Corey, and Herman Branson in 1951. Each structure has unique features.

α-helix

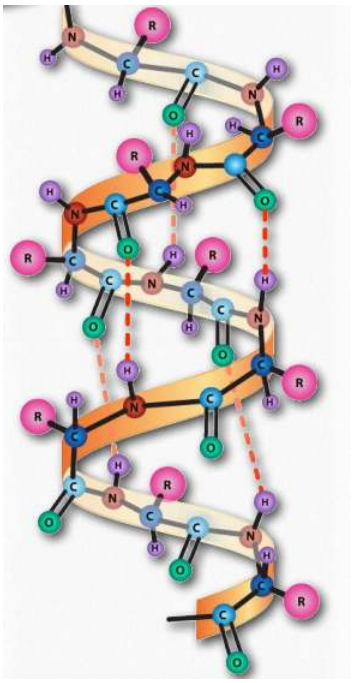


Figure 2.22 - The α-helix. Hydrogen bonds (dotted lines) between the carbonyl oxygen and the amine hydrogen stabilize the structure. Image by Aleia Kim

The α-helix has a coiled structure, with 3.6 amino acids per turn of the helix (5 helical turns = 18 amino acids). Helices are predominantly right handed - only in rare cases, such as in sequences with many glycines can left handed α- helices form. In the α-helix, hydrogen bonds form between C=O groups and N-H groups in the polypeptide backbone that are four amino acids distant. These hydrogen bonds are the primary forces stabilizing the α-helix.

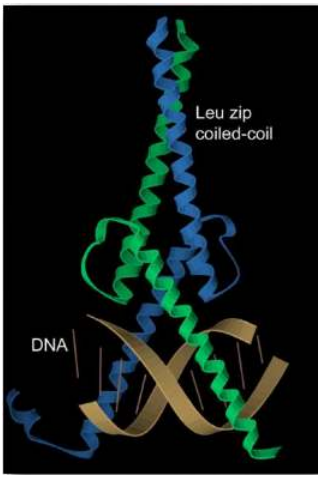


Figure 2.23 - α-helices in a protein with a leucine zipper structural domain. The α-helices are shown in blue and green and are bound to a DNA double helix in brown.

We use the terms rise, repeat, and pitch to describe the parameters of any helix. The repeat is the number of residues in a helix before it begins to repeat itself. For an α-helix, the repeat is 3.6 amino acids per turn of the helix. The rise is the distance the helix elevates with addition of each residue. For an α-helix, this is 0.15 nm per amino acid. The pitch is the distance between complete turns of the helix. For an α-helix, this is 0.54 nm. The stability of an α-helix is enhanced by the presence of the amino acid aspartate.

Figure 2.24 - α-helix sculpture outside Linus Pauling's boyhood home Wikipedia

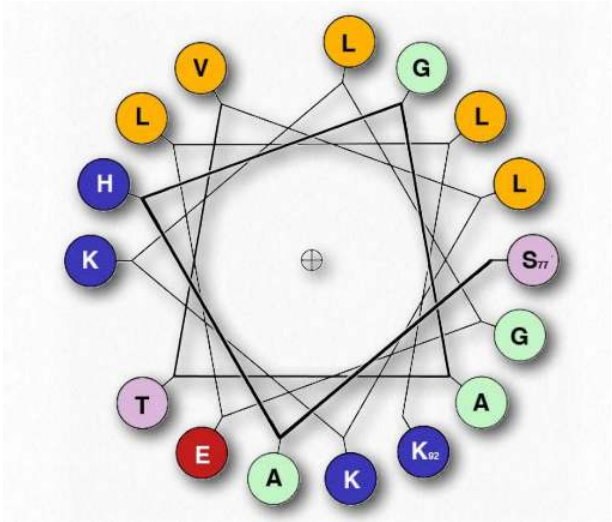


Figure 2.25 - Helical Wheel Representation of an α-Helix. The one letter genetic code is used. The helix starts at Serine #77 at the right and ends at lysine #92 in the lower right. Hydrophobic amino acids are shown in yellow and ionizing amino acids are shown in blue. Hydrophobic amino acids tend to interact with each other and not with ionizing amino acids. Wikipedia

β strand/sheet

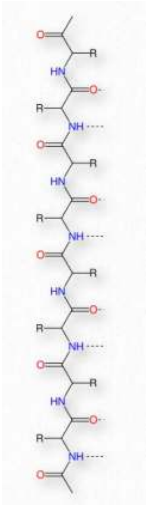


Figure 2.26 - β strand

A helix is, of course, a three-dimensional object. A flattened form of helix in two dimensions is a common description for a β-strand. Rather than coils, β-strands have bends and these are sometimes referred to as pleats, like the pleats in a curtain. β-strands can be organized to form elaborately organized structures, such as sheets, barrels, and other arrangements.

Higher order β-strand structures are sometimes called supersecondary structures), since they involve interactions between amino acids not close in primary sequence. These structures, too, are stabilized by hydrogen bonds between carbonyl oxygen atoms and hydrogens of amine groups in the polypeptide backbone (Figure 2.28). In a higher order structure, strands can be arranged parallel (amino to carboxyl orientations the same) or anti-parallel (amino to carboxyl orientations opposite of each other (in Figure 2.27, the direction of the strand is shown by the arrowhead in the ribbon diagrams).

Turns

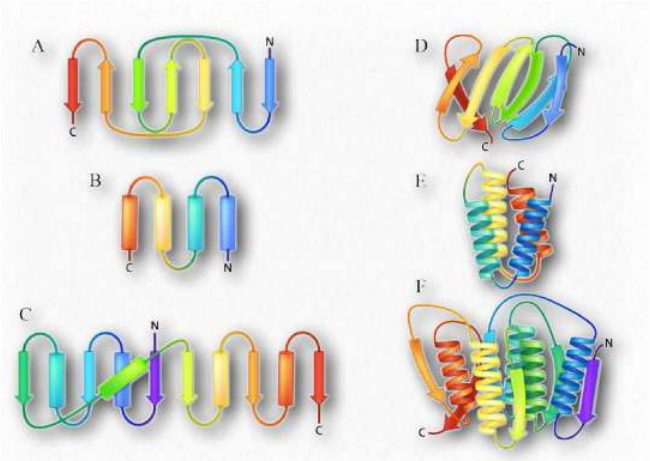


Figure 2.27 - Ribbon depictions of supersecondary β-sheets (A-D) and α-helix arrangements (E-F) Image by Aleia Kim

Turns (sometimes called reverse turns) are a type of secondary structure that, as the name suggests, causes a turn in the structure of a polypeptide chain. Turns give rise to tertiary structure ultimately, causing interruptions in the secondary structures (α- helices and β-strands) and often serve as connecting regions between two regions of secondary structure in a protein. Proline and glycine play common roles in turns, providing less flexibility (starting the turn) and greater flexibility (facilitating the turn), respectively.

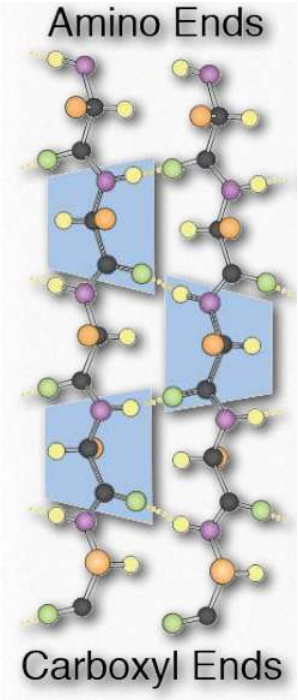


Figure 2.28 - Components of a β-sheet in a parallel arrangement. H-bonds in yellow. Image by Aleia Kim

There are at least five types of turns, with numerous variations of each giving rise to many different turns. The five types of turns are

• δ-turns - end amino acids are separated by one peptide bond

• γ-turns - separation by two peptide bonds

•β-turns - separation by three peptide bonds

•α-turns - separation by four peptide bonds

•π-turns - separation by five bonds

Of these, the β-turns are the most common form and the δ-turns are theoretical, but unlikely, due to steric limitations. Figure 2.29 depicts a β- turn.

3_{10} helices

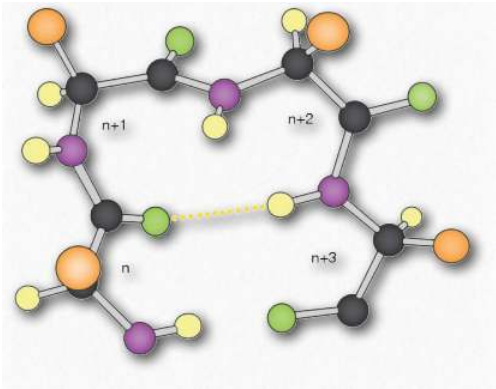


Figure 2.29 - β-turn. R-groups are shown in orange, hydrogens in yellow, carbons in charcoal, nitrogens in purple, and oxygens in green. A stabilizing hydrogen bond is indicated with the dotted line. Image by Aleia Kim

In addition to the α-helix, β-strands, and various turns, other regular, repeating structures are seen in proteins, but occur much less commonly. The 3_{10} helix is the fourth most abundant secondary structure in proteins, constituting about 10-15% of all helices. The helix derives its name from the fact that it contains 10 amino acids in 3 turns. It is right-handed. Hydrogen bonds form between amino acids that are three residues apart. Most commonly, the 3_{10} helix appears at the amine or carboxyl end of an α-helix. Like the α-helix, the 3_{10} helix is stabilized by the presence of aspartate in its sequence.

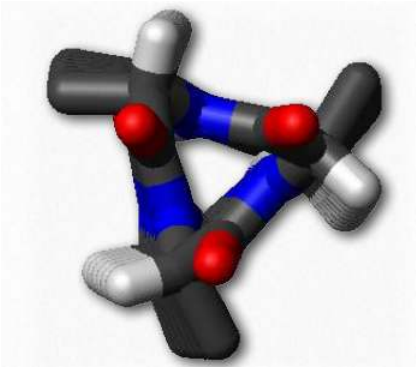


Figure 2.30 - Top view of a 310 Helix. Carbonyl groups are in red and pointed upwards. Note the almost perfect 3-fold symmetry Wikipedia

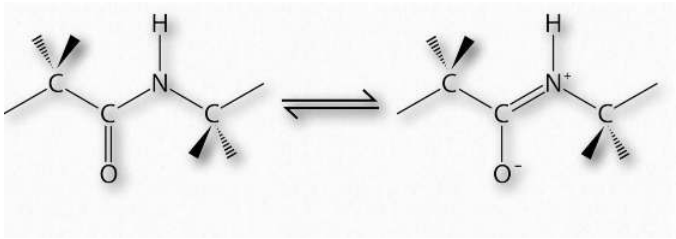


Figure 2.31 - Resonance of the peptide bond Wikipedia

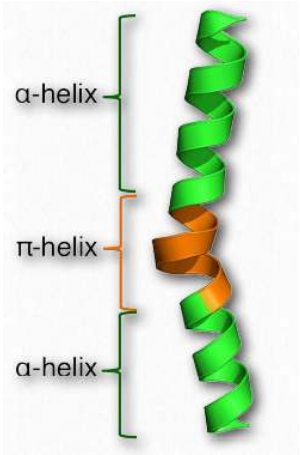


Figure 2.32 - π helix Wikipedia

π-helices

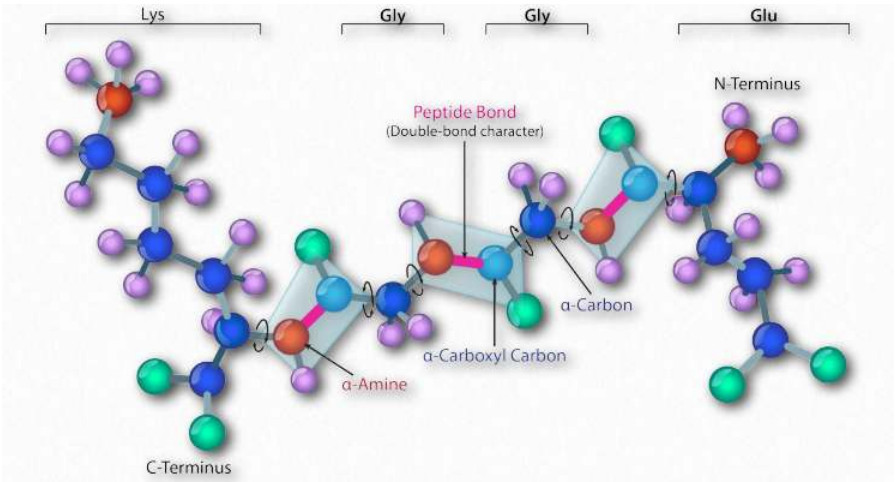


Figure 2.33 - Planes (light blue) defined by the double-bonded character of the peptide bond Image by Aleia Kim

A π-helix may be thought of as a special type of α- helix. Some sources describe it as an α-helix with an extra amino acid stuck in the middle of it (Figure 2.32). π-helices are not exactly rare, occurring at least once in as many as 15% of all proteins. Like the α-helix, the π-helix is right-handed, but where the α-helix has 18 amino acids in 5 turns, the π-helix has 22 amino acids in 5 turns. π-helices typically do not stretch for very long distances. Most are only about 7 amino acids long and the sequence almost always occurs in the middle of an α-helical region.

Ramachandran plots

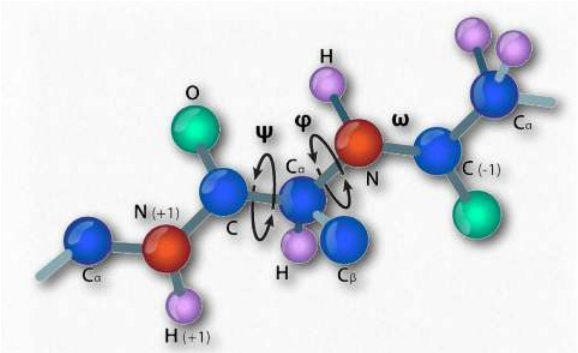


Figure 2.34 - ω, ψ, and φ rotational angles in a peptide Image by Aleia Kim

In 1963, G.N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan described a novel way to describe protein structure. If one considers the backbone of a polypeptide chain, it consists of a repeating set of three bonds. Sequentially (in the amino to carboxyl

direction) they are 1) a rotatable bond (ψ) between α-carbon and α-carboxyl preceding the peptide bond (see HERE), 2) a non-rotatable peptide bond (ω) between the α-carboxyl and α-amine groups), and 3) a rotatable bond (φ) between the α-amine and α-carbon following the peptide bond (see HERE). Note in Figures 2.33 and 2.34 that the amino to carboxyl direction is right to left.

The presence of the carbonyl oxygen on the α-carboxyl group allows the peptide bond to exist as a resonant structure, meaning that it behaves some of the time as a double bond. Double bonds cannot, of course, rotate, but the bonds on either side of it have some freedom of rotation. The φ and ψ angles are restricted to certain values, because some angles will result in steric hindrance. In addition, each type of secondary structure has a characteristic range of values for φ and ψ.

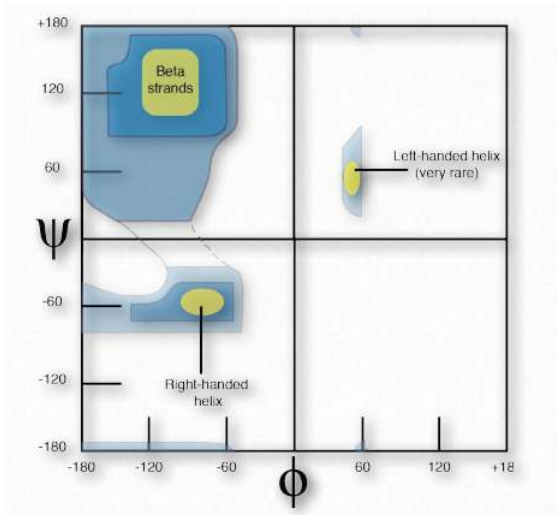


Figure 2.35 - Theoretical Ramachandran plot Image by Penelope Irving

Ramachandran and colleagues made theoretical calculations of the energetic stability of all possible angles from 0° to 360° for each of the φ and ψ angles and plotted the results on a Ramachandran Plot (also called a φ-ψ plot), delineating regions of angles that were theoretically the most stable (Figure 2.35).

Three primary regions of stability were identified, corresponding to φ-ψ angles of β-strands (top left), right handed α- helices (bottom left), and lefthanded α-helices (upper right). The plots of predicted stability are remarkably accurate when compared to φ-ψ angles of actual proteins.

Secondary structure prediction

Amino Acid	α helix	Reverse turn	β sheet
Ala	1.41	0.82	0.72
Arg	1.21	0.90	0.84
Asn	0.76	1.34	0.48
Asp	0.99	1.24	0.39
Cys	0.66	0.54	1.40
Gln	1.27	0.84	0.98
Glu	1.59	1.01	0.52
Gly	0.43	1.77	0.58
His	1.05	0.81	0.80
Ile	1.09	0.47	1.67
Leu	1.34	0.57	1.22
Lys	1.23	1.07	0.69
Met	1.30	0.52	1.14
Phe	1.16	0.59	1.33
Pro	0.34	1.32	0.31
Ser	0.57	1.22	0.96
Thr	0.76	0.96	1.17
Trp	1.02	0.65	1.35
Tyr	0.74	0.76	1.45
Val	0.90	0.41	1.87

Table 2.3 - Relative tendencies of each amino acid to be in a secondary structure. Higher values indicate greater tendency Image by Penelope Irving

By comparing primary structure (amino acid sequences) to known 3D protein structures, one can tally each time an amino acid is found in an α-helix, β-strand/sheet, or a turn. Computer analysis of thousands of these sequences allows one to assign a likelihood of any given amino acid appearing in each of these structures. Using these tendencies, one can, with up to 80% accuracy, predict regions of secondary structure in a protein based solely on amino acid sequence.

This is seen in Table 2.3. Occurrence in primary sequence of three consecutive amino acids with relative tendencies higher than one is an indicator that that region of the polypeptide is in the corresponding secondary structure. An online resource for predicting secondary structures called PSIPRED is available HERE.

Hydrophobicity

Amino Acid Hydropathy Scores

Amino Acid	One Letter Code	Hydropathy Score
Isoleucine	I	4.5
Valine	V	4.2
Leucine	L	3.8
Phenylalanine	F	2.8
Cysteine	C	2.5
Methionine	M	1.9
Alanine	A	1.8
Glycine	G	-0.4
Threonine	T	-0.7
Tryptophan	W	-0.9
Serine	S	-0.8
Tyrosine	Y	-1.3
Proline	P	-1.6
Histidine	H	-3.2
Glutamic acid	E	-3.5
Glutamine	Q	-3.5
Aspartic acid	D	-3.5
Asparagine	N	-3.5
Lysine	K	-3.9
Arginine	R	-4.5

Table 2.4 - Hydropathy Scores

The chemistry of amino acid Rgroups affects the structures they are most commonly found in. Subsets of their chemical properties can give clues to structure and, sometimes, cellular location. A prime example is the hydrophobicity (wateravoiding tendencies) of some Rgroups. Given the aqueous environment of the cell, such R-groups are not likely to be on the outside surface of a folded protein.

However, this rule does not hold for regions of protein that may be embedded within the lipid bilayers of cellular/ organelle membranes. This is because the region of such proteins that form the transmembrane domains are are buried in the hydrophobic environment in the middle of the lipid bilayer.

Not surprisingly, scanning primary sequences for specifically sized/spaced stretches of hydrophobic amino acids can help to identify proteins found in membranes. Table 2.4 shows hydrophobicity values for R-groups of the amino acids. In this set, the scale runs from positive values (hydrophobic) to negative values (hydrophilic). A KyteDoolittle Hydropathy plot for the RET protooncogene membrane protein is shown in Figure 2.36. Two regions of the protein are very hydrophobic as can be seen from the peaks near amino acids 5-10 and 630-640. Such regions might be reasonably expected to be situated either within the interior of the folded protein or to be part of transmembrane domains.

Random coils

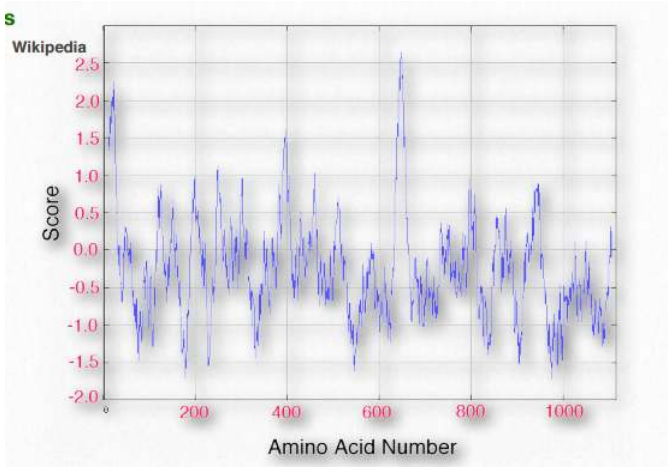


Figure 2.36 Kyte-Doolittle hydropathy plot for the RET protooncogene Wikipedia

Some sections of a protein assume no regular, discernible structure and are sometimes said to lack secondary structure, though they may have hydrogen bonds. Such segments are described as being in random coils and may have fluidity to their structure that results in them having multiple stable forms. Random coils are identifiable with spectroscopic methods, such as circular dichroism Wikipedia and nuclear magnetic resonance (NMR) in which distinctive signals are observed. See also metamorphic proteins (HERE) and intrinsically disordered proteins (HERE).

Figure 2.37 - Ribbon depiction of a β-hairpin. Shown are two β strands in turquoise interacting with each other.

Supersecondary structure

Another element of protein structure is harder to categorize because it incorporates elements of secondary and tertiary structure. Dubbed supersecondary structure (or structural motifs), these structures contain multiple nearby secondary structure components arranged in a specific way and that appear in multiple proteins. Since there are many ways of making secondary structures from different primary structures, so too can similar motifs arise from different primary sequences. An example of a structural motif is shown in Figure 2.37.

Tertiary structure

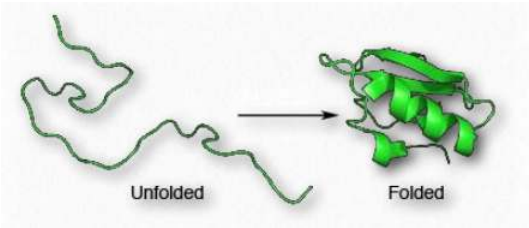


Figure 2.38 - Folding of a polypeptide chain

Proteins are distinguished from each other by the sequence of amino acids comprising them. The sequence of amino acids of a protein determines protein shape, since the chemical properties of each amino acid are forces that give rise to intermolecular

interactions to begin to create secondary structures, such as α-helices and β-strands. The sequence also defines turns and random coils that play important roles in the process of protein folding.

Since shape is essential for protein function, the sequence of amino acids gives rise to all of the properties a protein has. As protein synthesis proceeds, individual components of secondary structure start to interact with each other, giving rise to folds that bring amino acids close together that are not near each other in primary structure (Figure 2.38). At the tertiary level of structure, interactions among the R-groups of the amino acids in the protein, as well as between the polypeptide backbone and amino acid side groups play a role in folding.

Globular proteins

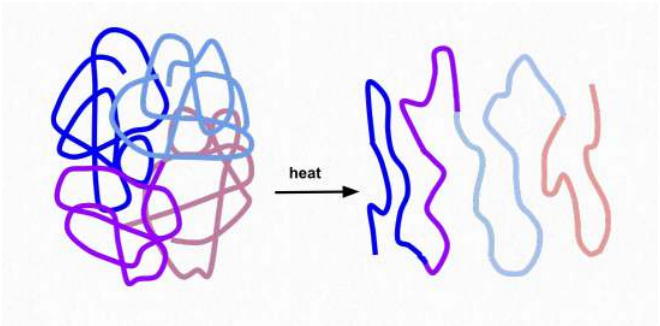


Figure 2.39 - Unfolding (denaturation) of a protein Wikipedia

Folding gives rise to distinct 3-D shapes in proteins that are non-fibrous. These proteins are called globular. A globular protein is stabilized by the same forces that drive its formation. These include ionic interactions, hydrogen bonding, hydrophobic forces, ionic bonds, disulfide bonds and metallic bonds. Treatments such as heat, pH changes, detergents, urea and mercaptoethanol overpower the stabilizing forces and cause a protein to unfold, losing its structure and (usually) its function (Figure 2.39). The ability of heat and detergents to denature proteins is why we cook our food and wash our hands before eating - such treatments denature the proteins in the microorganisms on our hands. Organisms that live in environments of high temperature (over 50°C) have proteins with changes in stabilizing forces - additional hydrogen bonds, additional salt bridges (ionic interactions), and compactness may all play roles in keeping these proteins from unfolding.

Protein stabilizing forces

Before considering the folding process, let us consider some of the forces that help to stabilize proteins.

Hydrogen bonds

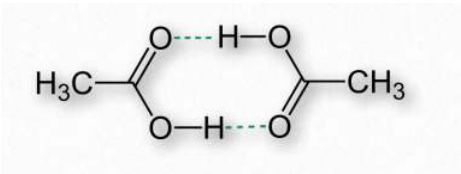


Figure 2.40 - Hydrogen bonds (dotted lines) between two molecules of acetic acid

Hydrogen bonds arise as a result of partially charged hydrogens found in covalent bonds. This occurs when the atom the hydrogen is bonded to has a greater electronegativity than hydrogen itself does, resulting in hydrogen having a partial positive charge because it is not able to hold electrons close to itself (Figure 2.40).

Hydrogen partially charged in this way is attracted to atoms, such as oxygen and nitrogen that have partial negative charges, due to having greater electronegativities and thus holding electrons closer to themselves. The partially positively charged hydrogens are called donors, whereas the partially negative atoms they are attracted to are called acceptors. (See Figure 1.30).

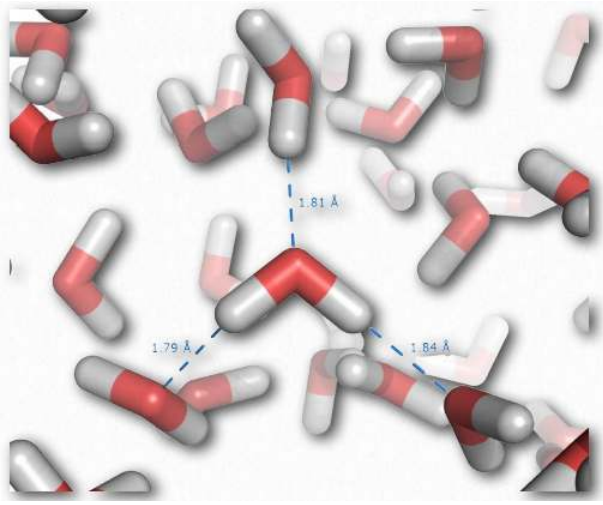


Figure 2.41 - Hydrogen bonding in liquid water Wikipedia

Individual hydrogen bonds are much weaker than a covalent bond, but collectively, they can exert strong forces. Consider liquid water, which contains enormous numbers of hydrogen bonds (Figure 2.41). These forces help water to remain liquid at room temperature. Other molecules lacking hydrogen bonds of equal or greater molecular weight than water, such as methane or carbon dioxide, are gases at the same temperature. Thus, the intermolecular interactions between water molecules help to “hold” water together and remain a liquid. Notably, only by raising the temperature of water to boiling are the forces of hydrogen bonding overcome, allowing water to become fully gaseous.

Hydrogen bonds are important forces in biopolymers that include DNA, proteins, and cellulose. All of these polymers lose their native structures upon boiling. Hydrogen bonds between amino acids that are close to each other in primary structure can give rise to regular repeating structures, such as helices or pleats, in proteins (secondary structure).

Ionic interactions

Ionic interactions are important forces stabilizing protein structure that arise from ionization of R-groups in the amino acids comprising a protein. These include the carboxyl amino acids (HERE), the amine amino acids as well as the sulfhydryl of cysteine and sometimes the hydroxyl of tyrosine.

Hydrophobic forces

Hydrophobic forces stabilize protein structure as a result of interactions that favor the exclusion of water. Non-polar amino acids (commonly found in the interior of proteins) favor associating with each other and this has the effect of excluding water. The excluded water has a higher entropy than water interacting with the hydrophobic side chains. This is because water aligns itself very regularly and in a distinct pattern when interacting with hydrophobic molecules.

When water is prevented from having these kinds of interactions, it is much more disordered that it would be if it could associate with the hydrophobic regions. It is partly for this reason that hydrophobic amino acids are found in protein interiors - so they can exclude water and increase entropy.

Disulfide bonds

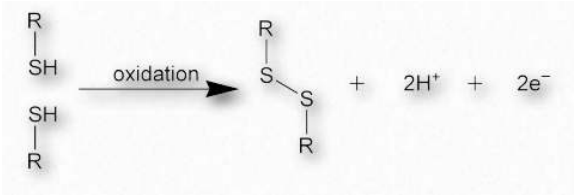


Figure 2.42 - Formation of a disulfide bond

Disulfide bonds, which are made when two sulfhydryl side-chains of cysteine are brought into close proximity, covalently join together different protein regions and can give great strength to the overall structure (Figures 2.42 & 2.43). An Ode to Protein Structure by Kevin Ahern The twenty wee amino A's Define a protein many ways Their order in a peptide chain Determines forms

that proteins gain And when they coil, it leaves me merry Cuz that makes structures secondary It's tertiary, I am told That happens when a protein folds But folded chains are downright scary When put together quaternary They're nature's wonders, that's for sure Creating problems, making cures A fool can fashion peptide poems But proteins come from ribosoems These joined residues of cysteine are sometimes referred to as cystine. Disulfide bonds are the strongest of the forces stabilizing protein structure.

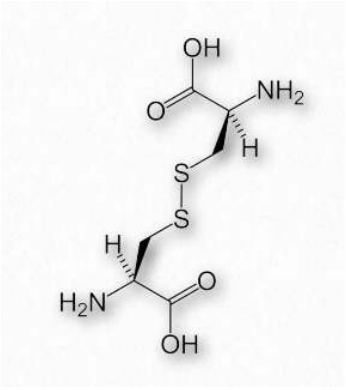


Figure 2.43 - Cystine - Two cysteines joined by a disulfide bond

van der Waals forces

van der Waals forces is a term used to describe various weak interactions, including those caused by attraction between a polar molecule and a transient dipole, or between two temporary dipoles. van der Waals forces are dynamic because of the fluctuating nature of the attraction, and are generally weak in comparison to covalent bonds, but can, over very short distances, be significant.

Post-translational modifications

Post-translational modifications can result in formation of covalent bonds stabilizing proteins as well. Hydroxylation of lysine and proline in strands of collagen can result in cross-linking of these groups and the resulting covalent bonds help to strengthen and stabilize the collagen.

Folding models

Two popular models of protein folding are currently under investigation. In the first (diffusion collision model), a nucleation event begins the process, followed by secondary structure formation. Collisions between the secondary structures (as in the β-hairpin in Figure 2.37) allow for folding to begin. By contrast, in the nucleation-condensation model, the secondary and tertiary structures form together.

Folding in proteins occurs fairly rapidly (0.1 to 1000 seconds) and can occur during synthesis - the amino terminus of a protein can start to fold before the carboxyl terminus is even made, though that is not always the case.

Folding process

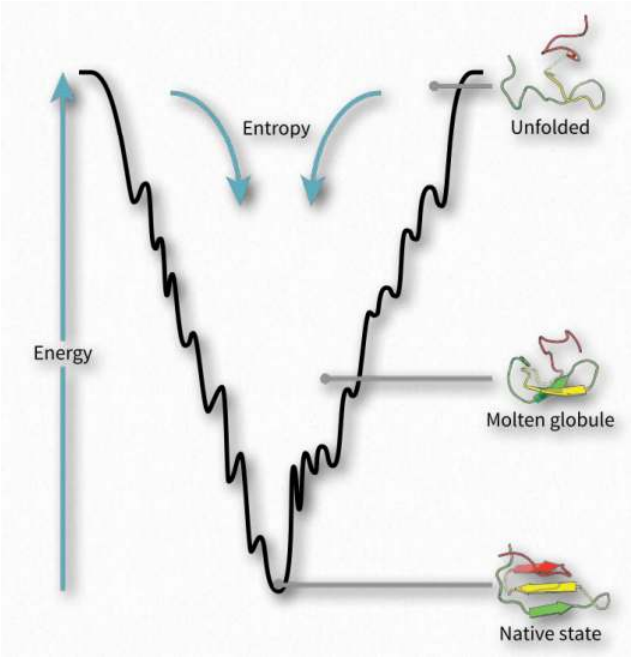


Figure 2.44 Folding funnel energy model of folding Wikipedia

Protein folding is hypothesized to occur in a “folding funnel” energy landscape in which a folded protein’s native state corresponds to the minimal free energy possible in conditions of the medium (usually aqueous solvent) in which the protein is dissolved. As seen in the diagram (Figure 2.44), the energy funnel has numerous local minima (dips) in which a folding protein can become trapped as it moves down the energy plot. Other factors, such as temperature, electric/magnetic fields, and spacial considerations likely play roles.

If external forces affect local energy minima during folding, the process and end-product can be influenced. As the speed of a car going down a road will affect the safety of the journey, so too do energy considerations influence and guide the folding process, resulting in fully functional, properly folded proteins in some cases and misfolded “mistakes” in others.

Getting stuck

As the folding process proceeds towards an energy minimum (bottom of the funnel in Figure 2.44), a protein can get “stuck” in any of the local minima and not reach the final folded state. Though the folded state is, in general, more organized and therefore has reduced entropy than the unfolded state, there are two forces that overcome the entropy decrease and drive the process forward.

The first is the magnitude of the decrease in energy as shown in the graph. Since $\Delta G = \Delta H - T\Delta S$, a decrease in ΔH can overcome a negative ΔS to make ΔG negative and push the folding process forward. Favorable (decreased) energy conditions arise with formation of ionic bonds, hydrogen bonds, disulfide bonds, and metallic bonds during the folding process. In addition, the hydrophobic effect increases entropy by allowing hydrophobic amino acids in the interior of a folded protein to exclude water, thus countering the impact of the ordering of the protein structure by making the ΔS less negative.

Structure prediction

Computer programs are very good at predicting secondary structure solely based on amino acid sequence, but struggle with determining tertiary structure using the same information. This is partly due to the fact that secondary structures have repeating points of stabilization based on geometry and any regular secondary structure (e.g., α-helix) varies very little from one to another. Folded structures, though, have an enormous number of possible structures as shown by Levinthal’s Paradox.

Spectroscopy

Because of our inability to accurately predict tertiary structure based on amino acid sequence, proteins structures are actually determined using techniques of spectroscopy. In these approaches, proteins are subjected to varied forms of electromagnetic radiation and the ways they interact with the radiation allows researchers to determine atomic coordinates at Angstrom resolution from electron densities (see X-ray crystallography) and how nuclei spins interact (see NMR).

Levinthal's paradox

In the late 1960s, Cyrus Levinthal outlined the magnitude of the complexity of the protein folding problem. He pointed out that for a protein with 100 amino acids, it would have 99 peptide bonds and 198 considerations for φ and ψ angles. If each of these had only three conformations, that would result in 3198 different possible foldings or 2.95x1094.

Even allowing a reasonable amount of time (one nanosecond) for each possible fold to occur, it would take longer than the age of the universe to sample all of them, meaning clearly that the process of folding is not occurring by a sequential random sampling and that attempts to determine protein structure by random sampling were doomed to fail. Levinthal, therefore, proposed that folding occurs by a sequential process that begins with a nucleation event that guides the process rapidly and is not unlike the funnel process depicted in Figure 2.44.

Diseases of protein misfolding

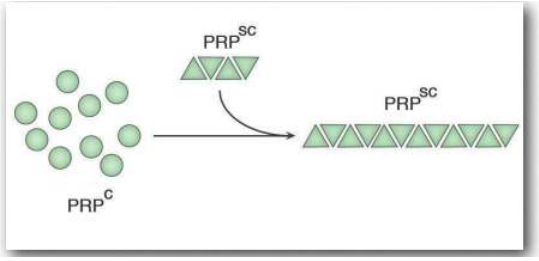


Figure 2.45 - Misfolding of the normal PRPc protein induced by PRPsc Image by Penelope Irving

The proper folding of proteins is essential to their function. It follows then that misfolding of proteins (also called proteopathy) might have consequences. In some cases, this might simply result in an inactive protein. Protein misfolding also plays a role in numerous diseases, such as Mad Cow Disease, Alzheimers, Parkinson's Disease, and CreutzfeldJakob disease. Many, but not all, misfolding diseases affect brain tissue.

Figure 2.46 - Cows with Mad Cow Disease lose their ability to stand

Insoluble deposits

Misfolded proteins will commonly form aggregates called amyloids that are harmful to tissues containing them because they change from being soluble to insoluble in water and form deposits. The process by which misfolding (Figure 2.45) occurs is not completely clear, but in many cases, it has been demonstrated that a "seed" protein which is misfolded can induce the same misfolding in other copies of the same protein. These seed proteins are known as prions and they act as infectious agents, resulting in the spread of disease. The list of human diseases linked to protein misfolding is long and continues to grow. A Wikipedia link is HERE.

Prions

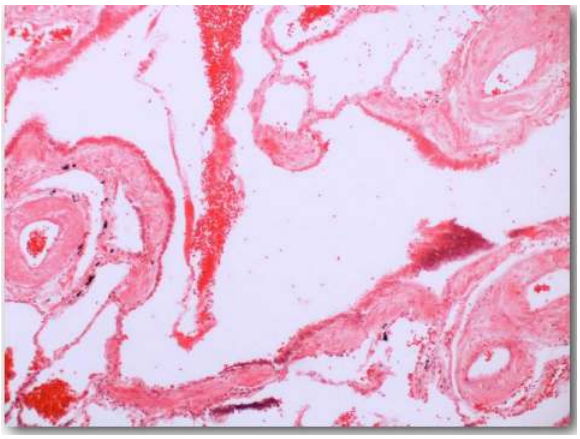


Figure 2.47 - Diffuse amyloidosis in a blood vessel (red dots) Wikipedia

Prions are infectious protein particles that cause transmissible spongiform encephalopathies (TSEs), the best known of which is Mad Cow disease. Other manifestations include the disease, scrapie, in sheep, and human diseases, such as CreutzfeldtJakob disease (CJD), Fatal Familial Insomnia, and kuru. The protein involved in these diseases is a membrane protein called PrP. PrP is encoded in the genome of many organisms and is found in most cells of the body. PrPc is the name given to the structure of PrP that is normal and not associated with disease. PrPSc is the name given to a misfolded form of the same protein, that is associated with the development of disease symptoms (Figure 2.45).

Misfolded

The misfolded PrPSc is associated with the TSE diseases and acts as an infectious particle. A third form of PrP, called PrPres can be found in TSEs, but is not infectious. The 'res' of PrPres indicates it is protease resistant. It is worth noting that all three forms of PrP have the same amino acid sequence and differ from each other only in the ways in which the polypeptide chains are folded. The most dangerously misfolded form of PrP is PrPSc, because of its ability to act like an infectious agent - a seed protein that can induce misfolding of PrPc , thus converting it into PrPSc.

Function

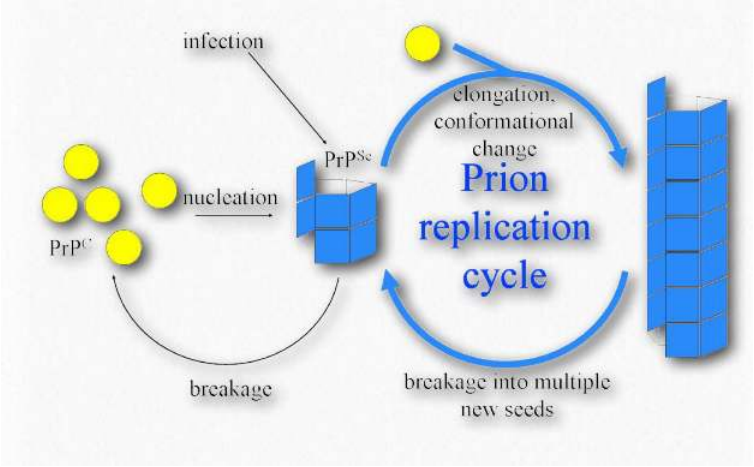


Figure 2.48 - One model of prion propagation Wikipedia

The function of PrPc is unknown. Mice lacking the PrP gene do not have major abnormalities. They do appear to exhibit problems with long term memory, suggesting a function for PrPc . Stanley Prusiner, who discovered prions and coined the term, received the Nobel Prize in Medicine in 1997 for his work. I think that if I chanced to be on A protein making up a prion I'd twist it and for goodness sakes Stop it from making fold mistakes

Amyloids

Amyloids are a collection of improperly folded protein aggregates that are found in the human body. As a consequence of their misfolding, they are insoluble and contribute to some twenty human diseases including important neurological ones involving

prions. Diseases include (affected protein in parentheses) - Alzheimer's disease (Amyloid β), Parkinson's disease (α-synuclein), Huntington's disease (huntingtin), rheumatoid arthritis (serum amyloid A), fatal familial insomnia (PrPSc), and others.

Amino acid sequence plays a role in amyloidogenesis. Glutamine-rich polypeptides are common in yeast and human prions. Trinucleotide repeats are important in Huntington's disease. Where sequence is not a factor, hydrophobic association between β-sheets can play a role.

Amyloid β

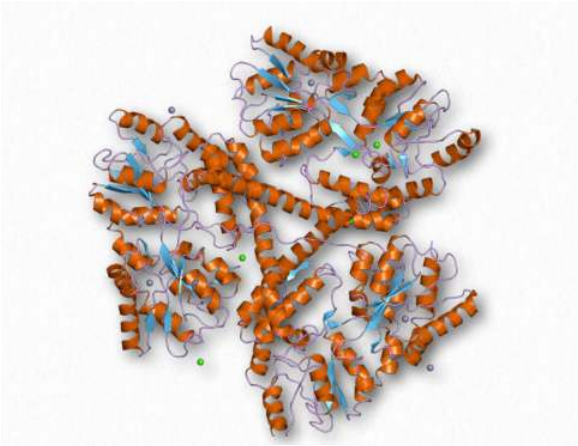


Figure 2.49 - Huntingtin

Amyloid β refers to collections of small proteins (36-43 amino acids) that appear to play a role in Alzheimer's disease. (Tau protein is the other factor.) They are, in fact, the main components of amyloid plaques found in the brains of patients suffering from the disease and arise from proteolytic cleavage of a larger amyloid precursor glycoprotein called Amyloid Precursor Protein, an integral membrane protein of nerve cells whose function is not known. Two proteases, β-secretase and γ- secretase perform this function. Amyloid β proteins are improperly folded and appear to induce other proteins to misfold and thus precipitate and form the amyloid characteristic of the disease. The plaques are toxic to nerve cells and give rise to the dementia characteristic of the disease.

It is thought that aggregation of amyloid β proteins during misfolding leads to generation of reactive oxygen species and that this is the means by which neurons are damaged. It is not known what the actual function of amyloid β is. Autosomal dominant mutations in the protein lead to early onset of the disease, but this occurs in no more than 10% of the cases. Strategies for treating the disease include inhibition of the secretases that generate the peptide fragments from the amyloid precursor protein.

Huntingtin

Huntingtin is the central gene in Huntington's disease. The protein made from it is glutamine rich, with 6-35 such residues in its wild-type form. In Huntington's disease, this gene is mutated, increasing the number of glutamines in the mutant protein to between 36 and 250. The size of the protein varies with the number of glutamines in the mutant protein, but the wild-type protein has over 3100 amino acids and a molecular weight of about 350,000 Da. Its precise function is not known, but huntingtin is found in nerve cells, with the highest level in the brain. It is thought to possibly play roles in transport, signaling, and protection against apoptosis. Huntingtin is also required for early embryonic development. Within the cell, huntingtin is found localized primarily with microtubules and vesicles.

Trinucleotide repeat

The huntingtin gene contains many copies of the sequence CAG (called trinucleotide repeats), which code for the many glutamines in the protein. Huntington's disease arises when extra copies of the CAG sequence are generated when the DNA of the gene is being copied. Expansion of repeated sequences can occur due to slipping of the polymerase relative to the DNA template during replication. As a result, multiple additional copies of the trinucleotide repeat may be made, resulting in proteins with variable numbers of glutamine residues. Up to 35 repeats can be tolerated without problem. The number of repeats can expand over the course of a person's lifetime, however, by the same mechanism. Individuals with 36-40 repeats begin to show signs of the disease and if there are over 40, the disease will be present.

Molecular chaperones

The importance of the proper folding of proteins is highlighted by the diseases associated with misfolded proteins, so it is no surprise, then, that cells expend energy to facilitate the proper folding of proteins. Cells use two classes of proteins known as molecular chaperones, to facilitate such folding in cells. Molecular chaperones are of two kinds, the chaperones, and the chaperonins. An example of the first category is the Hsp70 class of proteins. Hsp stands for “heat shock protein”, based on the fact that these proteins were first observed in large amounts in cells that had been briefly subjected to high temperatures. Hsps function to assist cells in stresses arising from heat shock and exposure to oxidizing conditions or toxic heavy metals, such as cadmium and mercury. However, they also play an important role in normal conditions, where they assist in the proper folding of polypeptides by preventing aberrant interactions that could lead to misfolding or aggregation. The Hsp70 proteins are found in almost all cells and use ATP hydrolysis to stimulate structural changes in the shape of the chaperone to accommodate binding of substrate proteins. The binding domain of Hsp70s contains a β-barrel structure which wraps around the polypeptide chain of the substrate and has affinity for hydrophobic side chains of amino acids. As shown in Figure 2.50, Hsp70 binds to polypeptides as they emerge from ribosomes during protein synthesis. Binding of substrate stimulates ATP hydrolysis and this is facilitated by another heat shock protein known as Hsp40. The hydrolysis of ATP causes the Hsp70 to taken on a closed conformation that helps shield exposed hydrophobic residues and prevent aggregation or local misfolding.

After protein synthesis is complete, ADP is released and replaced by ATP and this results in release of the substrate protein, which then allows the full length polypeptide to fold correctly.

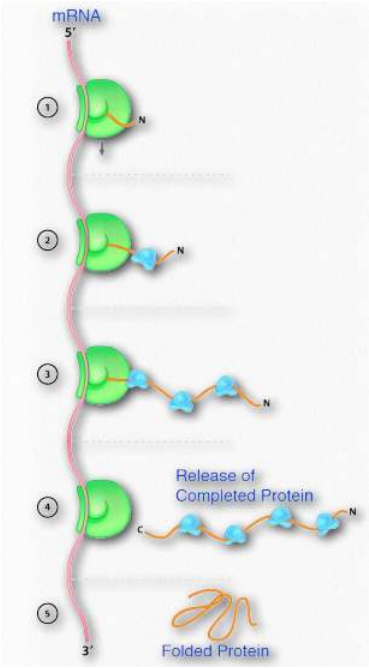


Figure 2.50 - Action of Hsp70 (blue) to facilitate proper folding of a protein (orange) Image by Aleia Kim

In heat shock

In times of heat shock or oxidative stress, Hsp70 proteins bind to unfolded hydrophobic regions of proteins to similarly prevent them from aggregating and allowing them to properly refold. When proteins are damaged, Hsp70 recruits enzymes that ubiquitinate the damaged protein to target them for destruction in proteasomes. Thus, the Hsp70 proteins play an important role in ensuring not only that proteins are properly folded, but that damaged or nonfunctional proteins are removed by degradation in the proteasome.

Chaperonins

A second class of proteins involved in assisting other proteins to fold properly are known as chaperonins. There are two primary categories of chaperonins - Class I (found in bacteria, chloroplasts, and mitochondria) and Class II (found in the cytosol of eukaryotes and archaebacteria). The best studied chaperonins are the GroEL/GroES complex proteins found in bacteria (Figure 2.51).

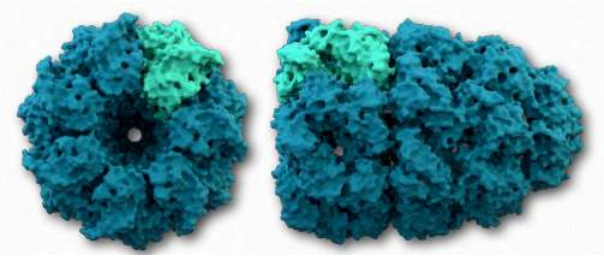


Figure 2.51 - View from bottom of GroEL (left) and GroEL/ GroES complex (right) Wikipedia

GroEL/GroES may not be able to undo aggregated proteins, but by facilitating proper folding, it provides competition for misfolding as a process and can reduce or eliminate problems arising from improper folding. GroEL is a double-ring 14mer with a hydrophobic region that can facilitate folding of substrates 15-60 kDa in size. GroES is a singlering heptamer that binds to GroEL in the presence of ATP and functions as a cover over GroEL. Hydrolysis of ATP by chaperonins induce large conformational changes that affect binding of substrate proteins and their folding. It is not known exactly how chaperonins fold proteins. Passive models postulate the chaperonin complex functioning inertly by preventing unfavorable intermolecular interactions or placing restrictions on spaces available for folding to occur. Active models propose that structural changes in the chaperonin complex induce structural changes in the substrate protein.

Protein breakdown

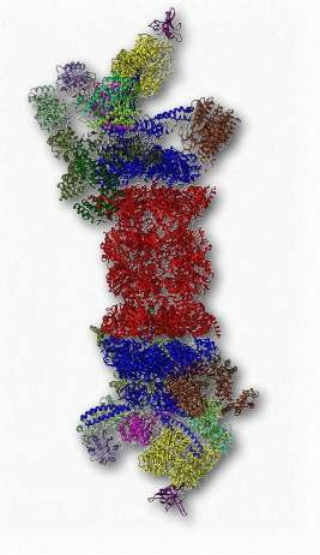


Figure 2.52 - 26S proteasome. Active site shown in red Wikipedia

Another protein complex that has an important function in the lifetime dynamics of proteins is the proteasome (Figure 2.52). Proteasomes, which are found in all eukaryotes and archaeans, as well as some bacteria, function to break down unneeded or damaged proteins by proteolytic degradation. Proteasomes help to regulate the concentration of some proteins and degrade ones that are misfolded. The proteasomal degradation pathway plays an important role in cellular processes that include progression through the cell cycle, modulation of gene expression, and response to oxidative stresses.

Degradation in the proteasome yields short peptides seven to eight amino acids in length. Threonine proteases play important roles. Breakdown of these peptides yields individual amino acids, thus facilitating their recycling in cells. Proteins are targeted for degradation in eukaryotic proteasomes by attachment to multiple copies of a small protein called ubiquitin (8.5 kDa - 76 amino acids). The enzyme catalyzing the reaction is known as ubiquitin ligase. The resulting polyubiquitin chain is bound by the proteasome and degradation begins. Ubiquitin was named due to it ubiquitously being found in eukaryotic cells.

Ubiquitin

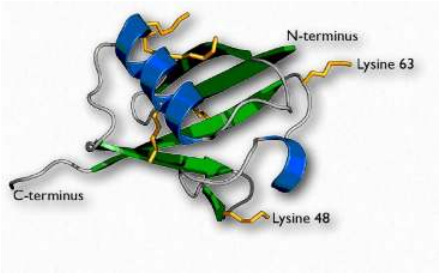


Figure 2.53 - Ubiquitin (lysine side chains shown in yellow) Wikipedia

Ubiquitin (Figure 2.53) is a small (8.5 kDa) multi-functional protein found in eukaryotic cells. It is commonly added to target proteins by action of ubiquitin ligase enzymes (E3 in Figure 2.54). One (ubiquitination) or many (polyubiquitination) ubiquitin molecules may be added. Attachment of the ubiquitin is through the side chain of one of seven different lysine residues in ubiquitin.

The addition of ubiquitin to proteins has many effects, the best known of which is targeting the protein for degradation in the proteasome. Proteasomal targeting is seen when polyubiquitination occurs at lysines #29 and 48. Polyubiquitination or monoubiquitination at other lysines can result in altered cellular location and changed protein-protein interactions. The latter may alter affect inflammation, endocytic trafficking, translation and DNA repair.

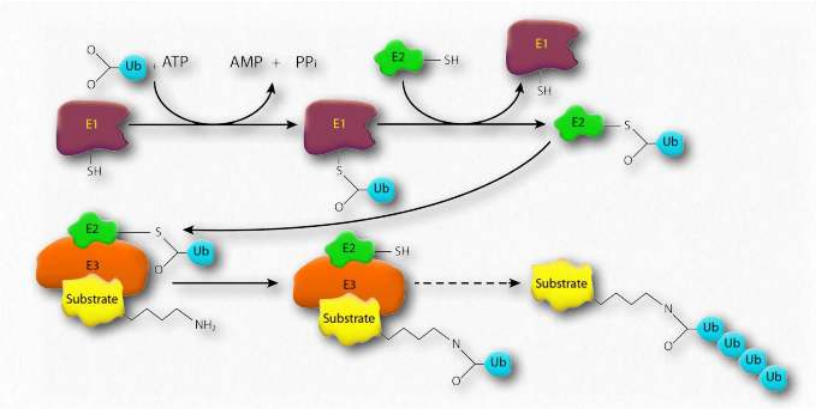
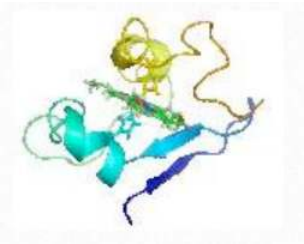


Figure 2.54 - Pathway for ubiquitination of a target substrate protein Image by Pehr Jacobson

Ubiquitin ligase malfunction

Parkin is a Parkinson's disease-related protein that, when mutated, is linked to an inherited form of the disease called autosomal recessive juvenile Parkinson's disease. The function of the protein is not known, but it is a component of the E3 ubiquitin ligase system responsible for transferring ubiquitin from the E2 protein to a lysine side chain on the target protein. It is thought that mutations in parkin lead to proteasomal dysfunction and a consequent inability to break down proteins harmful to dopaminergic neurons. This results in the death or malfunction of these neurons, resulting in Parkinson's disease.

Intrinsically disordered proteins

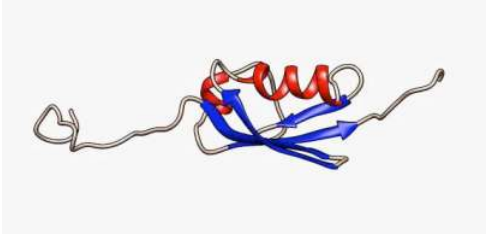


Movie 2.1 - Dynamic movement of cytochrome C in solution Wikipedia

As is evident from the many examples described elsewhere in the book, the 3-D structure of proteins is important for their function. But, increasingly, it is becoming evident that not all proteins fold into a stable structure. Studies on the so-called intrinsically

disordered proteins (IDPs) in the past cou- ple of decades has shown that many proteins are biologically active, even thought they fail to fold into stable structures. Yet other proteins exhibit regions that remain unfolded (IDP regions) even as the rest of the polypeptide folds into a structured form.

Intrinsically disordered proteins and disordered regions within proteins have, in fact, been known for many years, but were regarded as an anomaly. It is only recently, with the realization that IDPs and IDP regions are widespread among eukaryotic proteins, that it has been recognized that the observed disorder is a "feature, not a bug".



Movie 2.2 SUMO-1, a protein with intrinsically disordered sections Wikipedia

Comparison of IDPs shows that they share sequence characteristics that appear to favor their disordered state. That is, just as some amino acid sequences may favor the folding of a polypeptide into a particular structure, the amino acid sequences of IDPs favor their remaining unfolded. IDP regions are seen to be low in hydrophobic residues and unusually rich in polar residues and proline. The presence of a large number of charged amino acids in the IDPs can inhibit folding through charge repulsion, while the lack of hydrophobic residues makes it difficult to form a stable hydrophobic core, and proline discourages the formation of helical structures. The observed differences between amino acid sequences in IDPs and structured proteins have been used to design algorithms to predict whether a given amino acid sequence will be disordered.

What is the significance of intrinsically disordered proteins or regions? The fact that this property is encoded in their amino acid sequences suggests that their disorder may be linked to their function. The flexible, mobile nature of some IDP regions may play a crucial role in their function, permitting a transition to a folded structure upon binding a protein partner or undergoing post-translational modification. Studies on several wellknown proteins with IDP regions suggest some answers. IDP regions may enhance the ability of proteins like the lac repressor to translocate along the DNA to search for specific binding sites. The flexibility of IDPs can also be an asset in protein-protein interactions, especially for proteins that are known to interact with many different protein partners.

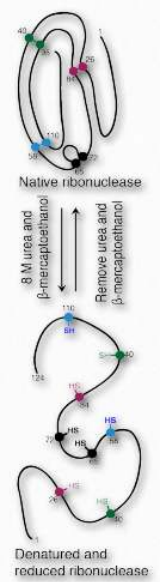


Figure 2.55 - Denaturation and renaturation of ribonuclease Wikipedia

For example, p53 has IDP regions that may allow the protein to interact with a variety of functional partners. Comparison of the known functions of proteins with predictions of disorder in these proteins suggests that IDPs and IDP regions may

disproportionately function in signaling and regulation, while more structured proteins skew towards roles in catalysis and transport. Interestingly, many of the proteins found in both ribosomes and spliceosomes are predicted to have IDP regions that may play a part in correct assembly of these complexes. Even though IDPs have not been studied intensively for very long, what little is known of them suggests that they play an important and underestimated role in cells.

Metamorphic proteins

Another group of proteins that have recently changed our thinking about protein structure and function are the so-called metamorphic proteins. These proteins are capable of forming more than one stable, folded state starting with a single amino acid sequence. Although it is true that multiple folded conformations are not ruled out by the laws of physics and chemistry, metamorphic proteins are a relatively new discovery. It was known, of course, that prion proteins were capable of folding into alternative structures, but metamorphic proteins appear to be able to toggle back and forth between two stable structures. While in some cases, the metamorphic protein undergoes this switch in response to binding another molecule, some proteins that can accomplish this transition on their own. An interesting example is the signaling molecule, lymphotactin. Lymphotactin has two biological functions that are carried out by its two conformers- a monomeric form that binds the lymphotactin receptor and a dimeric form that binds heparin. It is possible that this sort of switching is more widespread than has been thought.

Refolding denatured proteins

All information for protein folding is contained in the amino acid sequence of the protein. It may seem curious then that most proteins do not fold into their proper, fully active form after they have been+++ denatured and the denaturant is removed. A few do, in fact. One good example is bovine ribonuclease (Figure 2.55). Its catalytic activity is very resistant to heat and urea and attempts to denature it don't work very well. However, if one treats the enzyme with β-mercaptoethanol (which breaks disulfide bonds) prior to urea treatment and/or heating, activity is lost, indicating that the covalent disulfide bonds help stabilize the overall enzyme structure and when they are broken, denaturation can readily occur. When the mixture cools back down to room temperature, over time some enzyme activity reappears, indicating that ribonuclease re-folded under the new conditions.

Interestingly, renaturation will occur maximally if a tiny amount of β-mercaptoethanol is left in the solution during the process. The reason for this is because β- mercaptoethanol permits reduction (and breaking) of accidental, incorrect disulfide bonds during the folding process. Without it, these disulfide bonds will prevent proper folds from forming.

Irreversible denaturation

Most enzymes, however, do not behave like bovine ribonuclease. Once denatured, their activity cannot be recovered to any significant There are not very many ways Inactivating RNase It's stable when it's hot or cold Because disulfides tightly hold If you desire to make it stall Use hot mercaptoethanol extent. This may seem to contradict the idea of folding information being inherent to the sequence of amino acids in the protein. It does not.

Most enzymes don't refold properly after denaturation for two reasons. First, normal folding may occur as proteins are being made. Interactions among amino acids early in the synthesis are not "confused" by interactions with amino acids later in the synthesis because those amino acids aren't present as the process starts.

Chaperonins' role

In other cases, the folding process of some proteins in the cell relied upon action of chaperonin proteins (see HERE). In the absence of chaperonins, interactions that might result in misfolding occur, thus preventing proper folding. Thus, early folding and the assistance of chaperonins eliminate some potential "wrong-folding" interactions that can occur if the entire sequence was present when folding started.

Quaternary structure

A fourth level of protein structure is that of quaternary structure. It refers to structures that arise as a result of interactions between multiple polypeptides. The units can be identical multiple copies or can be different polypeptide chains. Adult hemoglobin is a good example of a protein with quaternary structure, being composed of two identical chains called α and two identical chains called β.

Though the α-chains are very similar to the β- chains, they are not identical. Both of the α- and the β-chains are also related to the single polypeptide chain in the related protein called myoglobin. Both myoglobin and hemoglobin have similarity in binding

oxygen, but their behavior towards the molecule differ significantly. Notably, hemoglobin's multiple subunits (with quaternary structure) compared to myoglobin's single subunit (with no quaternary structure) give rise to these differences.

References

1. https://en.Wikipedia.org/wiki/Van_der_W aals_force 105

This page titled 4.2: Levels of Protein Structure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: Structure & Function- Proteins I** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **4.2: Levels of Protein Structure** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

4.3: Secondary Structure

Secondary structures are those repetitive structures involving H bond between amide H and carbonyl O in- the main chain. These include alpha helices, beta strands (sheets) and reverse turns.

Alpha Structure

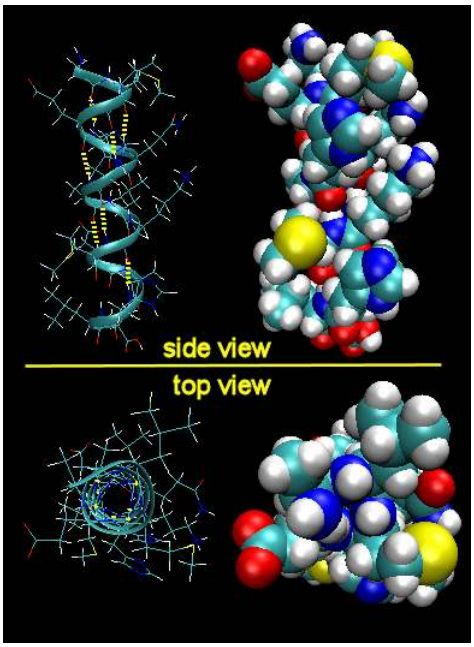


***Figure**: Right Handed Alpha helices - image made with VMD*

These helices are formed when the carbonyl O of the i th amino acid H bonds to the amide H of the i th +4 aa (4 amino acids away). The phi/psi angles for those amino acids in the alpha helix are - 57,-47, which emphasizes the regular repeating nature of the structure. It can also be characterized by n (the number of amino acid units/turn = 3.6) and pitch (the helix rise/turn = 5.4 angstroms). Some facts:

- the alpha helix is more compact than the fully extended polypeptide chain with phi/psi angles of 180o
- in proteins, the average number of amino acids in a helix is 11, which gives 3 turns.
- the left-handed alpha helix, although allowed from inspections of a Ramachandran plot, is never observed, since the side chains are too close to the backbone.
- the core of the helix is packed tightly. There are not holes or pores in the helix.
- All the R-groups extend backward and away from the helix axis.
- Some amino acids are more commonly found in alpha helices than other. Amino acids can be divided into two kinds, those with branches at the beta C and those with none. Consider first those that aren't branched . Gly is too conformationally flexible to be found with high frequency in alpha helices, while Pro is too rigid. The amino acids with side chains that can H-bond (Ser, Asp, and Asn) and aren't too long appear to act as competitors of main chain H bond donor and acceptors, and destabilize alpha helices. The rest with no branches at the beta C can form helices. Those with branches at the beta carbon (Val, Ile) destabilize the alpha helix due to steric interactions of the bulky side chains with the helix backbone. (Remember left-handed alpha helices are not found in nature for similar reasons.) Summary of amino acids propensities for alpha helices (and beta structure as well)
- alpha keratins, the major component of hair, skin, fur, beaks, and fingernails, are almost all alpha helix.

Jmol: Updated An isolated helix from an Antifreeze Protein Jmol14 (Java) | JSMol (HTML5)

helices

There are other kinds of helices that can occur. These include a 310 helix and a \(\pi\) helix, which are stabilized by H-bonds between the amide NH and carbonyl O of residues (i, i+3) and (i, i+5), respectively. Likewise, they have 3 and 4.3 residues/turn, respectively, and a rise per residue of 6 and 4.7 angstrom, respectively. These structures are much rarer than right handed alpha helices.

Helix Type	H bond btw ith and ith+X AA, where X =	Residue/turn	Rise (Angstrom)/turn
310	3	3	6
α	4	3.6	5.4
π	5	4.3	4.7

Beta Structure

Beta Structure: Parallel and antiparallel beta strands are much more extended than alpha helices (phi/psi of -57,-47) but not as extended as a fully extended polypeptide chain (with phi/psi angles of +/- 180). The beta sheets are not quit so extended (parallel -119, +113 ; antiparallel, -139, +135), and can be envisioned as rippled sheets. They can be visualized by laying thin, pleated strips of paper side by side to make a "pleated sheet" of paper. Each strip of paper can be pictured as a single peptide strand in which the peptide backbone makes a zig-zag along the strip, with the alpha carbons lying at the folds of the pleats. Each single strand of the beta-sheet can be pictured as a twofold helix, i.e. a helix with 2 residues/turn. The arrangement of each successive peptide plane is pleated due to the tetrahedral nature of the alpha C. The H bonds are interstrand, not intrastrand as in the alpha helix.

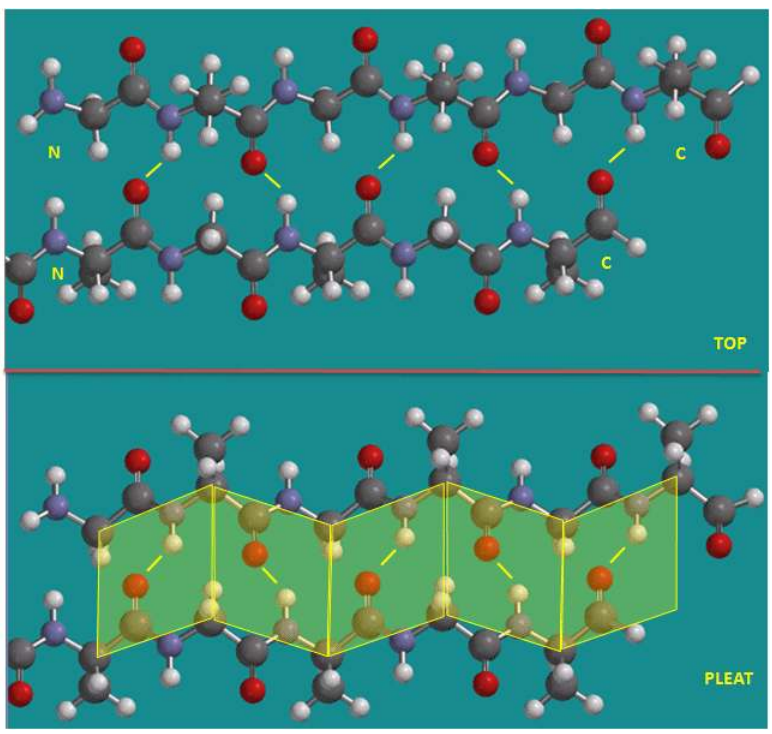


Figure: Parallel beta strands (image made with Spartan)

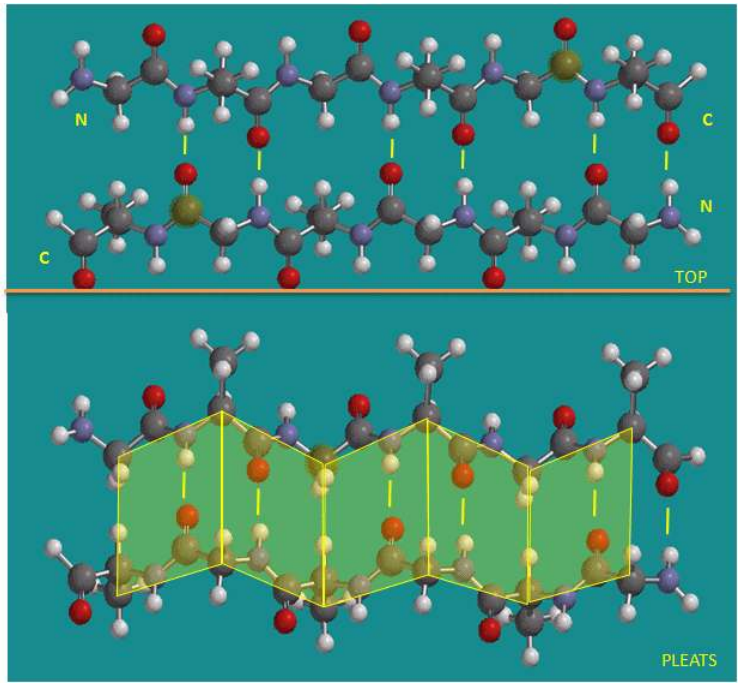


Figure: Antiparallel beta strands (image made with Spartan)

Note

Consider a strand as a continuous and contiguous polypeptide backbone propagating in one direction. Hence, using this definition, a helix consist of a single strand, and all the H-bonds are within the strand (or intrastrand). A beta sheet would then consist of multiple strands, since each "strand" is separated from other "strands" by an intervening contiguous stretch of amino acid which bends within the protein in a way which allows the next section of the peptide backbone, the next "strand", to H-bond with the first "strand". But remember, even in this case, all the H-bonds holding the alpha and beta structure together are intramolecular.

In a parallel beta sheet structure, the optimal H bond pattern leads to a less extended structure (phi/psi of -119, +113) than the optimal arrangement of the H bonds in the antiparallel structure (phi/psi of -139, +135). Also the H bonds in the parallel sheet are bent significantly. (i.e. the carbonyl O on one strand is not exactly opposite the amide H on the adjacent strand, as it is in the antiparallel sheet.) Hence antiparallel beta strands are presumably more stable, even though both are abundantly found in nature. Short parallel beta sheets of 4 strands or less are not common, which might reflect their lower stability.

The side chains in the beta sheet are normal to the plane of the sheet, extending out from the plane on alternating sides. Parallel sheets characteristically distribute hydrophobic side chains on both side of the sheet, while antiparallel sheets are usually arranged with all the hydrophobic residues on one side. This requires an alternation of hydrophilic and hydrophobic side chains in the primary sequence. Antiparallel sheets are found in silk with the sheets running parallel to the silk fibers. The following repeat is found in the primary sequence: (Ser-Gly-Ala-Gly)n), with Gly pointing out from one face, and Ser or Ala from the other.

EXTERNAL Jmol: Silk

Beta strands have a tendency to twist in the right hand direction. This leads to important consequences in how the beta strands are connected. Parallel strands can form twisted sheets or saddles as well as beta barrels.

Figure: Twisted Beta Sheet/Saddle (image made with VMD)

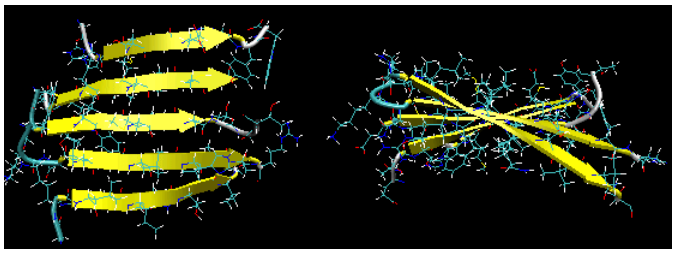
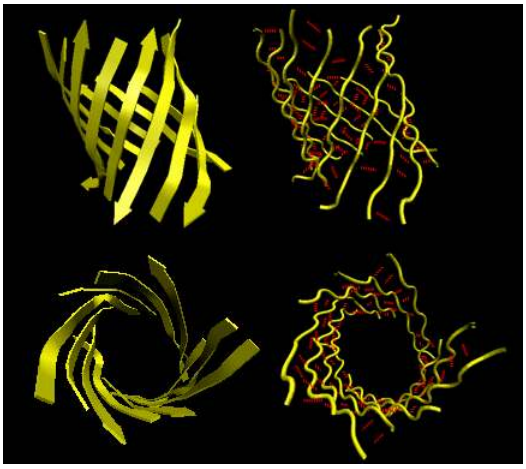


Figure: Beta Barrel (image made with VMD)



- in parallel strands, right handed connectivity is common.
- in a protein with parallel strand in register, and given the inherent twist in the stands, the strands arrange in a way to have the H bonds stretched equally at the ends of the chains, giving rise to a twisted saddle shape (top structure above).

Jmol: Updated Twisted beta sheet from Arabinose Binding Protein Jmol14 (Java) | JSMol (HTML5)

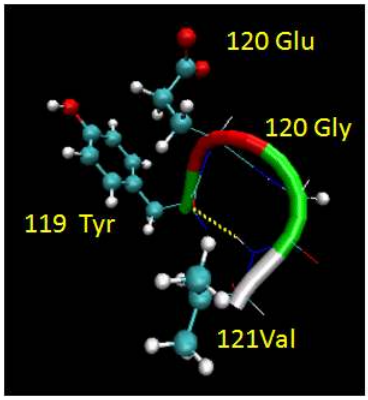
- in a protein with parallel strand out of register, and given the inherent twist in the stands, the strands arrange in a way to have the H bonds stretched equally at the ends of the chains, giving rise to a beta barrel (bottom structure above).

Jmol: Beta barrel from triose phosphate isomerase

Reverse Turns

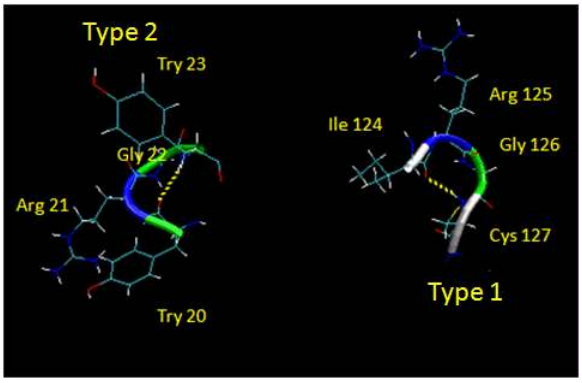
About 50% of the amino acids in a globular protein are in regular secondary structure (alpha or beta). The remaining amino acids are not less ordered, just less regular. An additional example of secondary structures is reverse turns (or beta-bends or beta turns). Reverse turns often connect successive antiparallel beta strands and are then called **beta hairpins**.

Figure: Reverse Turns (image made with VMD)



They are almost always at the surface, and consist of 4 amino acids. There are two types. (I - f2 = -60, y2=-30; f3 = -90, y3 = 0; II - f2 = -60, y2=120; f3 = 90, y3 = 0) Residue 2 of both is often Pro. (Why?) Both have an H bond between the carbonyl O of the i th a.a and the amide H of the i th+3 aa (three amino acids away). In the type 2, the O of residue 2 crowds the beta C of residue 3, so aa2 is usually Gly. Why? Those amino acids which destabilize alpha helices are often found in beta sheets, since the side chains project out of the plan which holds the main chain.

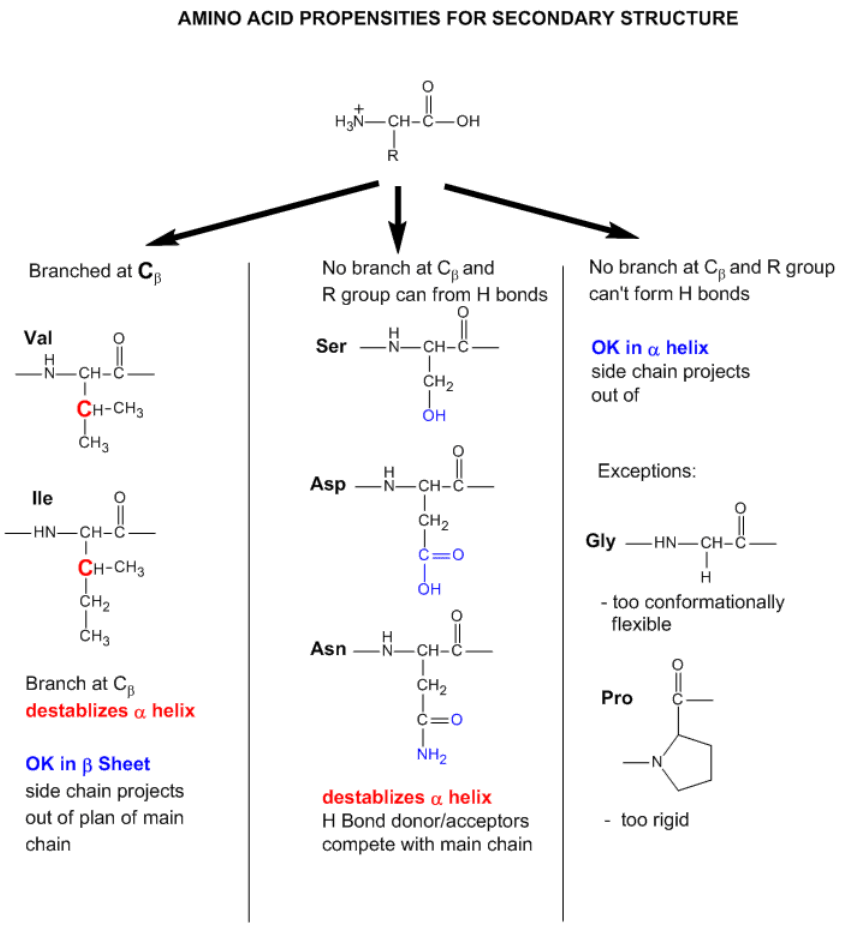
Figure: Type I and Type II Reverse Turns - from hen egg white lysozyme (image make with VMD)



Jmol: Updated Reverse Turn: Trypsin Inhibitor Jmol14 (Java) | JSMol (HTML5)

(Notice the tightness of the reverse turn and the presence of Pro and Gly.)

Figure: Why do amino acid propensities for secondary structure differ?



- EXTERNAL Secondary Structure and Backbone Conformation from ExPASy.(concentrate on second part on secondary structure)

This page titled 4.3: Secondary Structure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

- **C2. Secondary Structure** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.
- **4.3: Secondary Structure** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.

4.4: Protein folding

We've seen many static images of lipids and their aggregates as well as proteins. To think about how proteins fold, we have to think dynamically. Luckily we have the tools of molecular dynamics (MD) at our fingertips which helps us imagine how these processes take place and concomitantly how to probe protein folding experimentally. View the following two MD simulations and compare the spontaenously formation of a micelle and the folding of a protein before we delve into the complex topic of protein folding and stability.

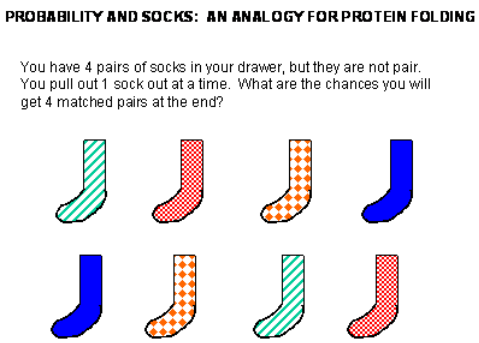
- MD simulation of micelle formation
- MD simulation of protein folding

Given the number of possibly nonnative states, it is amazing that proteins fold to the native state at all, let alone in a reasonable time frame. Consider this greatly simplified view of protein folding for a protein containing 100 amino acids. If each amino acid can adopt only 3 possible conformations, the total number of conformations could be 3100 = 5 x 10^{47}. Assuming that it would take 10-13s to change each conformation, the time required to "test" all conformations would be 5 x 1034s or 1027 years, longer than the age of the universe (14 x 109 yr). Yet the protein can fold within seconds. This paradox is called the *Levinthal paradox*, after Cyrus Levinthal.

Lubert Stryer (in his classic Biochemistry text), shows a way out of this dilemma by using an analogy of a monkey sitting at a typewriter, and typing this line out of Hamlet: "Me thinks it is like a weasel." Random typing would produce that line after 1040 keystrokes on average, but if the correct letters were maintained, the number of keystrokes would be in the realm of a few thousand. Proteins could fold more quickly if they retain native-like intermediates along the way. Also remember that much of conformation space is already restricted by allowed phi/psi angles. (Remember the blank areas in the Ramachandran plot?)

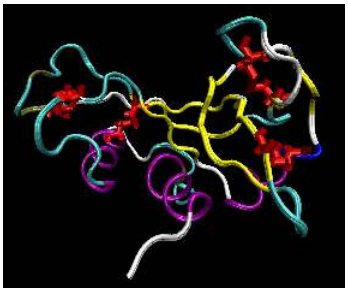
Before we study the classic experiment of protein folding conducted by Anfinsen, study the simpler analogy below:

Figure: Socks and protein folding



The classic experiment of Anfinsen has shown that, at least for some proteins, all the necessary and sufficient information required to direct the folding of a protein into the native state is present in the primary sequence of a protein. Anfinsen studied the *in vitro* (outside the cell, as opposed to *in vivo*, which is inside the cell, tissue, organ) folding of a single chain protein, RNase, which has 4 intrachain disulfide bonds.

Figure: RNase A with 4 Disulfide Bonds in red (image with VMD)



We have previously discussed how chemical agents (such as beta-mercaptoethanol, a disulfide reducing agent) can covalently interact with specific protein functional groups. Other substances can bind through complementary intermolecular forces to the active site or other cavities on the surface. Other reagents, like urea, acting through generalized solvent changes or nonspecific interactions with the protein, can alter protein folding. Anfinsen used two different reagents, 8 M urea and beta-mercaptoethanol, in combination to unfold, or denature, RNase to the nonnative or denatured state. He then removed the bME using dialysis, allowing the disulfides to reform. Next he removed the denaturing reagent, urea. To monitor if the protein was correctly refolded or renatured, he tested the activity of the protein compared to native protein. He found that the "refolded" protein retained only 1% of its initial activity. If, however, he added a catalytic amounts of bME, the protein soon retained 100% of its initial activity. For his work, he was awarded the Nobel Prize in Chemistry in 1972.

Figure: Anfinsen Experiment: Folding of RNase

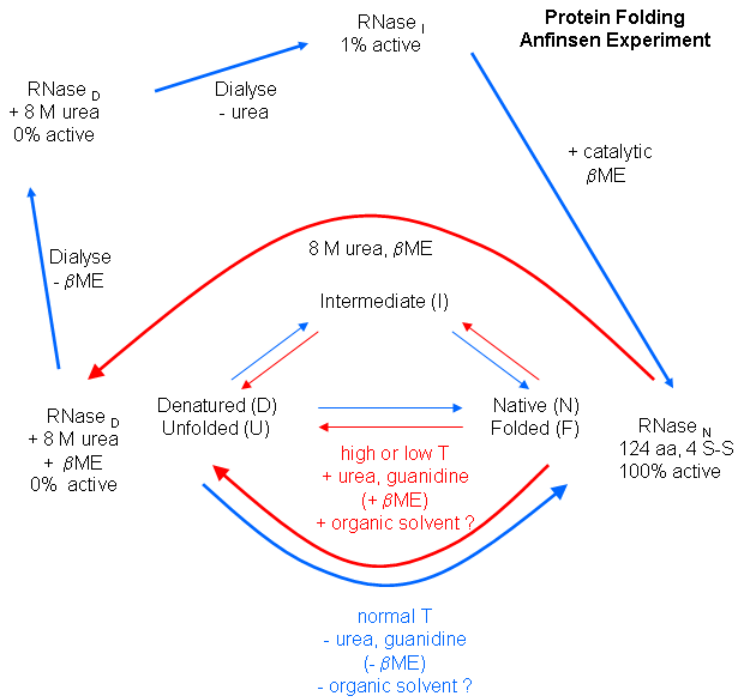
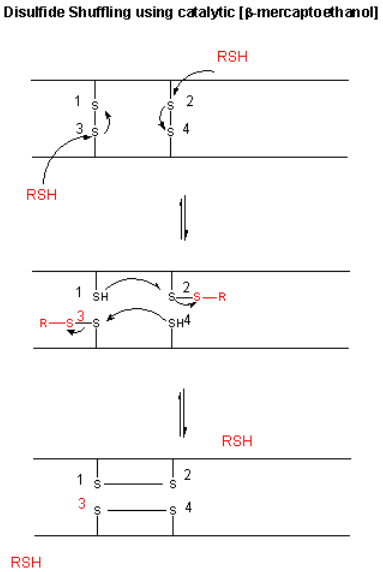


Figure: CATALYTIC SHUFFLING OF DISULFIDES WITH BETA-MERCAPTOETHANOL



Jmol: Updated RNase A Jmol14 (Java) | JSMol (HTML5)

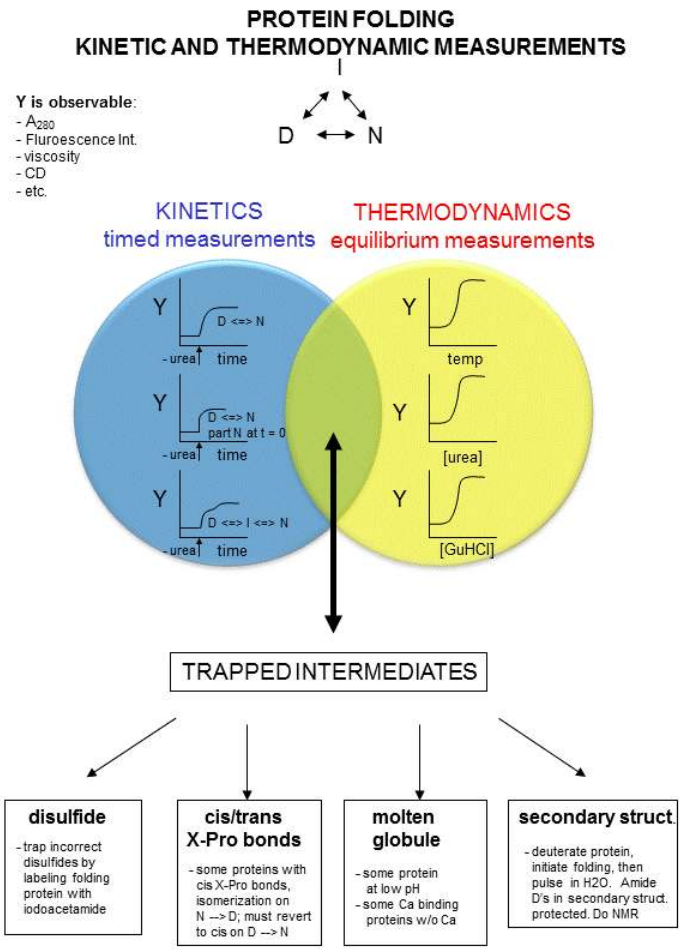
Scientists have investigated the folding of proteins both in vitro and in vivo. In vitro experiments involve denaturing the protein with urea, guanidine hydrochloride, or heat, then refolding the protein by removing the perturbant (denaturing agent), using spectral techniques to follow the process. In vivo experiments involve the study of intracellular proteins that assist folding. The in vitro experiments involve unfolding the native state and then refolding it, while the in vivo ones involve folding of the newly synthesized protein. An understanding of protein folding can not be separated from an understanding of protein stability, and an understanding of the nature of the native and denatured state.

In studying protein folding and stability/structure of the native and denatured states, both equilibrium (thermodynamic) and timed (kinetic) measurements are made. Folding occurs in the ms to second range, which limits the ability to study the presence of intermediates in the process. Some clever methods have been developed to study intermediates in protein folding by trapping specific intermediate structures, and investigating their structure and stability in a "leisurely" fashion. Alternatively, intermediates can be studied as they occur using stop flow kinetics. In this technique, a protein under denaturing conditions is rapidly mixed with a solution containing no denaturant or protein by injecting both solutions into a mixer/cuvette using syringes. The denaturant in the protein solution is now diluted such that renaturation can occur. Spectral measurements can begin at once.

- stop flow Java applet

A diagram summarizing these methods is shown below. Study it in conjunction with the text which follows.

Figure: Kinetic and thermodynamic measurements of proteins stability and folding

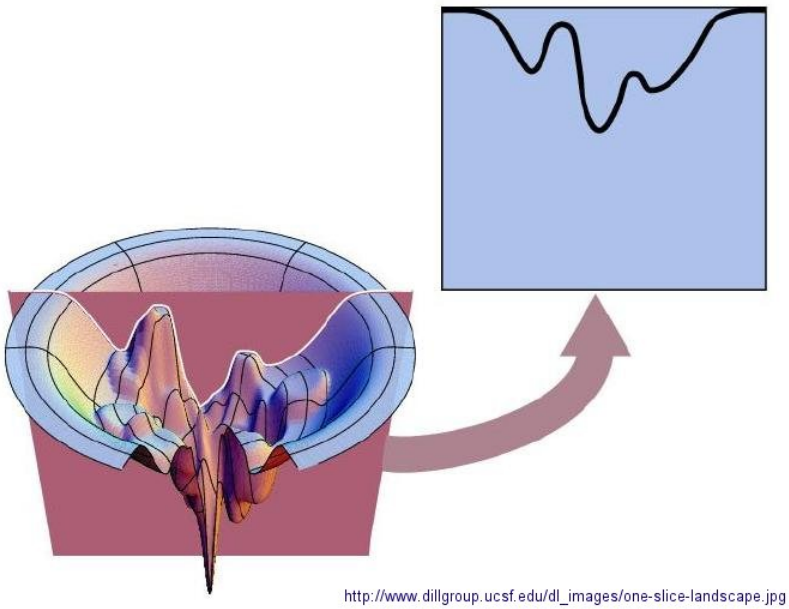


In considering the folding pathway, we will consider that the native protein represents the global energy minimum. All other states represent variations of the denatured state. Some, closer in energy to the native state, could be considered intermediates in the

folding process. Instead of considering a folding "trajectory", consider protein folding occurring within a large folding landscape of free energy. Folding appears to proceed not by an obligatory pathway but a probabilistic or stochastic search of possible conformation. The free energy landscape must be shaped somewhat like a funnel such that a proteins could adopt a "reasonable" number of conformations which lead to the native state. Evolution has surely selected for sequences that can make it to that state. Localized secondary structure motifs (like a short alpha helix and beta turns) can form quickly (about 1 ms). Small proteins folding occurs, depending on the structure, over a wide time frame (ms to minutes). Mostly likely, a small number of amino acids coalesce into a core which nucleates folding into structures that are similar to the native state. Finally packing interactions collapse the structure into the native state.

In general, the more complex the fold of the backbone, the longer it takes the protein to fold. If complexity requires more interactions among distal regions of the polypeptide change, then the more complex the fold, the less probable that random interactions would lead to quick protein folding. The mechanisms of folding for larger proteins (greater than 100 amino acids) appear to proceed through intermediates, suggesting that different domains of the protein can fold independently.

Figure: Protein Folding Landscape: One View from Ken Dill



Contributors and Attributions

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

This page titled 4.4: Protein folding is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

- **D1. Introduction** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.
- **4.4: Protein folding** by Henry Jakubowski is licensed CC BY-NC-SA 4.0.

4.5: Analyses of Protein Structure

Notes to students and instructors:

This section is placed in Chapter 3 before the in-depth look at protein structure in Chapter 4. It could be placed in either chapter from a learning and teaching perspective. Proteins are very complicated. This section could help students understand the tools used to explore and understand protein structure/function before a deeper dive into structural features found in Chapter 4 and throughout the book. Alternatively, it could also deepen an acquired understanding of proteins if used at the end of Chapter 4.

This comprehensive section covers topics (such as mass spectrometry, NMR, and molecular dynamics) that might not be broadly discussed or even mentioned in a one-semester survey course. Some of the methods may be more historic and less relevant today. Instructors can choose parts that are appropriate to their learning goals.

Learning Goals (ChatGPT o1

- **Conceptualize Structural Resolution in Protein Analysis:**
  - Explain the concept of “resolution” and how different analytical techniques (from low to high resolution) reveal various levels of protein structure and function.
- **Master Low-Resolution Analytical Techniques:**
  - Describe methods for determining protein concentration (e.g., UV absorbance, dye-binding assays, BCA) and discuss their advantages, limitations, and sources of interference.
  - Outline techniques for amino acid composition analysis, N-/C-terminal analysis, and primary sequence determination, including the use of chemical assays and Edman degradation.
- **Understand Spectral Techniques for Structural Insights:**
  - Explain how circular dichroism (CD) spectroscopy provides information about secondary structure by differentiating between α-helices, β-sheets, and random coils.
  - Discuss the principles behind fluorescence spectroscopy, including intrinsic versus extrinsic fluorophores, fluorescence quenching, and fluorescence resonance energy transfer (FRET) for studying protein-ligand interactions and conformational changes.
  - Define key fluorescence concepts such as Stokes shift and anisotropy, and how they reflect changes in the protein’s environment or dynamics.
- **Grasp Mass Spectrometry for Protein Analysis:**
  - Describe the principles of mass spectrometry for protein characterization, including ionization techniques (ESI and MALDI), mass analyzers, and the calculation of molecular weight from m/z values.
  - Differentiate between top-down, bottom-up, and middle-down approaches in proteomic analysis, and explain how peptide mass fingerprinting and tandem MS/MS contribute to protein sequencing and post-translational modification mapping.
- **Explore High-Resolution Structural Determination:**
  - Compare X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy (cryo-EM) in terms of sample requirements, strengths, limitations, and the type of structural information provided.
  - Discuss how time-resolved crystallography and multidimensional NMR techniques offer insights into protein dynamics and conformational changes.
- **Integrate Computational and Dynamic Analyses:**
  - Explain the basic principles of molecular mechanics and molecular dynamics (MD) simulations, including the use of force fields to model bonded and nonbonded interactions.
  - Discuss how MD simulations help predict conformational changes, calculate free energy differences (ΔG), and complement experimental data from high-resolution methods.
- **Contextualize Proteome Complexity:**
  - Define the proteome and discuss the challenges of proteomic analysis due to dynamic protein expression, diverse post-translational modifications, and the wide range of protein abundances.
  - Describe how integrated workflows (e.g., 2D electrophoresis combined with LC-MS/MS) are used in proteomics to identify and quantify proteins in complex biological samples.
- **Synthesize Structure/Function Relationships:**
  - Connect the information obtained from various analytical techniques to understand how protein structure—from primary sequence to 3D architecture—dictates function.
  - Evaluate how conformational dynamics, ligand binding, and post-translational modifications modulate protein activity and interactions in a cellular context.

These learning goals will enable students to critically assess experimental approaches for studying proteins and to integrate biochemical, biophysical, and computational techniques to elucidate structure/function relationships in biological macromolecules.

4.5.1: Introduction

In the last chapter section, we discussed how to purify a protein (primarily through differential salt precipitation and column chromatography) and how to assess its purity (primarily through various electrophoresis methods) during purification. Now, we want to continue analyzing a "pure" protein to understand its structure and the function conferred by that structure. We need a lot of information to study protein structure/function relationships. Some are "low resolution" characteristics, such as knowing the concentration of a protein. At the highest "resolution" end, we would like to know the 3D structure of a protein with a specific ligand bound to a specific site. Our goal is then to understand proteins at varying levels of complexity or "resolution," some of which are illustrated in Figure **4.5.1**.

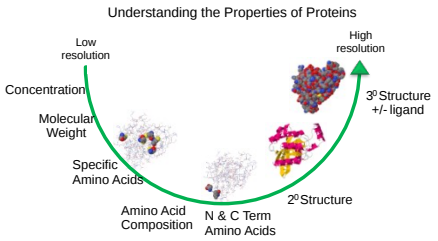


Figure **4.5.1**: Understanding proteins from low to high resolution

Various chemical and spectroscopy analysis techniques are used to achieve the specified level of structural elucidation. Spectral techniques can give us information on concentration (UV absorbance, fluorescence) and secondary structure (CD spectroscopy). Chemical analyses and mass spectroscopy give information on the amino acid composition and/or sequence. More sophisticated techniques (x-ray crystallography, cryoelectron microscopy, NMR spectroscopy) can give us 3D structural information. Each analysis shown in the figure above will be summarized below.

Sequencing the cDNA can give you some information (amino acid composition, N- and C-terminal amino acids, and the primary structure (omitted from the figure above). Even DNA sequencing won't provide information on post-translational modifications and other covalent processing (e.g., limited proteolysis, disulfide bond formation) that some of the methods below would.

In this section, we will explore some commonly used methods for protein analysis. In previous sections, we learned about the charge and chemical reactivity properties of isolated amino acids in proteins. Analyzing a whole protein is complicated because each amino acid is most often represented multiple times in the sequence. Each protein has an N-terminal and a C-terminal amino acid, as well as secondary structure. Some proteins exist biologically as multisubunit complexes, which adds to the complexity of analyses, since these proteins now have multiple N- and C-terminal ends. In addition, isolated proteins might have chemical modifications (post-translational), which add to the functionalities of the proteins and the complexities of the analyses.

Spectral techniques are widely used to provide information on protein concentration (UV absorbance, fluorescence) and secondary structure (CD spectroscopy). Chemical analyses and mass spectroscopy give information on the amino acid composition and/or sequence. More sophisticated techniques (X-ray crystallography, cryoelectron microscopy, NMR spectroscopy) can give us 3D structural information. More complete descriptions of two techniques, fluorescence spectroscopy and mass spectrometry, are presented, as their use in the analysis of biomacromolecules is underrepresented in curricula. In contrast, their use in actual laboratories becomes increasingly prevalent.

4.5.2: Low-Resolution Analyzes

4.5.2.1: Protein Concentration

There are multiple methods for determining protein concentrations in samples. Other components of a protein solution may interfere with the assays, so the choice of methods must be carefully determined.

1. Direct mass determination. A known, accurate amount of a dried protein is added to a solution of specific ionic strength and composition. The absorbance at a specific wavelength (usually 280 nm) is measured, and an extinction coefficient at that wavelength (ε1% = absorbance of a 1% protein solution = 1g protein/100 ml solution) is determined. If the protein's molecular weight is known from sequence analysis, a molar absorptivity can be determined. The concentration of the same protein in an unknown pure solution can then be determined. There are several problems with this technique. It requires relatively large amounts of protein to make accurate measurements. An even more difficult problem is that proteins bind water and counter-ions. Even a freeze-dried protein (frozen at -60 ^{0}C and placed in a vacuum, which causes sublimation of water and volatile salts in the solution) has probably 10% by weight of bound water (water of hydration).

2. ***Quantitative amino acid analysis***. The protein is hydrolyzed completely to amino acids with 6N HCl. The amino acids are then separated by high-performance liquid chromatography. As amino acids elute from the column, they are reacted with a fluorescent reagent, such as ninhydrin, fluorescamine, or orthophthaldehyde (OPA), to produce a fluorescent amino acid conjugate. The fluorescence intensity of the conjugates is proportional to the concentration of the amino acids in the protein. Before hydrolysis, a known quantity of an amino acid not present in proteins (norleucine, beta-alanine) is added, and its recovery is determined at the end of the hydrolysis and fluorescence conjugation to normalize the recovery of the other amino acids. Several problems are encountered using this technique. Incomplete peptide bond hydrolysis and partial or complete destruction of serine, threonine, tryptophan, and tyrosine occur during the acid hydrolysis. The OH-containing amino acids can be determined by quantifying them at several time points during hydrolysis and extrapolating the concentrations back to zero time. Incomplete reactions with the detecting reagent can also occur.

The most widely used, and perhaps less analytically accurate, are indirect, comparative protein assays based on the chemical properties of amide bonds or the spectrophotometric properties of the side chains Trp, Tyr, and Phe. Unknown concentrations can be determined from a standard curve derived from performing the same reactions or spectrophotometric measurements on a series of solutions of known protein concentration. Below is a discussion of each of these techniques.

3. ***Modified Lowry protein determination***. The Lowry Method is a modification of the biuret method, whose basis is described below. It is, however, much more sensitive. Biuret, as its name implies, derives from combining two urea molecules (***bi-ur-et***), as shown in the figure (panel A) below. A violet color results when copper sulfate is added to biuret in a concentrated hydroxide solution. An illustration of the copper (II) complex with biuret is shown in panel B of the figure. This biuret reaction also arises in any compound with three or more peptide bonds. Compare in Figure **4.5.2** the structure of Biuret and a polypeptide (Panel A).

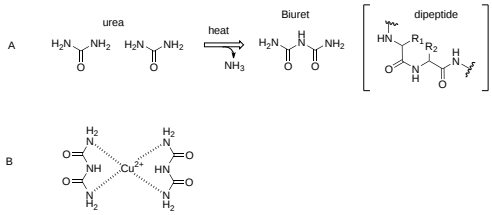


Figure **4.5.2**: The biuret reaction for protein quantitation

4. ***Dye binding assay (Bradford method)***. This method is based on the binding of the dye Coomassie Brilliant Blue G-250 to proteins, resulting in a change in the dye's absorbance. The structure of the dye is shown in Figure **4.5.3**.

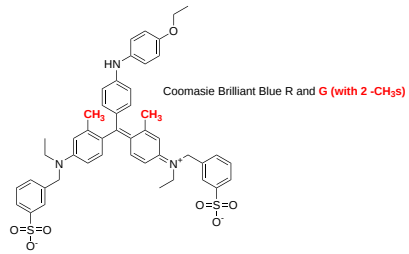


Figure **4.5.3**: Structures of Coomassie Brilliant Blue R-250 and G-250 (with 2 $-CH_3$ groups added)

The magnitude of the difference spectra at 595 nm is directly proportional to the protein concentration. The dye in the unbound free state has an absorbance maximum of 465 nm. Bradford initially developed the method (Analytic Biochemistry, *72*, 1976), which is available commercially. The dye appears to bind to proteins through hydrophobic and electrostatic interactions with a sulfonic acid group on the dye. The predominant advantage of this method is that it is cheap, simple, rapid, 3x more sensitive than the modified Lowry method, and less susceptible to interference from other compounds. The color fully develops in about 5 minutes, but decreases within 10-15 minutes when the proteins start to precipitate. Precipitation occurs more extensively at higher protein concentrations. Hence, the high-concentration standards will be affected more than the low-concentration standards.

5. ***Bicinchoninic Acid method (BCA)***. This method is based on the reduction of Cu^{++} to Cu^{+} by the peptide bond and the chelation of the Cu+ by BCA, which is monitored by absorbance at 562 nm. This method, which has been commercialized, is less prone to interference from other compounds than either the modified Lowry or the dye-binding assay. Two solutions are required: a BCA solution and a copper sulfate solution. The two are mixed to form an apple-green working solution. When protein is added, the resulting Cu^{+} chelates with two molecules of BCA, as shown in Figure **4.5.4**.

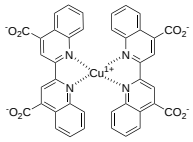


Figure **4.5.4**: BCA analysis for protein quantitation

This results in a purple color, which can be monitored spectrophotometrically at 562 nm. An absorbance at 562 nm of 0.012 per microgram of protein added to the working reagent gives this technique high sensitivity.

6. ***Absorbance A_{280} or ratios at different wavelengths***. This method is based on the fact that the three aromatic acids (Tyr, Phe, Trp) have significant absorbances in the UV. The absorption spectra of the three amino acids as a function of wavelength are shown in Figure **4.5.5**. Note the log scale on the y-axis.

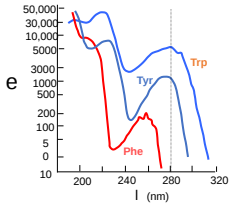


Figure **4.5.5**: Log absorbance spectra of tyrosine, phenylalanine, and tryptophan

The elution of proteins from chromatography columns is typically monitored by measuring the absorbance of the eluate at 280 nm as it leaves the column, a proxy for protein concentration. An average protein at a concentration of 1 mg/ml has an A_{280} of about 1. However, only two amino acids absorb significantly at this wavelength. Since proteins have variable numbers of these amino acids, this measurement can only estimate the concentration of an unknown protein.

The Beer-Lambert law shows that the absorbance of a chromophore in solution is given by

$$A = \epsilon l c \tag{4.5.1}$$

where A is the absorbance at a given wavelength, ε is the molar absorptivity, l is the path length of the cuvette, and c is the concentration (mol/L). Pace et al. have shown that, based on over a hundred measurements on 61 proteins in an aqueous solution, the ε(280), the molar absorptivity at 280 nm, is given by this empirical equation:

$$\epsilon(280)\left(\mathrm{M}^{-1}\ \mathrm{cm}^{-1}\right) = (\#\,\mathrm{Trp})(5,500) + (\#\mathrm{Tyr})(1,490) + (\#\text{ cystine })(125) \tag{4.5.2}$$

Proteins also absorb strongly at wavelengths less than 240 nm. This part of the absorption spectra arises from the above-mentioned amino acids and contributions from His, Met, Cys, and the peptide bond. At these wavelengths, the absorbance is less dependent on the protein's amino acid composition, but it becomes increasingly susceptible to interference from other substances. Contaminating nucleic acids, which absorb maximally at 260 nm, also contribute to the absorbance at 280 nm. Hence, the A_{280}/A_{260} ratio can be determined, and the contribution of nucleic acids can be removed through appropriate calculations. Optimal reliability is achieved by measuring A_{280}/A_{205} values since, at 205 nm, a large fraction of the absorbance due to the peptide bond. pH changes have little effect on the absorbance of the peptide bond but have a much larger effect on Tyr. At high pH, the side-chain hydroxyl is deprotonated (pKa = 10.5), with concomitant changes in A_{295}. These changes can be used to follow the titration of the Tyr residues in a protein.

4.5.2.2: Molecular Weight

Hydrodynamic techniques can estimate molecular weights, including size exclusion chromatography (under denaturing conditions using standards of known molecular weight), ultracentrifugation, and dynamic light scattering. In addition, they can be determined by polyacrylamide gel electrophoresis, again under denaturing conditions using standards. More precisely, it could be determined through protein or cDNA sequence analyses. As described below, mass spectrometry is the most accurate method for smaller proteins.

4.5.2.3: Specific Amino Acids

Aromatic amino acids can be detected by their characteristic absorbance profiles. As shown in Chapter **3.1: Amino Acids and Peptides**, amino acids with specific functional groups can be determined by chemical reactions with specific modifying groups.

4.5.2.4: Amino Acid Composition

At low resolution, we can determine the amino acid composition of the protein by hydrolyzing it in 6 N HCl at 100 °C under vacuum for various time intervals. After removing the HCl, the hydrolysate is applied to an ion-exchange or hydrophobic-interaction column, and the amino acids are eluted and quantified using known standards. A non-naturally occurring amino acid, like norleucine, is added in known amounts as an internal standard to monitor quantitative recovery during the reactions. The separated amino acids are often derivatized with ninhydrin or phenylisothiocyanate to facilitate their detection. The reaction is carried out for 24, 36, and 48 hours since amino acids with OH (like ser) are destroyed. A time course allows the concentration of Ser at time t=0 to be extrapolated. Trp is also destroyed during the process. In addition, the amide links in the side chains of Gln and Asn are hydrolyzed to form Glu and Asp, respectively.

4.5.2.5: N- and C-Terminal Amino Acid Analysis

The amino acid composition does not determine the protein's sequence. The N-terminus of the protein can be determined by reacting the protein with fluorodinitrobenzene (FDNB) or dansyl chloride, which reacts with any free amine in the protein, including the epsilon amino group of lysine. The amino group of the protein is linked to the aromatic ring of the dinitrobenzene through an amine and to the dansyl group by a sulfonamide and is hence stable to hydrolysis. The protein is hydrolyzed in 6 N HCl, and the amino acids are separated by TLC or HPLC. Two spots should result if the protein were a single chain with some Lys residues. The labeled amino acid other than Lys is the N-terminal amino acid. The C-terminal amino acid can be determined by adding carboxypeptidases, which cleave amino acids from the C-terminal. A time course must be performed to determine which amino acid is released first. N-terminal analysis can also be done as part of sequencing the entire protein using Edman degradation.

4.5.3: Primary Sequence

4.5.3.1: Protein Sequencing using Edman Degradation

Edman degradation, developed by Pehr Edman, is a method of sequencing amino acids in a peptide. In this method, the amino-terminal residue is labeled and cleaved from the peptide without disrupting the peptide bonds between other amino acid residues. The reaction is shown in Figure **4.5.6**.

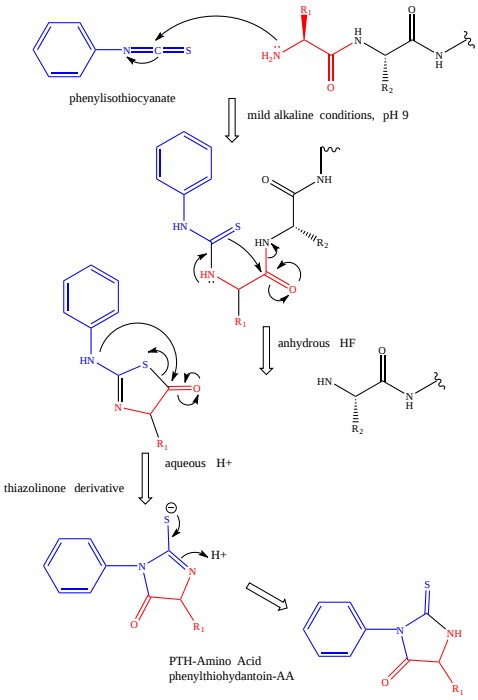


Figure **4.5.6**: Mechanism of the Edman Degradation

Phenyl isothiocyanate is reacted with an uncharged N-terminal amino group under mildly alkaline conditions to form a cyclical *phenylthiocarbamoyl* derivative. Under acidic conditions, this terminal amino acid derivative is cleaved to a thiazolinone. The thiazolinone amino acid is then selectively extracted into an organic solvent and treated with acid to form the more stable phenylthiohydantoin (PTH)- amino acid derivative that can be identified by using chromatography or electrophoresis. This procedure can then be repeated to identify the next amino acid.

A major drawback of Edman degradation is that the peptides sequenced in this manner cannot exceed 50-60 residues (and, in practice, are typically under 30). The peptide length is limited because cyclical derivatization does not always go to completion. The derivatization problem can be resolved by cleaving large peptides into smaller peptides before proceeding with the reaction. It can accurately sequence up to 30 amino acids, and modern instruments can achieve over 99% efficiency per amino acid. An advantage of the Edman degradation is that it requires only 10-100 picomoles of peptide for sequencing. The Edman degradation reaction was automated in 1967 by Edman and Beggs to speed up the process, and 100 automated devices were in use worldwide by 1973.

Because the Edman degradation proceeds from the N-terminus of the protein, it will not work if the N-terminus has been chemically modified (e.g., by acetylation or formation of pyroglutamic acid). Sequencing will stop if a non-α-amino acid is encountered (e.g., isoaspartic acid), since the favored five-membered ring intermediate cannot be formed. Edman

degradation is generally not useful in determining the positions of disulfide bridges. It also requires peptide amounts of 1 picomole or above for discernible results.

4.5.4: Secondary Structure

The percent and type of secondary structure can be determined using circular dichroism (CD) spectroscopy. In this method, right- and left-circularly polarized light illuminates a protein, which is chiral because it is composed entirely of L-amino acids. (The mirror image would be a protein of the same sequence made of D-amino acids.) Differential absorption of the right and left forms gives a CD spectrum

Circularly polarized light can be made when plane-polarized light of the same amplitude and wavelength meets out of phase by 90^0. (If they were out of phase by 180^0, they would cancel.)

- To see an animation of how circularly polarized light can be created, go to this page and select: 1. Superposition of plane-polarized waves 2.

If R and L circularly polarized light of the same wavelength and amplitude are passed through an optically inactive medium, the two waves combine (vectorially) to produce **plane-polarized light**.

- To see an animation of how circularly polarized light can be created, go to this page and select 1. Superposition of circularly polarized waves

Optical activity is observed only when the environment in which a transition occurs is asymmetric.

The peptide (amide) bond absorbs UV light in the range of 180 to 230 nm (far-UV range), so this region of the spectra gives information about the protein backbone and, more specifically, the secondary structure of the protein. The main electronic energy transitions are $n \rightarrow \pi^*$ at 220 nm and $\pi \rightarrow \pi^*$ at 190 nm for the peptide bond. There is a contribution from aromatic amino acid side chains, but it is small, given the large number of peptide bonds. The lone pair on the nitrogen adjacent to the pi bond can be considered rehybridized from sp^3 to sp^2, allowing for conjugation of the p electrons (which lowers the energy of the electrons). The Hückel diagram shown in Figure **4.5.7** below shows three molecular (not atomic) orbitals generated from the three atomic p orbitals.

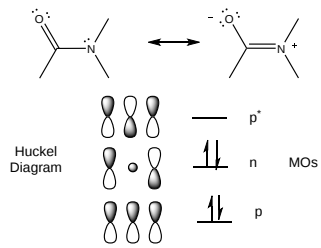


Figure **4.5.7**: Hückel molecular orbitals for the peptide bond

The middle one (with one node) has energy similar to the separate atomic p orbitals and is considered a nonbonding molecular orbital. This is consistent with the lone nonbonding pair on the nitrogen atom.

The peptide bonds in a protein's asymmetric environment will absorb light in this wavelength range (promoting electrons to higher energy levels). In different secondary structures, the peptide bond electrons absorb right- and left-circularly polarized light differently (for example, they exhibit different molar absorptivities). Hence, α, β, and random coil structures exhibit distinguishable far-UV CD spectra.

To see an animation of circularly polarized light, go to this page and select 1. Circularly Polarized Waves

Stated in another way, if plane-polarized light, which is a superposition of right and left circularly polarized light, passes through an asymmetric sample, which absorbs right and left circularly polarized light differently (i.e, they display circular dichroism), then the light passing through the sample after vector addition of the right and left circularly polarized light gives elliptically polarized light.

To see an animation of elliptically polarized light, go to this page and select 2. Plane-polarized waves in a medium with circular dichroism

If the chiral molecules also have a different index of refraction for R and L circularly polarized light, an added effect is the rotation of the angle of the elliptically polarized light. The far-UV CD spectrum of the protein is sensitive to its main chain conformation. The CD spectra of alpha and beta secondary structures are shown in Figure **4.5.8**.

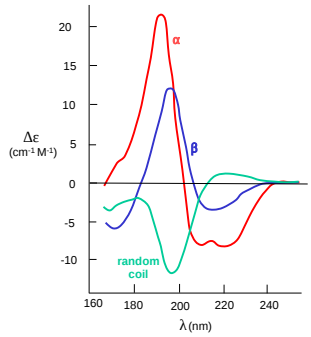


Figure **4.5.8**: CD spectra of alpha, beta, and random coil structures in proteins

Protein side chains also find themselves in such an asymmetric environment. If irradiated with circularly polarized UV light in the range of 250-300 nm (near UV), differential absorption of right and left circularly polarized light by the aromatic amino acids (Tyr, Phe, Trp) and disulfide bonds occurs, and a near UV CD spectra result. Suppose a protein's near-UV CD spectra are taken under two different sets of conditions, and the spectra differ. In that case, it can be inferred that the environment of the side chains differs, and hence the proteins have somewhat different conformations. It will not provide information about the backbone's secondary structure, as it requires lower wavelengths for absorption. Instead, it can show differences in tertiary structure.

4.5.5: Analysis of Proteins Using Fluorescence Spectroscopy

Fluorescence spectroscopy is widely used to study many aspects of protein chemistry. Although this technique is not often used in lower-level undergraduate classes, it has become so important in the study of biomolecules that a detailed explanation is necessary.

When molecules absorb energy, electrons are promoted to higher electronic energy states. This is the basis of absorption spectroscopy. These excited-state electrons can return to the ground state in processes that don't emit photons of light (ie, nonradiative processes) or radiative processes that emit light. In simple absorption spectroscopy, excited-state electrons relax to the ground state through collisions. In radiative deexcitation, light is emitted. This process of light emission is called **luminescence**, which can be divided into two categories:

- **fluorescence**: If one electron from a ground state electron pair is excited to a higher energy state, the excited electrons can still be spin paired with their ground state counterpart - i.e., they have opposite spins. The excited electron can return to the ground state without reversing its spin. (The excited state is a **singlet** state with S, the total spin state, given the formula S = 2s +1, where s = 0 (sum of +1/2 and -1/2) and S = 1 for a singlet.) This process, which results in rapid photon emission, is "spin allowed." The photon emission rate is about 108 s^{-1}, resulting in a lifetime (the average time between excitation and emission) of the excited state of about 10 ns.
- **phosphorescence**: If, in contrast to the above case, the spin of the excited electron is flipped, then its transition back to the ground state is "spin forbidden" since the excited state electron and its ground state counterpart have the same spin state. (The excited state is a **triplet** state with S, the total spin state, given by the formula S = 2s +1, where s = 1 (1/2 + 1/2) and S = 3 for a triplet. Hence, this transition occurs slowly (in the ms-s range). Toys that glow in the dark display even longer phosphorescence lifetimes. (Note: This guide will concentrate on fluorescence.)

Competing with the two deexcitation processes are nonradiative processes (such as through collisions). Given these competing processes, it might be expected that phosphorescence in liquid solutions at room temperature might not be detectable.

Molecules that fluoresce are typically aromatic, which absorb readily in the UV and visible light regions. Common fluorophores are quinine, found in tonic water (observe the faint blue glow at the surface when placed in direct sunlight), and fluorescein and rhodamine, two fluorophores often added to antifreeze. Atoms are usually nonfluorescent, except europium and terbium ions from the lanthanide series. These fluoresce when electronic transitions occur between f orbitals, which are shielded from solvent relaxation by the filled 5s and 5p orbitals.

Among biological molecules, some, especially macromolecules with aromatic groups, fluoresce. These groups are called **intrinsic fluorophores** and, in proteins, include the side chains of the aromatic amino acids tryptophan, tyrosine, and phenylalanine. The indole side chain of tryptophan is the most fluorescent, and its emission spectrum, which is sensitive to solvent conditions, is often blue-shifted when the ring is buried and red-shifted when solvent-exposed. Nucleic acids, although they also contain aromatic bases, are poor fluorophores. Many biological molecules can be made fluorescent by covalently modifying them (via nucleophilic groups on the biological molecule) with exogenously added fluorophores, such as fluorescein isothiocyanate, rhodamine isothiocyanate, or dansyl chloride. These are called **extrinsic fluorophores**. These include molecules that bind noncovalently to structures such as ds-DNA (ethidium bromide) or lipid membranes (diphenylhexatriene). Some biological fluorophores are substrates for enzyme reactions. Examples include oxidized flavins (FAD, FMN) and the reduced form of NAD (i.e., NADH). Another type of useful fluorophore is an indicator, whose fluorescent properties change with a parameter such as pH or [Ca2+].

The electronic transitions that occur during fluorescence can be represented by a Jablonski diagram as shown in the two-part Figure **4.5.9** (A and B) below.

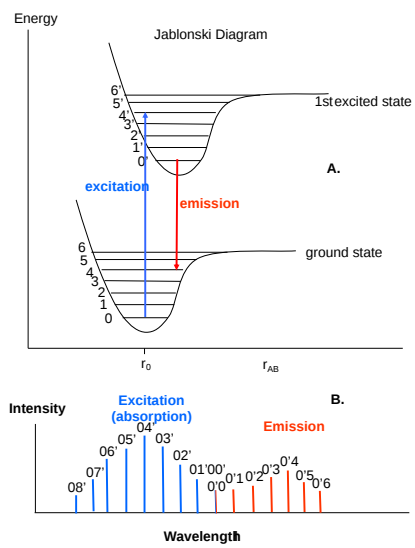


Figure **4.5.9**: Jablonski diagram and associated excitation and emission spectra. After Van Holde, et al. *Principles of Physical Biochemistry*, Pearson Prentice Hall (2006)

In panel A, the ground and first excited electronic states are shown. Each electronic state has multiple vibrational energy levels 0, 1, 2, ... and 0', 1', 2', ... This simple diagram ignores quenching of fluorescence, resonance energy transfer, etc. The transitions, represented by vertical lines, are considered to be instantaneous. They take about 10^{-15} s, so the nuclei don't move during that time. The ground-state electron is considered in the zero vibrational level, so thermal energy is insufficient to promote it to the next vibrational level. When light is absorbed, the electron is promoted to a higher vibrational level within a higher electronic level. Usually, the excited electrons relax quickly (< 1 ps) to the lowest vibrational level of S1 or S2 through **internal conversion**. Fluorescence emission may occur from the lowest vibrational state of S1 to any of the vibrational states of So. Hence, the photon emitted is lower in energy (longer in wavelength) than the absorbed photon. Also, since both processes involve the movement of an electron to different vibrational levels, with absorption or emission of a photon, and nonradiative vibrational relaxation within

those levels, the emission spectra are often the mirror image of the absorption spectra. (This assumes that the vibrational levels in So and S1 are similarly spaced. Alternatively, electrons in S1 may flip their spin and convert to the T1 state via intersystem crossing, leading to phosphorescence.

Panel B above shows blue lines for individual absorbances and red lines for emissions. Note that the excitation from 0 to 8' has the highest energy of absorption (lowest wavelength) but gives little intensity as it would occur with low frequency. If you were to draw a line over the tops of the lines in panel B, you would get simple excitation and emission spectra, which would be mirror images of each other. The emission peak is at a longer wavelength since energy was lost on vibrational, nonradiative relaxation of the excited electron. The difference in peak excitation and emission wavelengths is called the **Stokes Shift.** This shift is greatest for fluorophores in polar environments. Inferences can be made considering the disposition of a side chain (buried or surface) if changes in fluorescence properties (intensity, Stokes' shift) are noted on protein denaturation. Also, many probes are weakly fluorescent in aqueous solution but fluoresce intensely in nonpolar media (e.g., bound to a hydrophobic pocket in a protein, bilayer, or lipoprotein).

Emission spectra are usually independent of excitation wavelength (Kasha's rule). This occurs because the excited state rapidly relaxes into the lowest vibrational energy level. There are also exceptions to the mirror image rule. Deviations arise from a change in the geometry of nuclei in the excited state of the molecule. This may occur if the lifetime of the S1 state is long, allowing time for motion before emission.

An example is p-terphenyl in cyclohexane, where the rings become more coplanar in the excited state. Since there is an electron shift in the excited state, a complex between the excited fluorophore and another solution component may arise (charge-transfer complex). Alternatively, some fluorophores form dimers (e.g., pyrene) at higher concentrations. At high concentrations, changes in the emission spectra occur, arising from excited-state dimers or **excimers**. Acridine shows two emission spectra at different pHs, arising from changes in the pKa on excitation (5.45 to 10.7). Finally, exciting a fluorophore at different wavelengths (EX 1, EX 2, EX 3) does not change the emission profile. Still, it does produce variations in fluorescence emission intensity (EM 1, EM 2, EM 3) that correspond to the amplitude of the excitation spectrum.

Fluorophores can chemically modify nucleophilic side chains such as lysines and cysteines. Changes in intrinsic fluorescence in proteins can be used to measure ligand binding, conformational changes in the protein during binding interactions, changes in solution conditions, and protein denaturation. Let's explore a few fluorescence methods widely used to explore protein structure and function.

4.5.5.1: Fluorescence Quenching

Some chemical species (for example, iodide and monomeric, unpolymerized acrylamide), when added to a protein solution, can decrease the fluorescence of an intrinsic, surface-accessible fluorophore, such as the tryptophan side chain, providing information on the local environment of the intrinsic fluorophore (for example, tryptophan side-chain accessibility). For example, a buried tryptophan or probe will show little change in fluorescence intensity in the presence of a large, polar quencher. In contrast, a surface tryptophan or probe will show a significant decrease in fluorescent intensity. It is somewhat amazing that O_2, when added to a solution under increasing pressure, can quench the fluorescence of even buried tryptophan side chains, implying minimal diffusional barriers to O_2 access. This suggests significant conformational flexibility of the protein.

Quenching can be **dyn**amic, occurring when the quencher collides with the intrinsic fluorophore, or **static**, when the quencher binds to a site near the fluorophore as a prelude to quenching.

Collisional quenching is described by the **Stern-Volmer** equation.

$$\frac{F_0}{F} = 1 + k_q \tau_0 [Q] = 1 + K_D [Q] \tag{4.5.3}$$

where $\mathbf{F_o}$ and **F** are the fluorescent intensities in the absence and presence of the quencher, $\mathbf{k_q}$ is the biomolecular quenching constant, $\boldsymbol{\tau_o}$ is the lifetime of the fluorophore in the absence of the quencher, and **[Q]** is the concentration of the quencher. $\mathbf{k_q \tau_o = K_D}$ is the **Stern-Volmer quenching constant.**

A plot of F_o/F vs [Q] is linear, with a slope of K_D. $1/K_D$ is the quencher concentration at which $F_o/F = 2$, or 50% of the fluorescence intensity, is quenched. A linear plot indicates a single class of fluorophores, all of which are equally accessible to the quencher. A nonlinear plot would be found for the quenching of tryptophan fluorescence in proteins by charged or polar quenchers for proteins with more than one tryptophan, and in which some are buried. Static quenching also results in a linear SV plot. Different dependencies on temperature and viscosity can distinguish dynamic and static quenching. Since dynamic quenching depends on diffusion, higher temperatures result in higher diffusion coefficients; therefore, k_q should increase with temperature. If static quenching is involved, higher temperatures will probably reduce complex formation.

4.5.5.2: Fluorescence Resonance Energy Transfer (FRET)

If an absorbing species is near an excited-state fluorophore, and if the emission spectrum of the fluorophore overlaps the absorption spectrum of the second species, coupling of the two dipoles can occur, and energy can be transferred from the excited state of the fluorophore (donor D) to the second absorbing species (acceptor A). This energy transfer occurs through dipole coupling, **not** through the trivial emission and absorption of a photon. No photon is produced. This process is called **fluorescence resonance energy transfer (FRET).** Efficiency, E, of FRET for a single donor/acceptor pair at a fixed distance is given by:

$$\mathrm{E} = \frac{\mathrm{R}_0^6}{(\mathrm{R}_0^6 + \mathrm{r}^6)} = \frac{1}{1 + \left(\frac{\mathrm{r}}{\mathrm{R}_0}\right)^6} \tag{4.5.4}$$

R_o is the Förster distance (or radius) with a 50% transfer efficiency, and r is the distance between the donor and acceptor. R_o measures the spectral overlap of the donor and acceptor (for which most biological macromolecules have a value of 30-60 angstroms). This equation shows that efficiency depends on 1/r^6, making FRET exquisitely sensitive to distance. Figure **4.5.10** below illustrates FRET and its distance dependency.

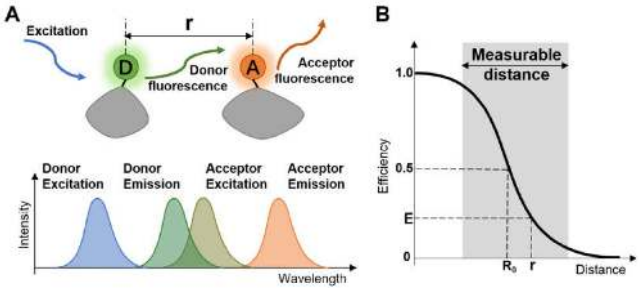


Figure **4.5.10**: Fluorescence Resonance Energy Transfer (FRET) Qiao, Y.; Luo, Y.; Long, N.; Xing, Y.; Tu, J. Single-Molecular Förster Resonance Energy Transfer Measurement on Structures and Interactions of Biomolecules. *Micromachines* **2021**, *12*, 492. https://doi.org/10.3390/mi12050492Creative Commons Attribution (CC BY) license (**https://creativecommons.org/licenses/by/4.0/**).

4.5.5.3: Anisotropy or polarization

These measure the extent of rotation of the fluorophore during its fluorescent lifetime. If a small fluorophore binds to a large molecule, its rotational diffusion constant decreases, and its anisotropy increases, as illustrated in Figure **4.5.11**.

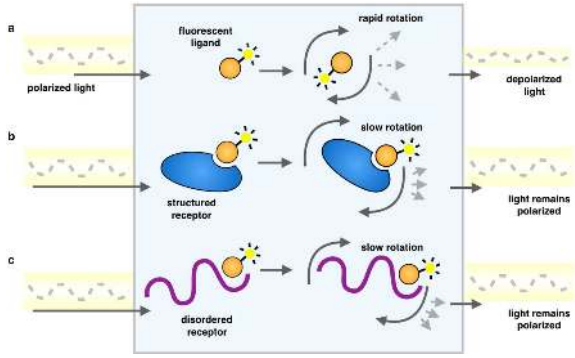


Figure **4.5.11**: Fluorescence anisotropy and rotational diffusion of a fluorophore. -2563-4Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

Schematic representation of a fluorescence polarization experiment. As a result of the rapid tumbling of molecules in solution, when a fluorescently labeled ligand is excited with plane-polarized light, the resulting emitted light is largely depolarized (a). Upon binding another species, a larger proportion of the emitted light remains in the same plane as the excitation energy because the rotation is slowed as the effective molecular size increases, whether it is an ordered molecular structure (b) or one that is disordered (c)". Since viscosity decreases rotational diffusion rates, changes in fluorescence (such as inside a bilayer) can be inferred from these measurements. For example, membranes more enriched in saturated fatty acids should show increased anisotropy of a hydrophobic, fluorescent probe compared to the same probe in a bilayer enriched in polyunsaturated fatty acids.

4.5.6: Analysis of Protein Using Mass Spectrometry

Mass spectrometry supplants traditional methods (see above) to determine a protein's molecular mass and structure. Its power comes from its exquisite sensitivity and modern computational methods for determining structure by comparing ion fragment data with databases of known protein structures. In mass spectrometry, a molecule is first ionized in an ion source. An electric field then accelerates the charged particles into a mass analyzer, where they are subjected to an external magnetic field. The external magnetic field interacts with the magnetic field generated by the charged particles' motion, deflecting them. The deflection is proportional to the mass-to-charge ratio, m/z. Ions then enter the detector, which is usually a photomultiplier. Sample introduction into the ion source occurs through simple diffusion of gases and volatile liquids from a reservoir, by injecting a liquid sample containing the analyte via a fine mist, or, for very large proteins, by desorbing a protein from a matrix using a laser. Complex mixtures are analyzed by coupling HPLC to mass spectrometry in an LC-MS system.

There are many methods for ionizing molecules, including atmospheric pressure chemical ionization (APCI), chemical ionization (CI), and electron impact (EI). The most common methods for protein/peptide analyses are electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI).

4.5.6.0.1: Electrospray ionization (ESI)

The analyte, dissolved in a volatile solvent such as methanol or acetonitrile, is injected into the ion source through a fine stainless steel capillary at a low flow rate. A high voltage (3-4 kV) is applied to the capillary, giving it a positive charge relative to the oppositely charged electrode. The flowing liquid becomes charged with the same polarity as the polarity of the positively charged capillary. The high field leads to the emergence of the sample as a charged aerosol spray of charged microdrops, thereby reducing electrostatic repulsions within the liquid. This method uses electrical energy to produce the aerosol, rather than mechanical energy to produce a liquid aerosol, as in a perfume atomizer. Surrounding the capillary is a flowing gas (nitrogen) that helps move the aerosol toward the mass analyzer. As the volatile solvent evaporates, the microdrops become smaller, increasing the positive charge density on them. Eventually, electrostatic repulsions cause the drops to explode in a series of steps, ultimately producing the analyte devoid of solvent. This gentle ionization method produces analytes that are not cleaved but ready for introduction into the mass analyzer.

Proteins emerge from this process with a roughly Gaussian distribution of positive charges on basic side chains. You studied mass spectra of small molecules induced by electron bombardment in organic chemistry. This produces ions with a +1 charge as an electron is stripped from the neutral molecule. The highest m/z peak in the spectrum corresponds to the parent ion, M^+. The highest m/z ratio detectable in the mass spectrum is in the thousands. However, large peptides and proteins with high molecular mass can be detected and resolved because their ions carry a charge greater than +1. In 2002, John Fenn was awarded a Nobel Prize in Chemistry for developing and using ESI to study biological molecules.

Figure **4.5.12** shows an example of an ESI spectrum of apo-myoglobin.

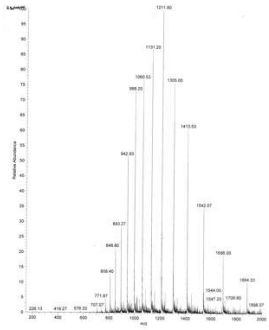


Figure 4.5.12: ESI mass spectrum of apo-myoglobin

Note the roughly Gaussian distribution of the peaks, each representing the intact protein with charges differing by +1. Proteins acquire positive charges by protonating amino acid side chains and by charges induced during electrospray. Based on the amino acid sequence of myoglobin and the assumption that the pKas of the side chains are the same in the protein as for isolated amino acids, the calculated average net charges of apomyoglobin (apoMb) would be approximately +30 at pH 3.5, +20 at pH 4.5, +9 at pH 6, and 0 at pH 7.8 (the calculated pI). The mass spectrum below was obtained by direct injection of apoMb into the MS in 0.1% formic acid (pH 2.8)—charges on the peptide result from those present before the electrospray and those induced during the process.

The protein's molecular mass can be determined by analyzing two adjacent peaks, as shown in Figure 4.5.13.

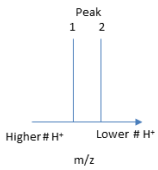


Figure 4.5.13: m/z ratio for two adjacent peaks in the mass spectrum of a protein

Suppose M is the molecular mass of the analyte protein, and n is the number of positive charges on the protein represented in a given m/z peak. In that case, the following equations give the molecular mass M of the protein for each peak:

$$\begin{aligned} M_{\text{peak 2}} &= n(m/z)_{\text{peak 2}} - n(1.008) \\ M_{\text{peak 1}} &= (n+1)(m/z)_{\text{peak 1}} - (n+1)(1.008) \end{aligned} \tag{4.5.5}$$

where 1.008 is the atomic weight of H. Since there is only one value of M, the two equations can be set equal to each other, giving:

$$n(m/z)_{\text{peak 2}} - n(1.008) = (n+1)(m/z)_{\text{peak 1}} - (n+1)(1.008) \tag{4.5.6}$$

Solving for n gives:

$$n = \frac{(m/z)_{\text{peak1}} - (1.008)}{(m/z)_{\text{peak2}} - (m/z)_{\text{peak1}}} \tag{4.5.7}$$

Knowing n, the molecular mass M of the protein can be calculated for each m/z peak. The best value of M can then be determined by averaging the M values determined from each peak (16,956 from the above figure). For peaks from m/z of 893-1542, the calculated values of n ranged from +18 to +10.

4.5.6.0.1: Matrix-assisted laser desorption ionization (MALDI)

This technique for larger biomolecules (proteins and polysaccharides) involves mixing the analyte with an absorbing matrix. Laser excitation of the matrix leads to energy transfer, resulting in ionization and the "launching" of the matrix and analyte in ion form from the solid mixture. Parent ion peaks of $(M+H)^+$ and $(M-H)^-$ are formed.

The Mass Analyzer separates the created ions based on m/z ratios. Let's consider two: the quadrupole ion trap and the time-of-flight. There are several general types of mass analyzers, including **magnetic sector, time-of-flight, quadrupole, and ion trap.**

Quadrupole ion trap (used in ESI) - A complex mixture of ions can be contained (or trapped) in this mass analyzer. Two common types are linear and 3D quadrupole. If a dipole has two poles (+ and -) separated by some distance, then a quadrupole has four poles (+, -, +, and -) arranged geometrically such that each + has a - on each side and vice versa. Figure 4.5.14 below shows linear and 3D quadrupoles.

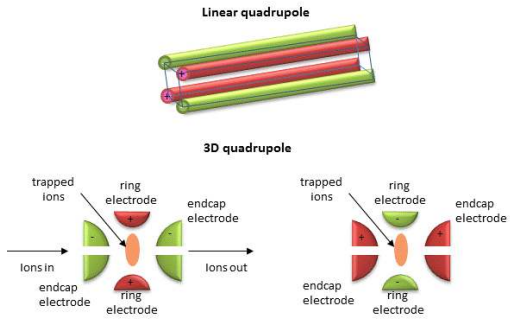


Figure 4.5.14: Linear and 3D quadrupoles

As dipoles display positive and negative charge separation on a linear axis, quadrupoles have either opposite electrical charges or opposite magnetic fields at the opposing ends of a square or cube. In charge separation, the monopole (sum of the charges) and dipoles cancel to zero, but the quadrupole moment does not. The quadrupole traps ions using a combination

of fixed and alternating electric fields. The trap contains helium at 1 mTorr. For the 3D trap, the ring electrode has an oscillating RF voltage, which keeps the ions trapped. The end caps also have an AC voltage. Ions oscillate in the trap at a "secular" frequency determined by the RF voltage frequency and, of course, the m/z ratio. By increasing the amplitude of the RF field, the ion motion in the trap becomes destabilized, leading to ion ejection into the detector. When the secular frequency of ion motion matches the applied AC voltage across the end cap electrodes, resonance occurs, increasing the ions' motion amplitude and allowing leakage out of the ion trap into the detector.

Tandem Mass Spectrometry (MS/MS): Quadrupole mass analyzers, which can select ions of varying m/z ratios in the ion traps, are commonly used in tandem mass spectrometry (MS/MS). In this technique, the selected ions are further fragmented into smaller ions by collision-induced dissociation (CID). When performed on all of the initial ions present in the ion trap, the sequence of a peptide/protein can be determined. This technique usually requires two mass analyzers with a collision cell in between, where selected ions are fragmented by collision with an inert gas. Using a quadrupole ion trap, it can also be done in a single mass analyzer.

In time-of-flight (TOF) mass spectrometry (used in MALDI), a long tube determines the time required for ion detection. The small molecular mass ions reach the detector the fastest.

4.5.6.1: Sequence Determination Using Mass Spectrometry

In a typical MS/MS experiment to determine a protein sequence, a protein is cleaved into protein fragments using an enzyme such as trypsin, which cleaves on the carboxyl side of Lys and Arg residues. The average protein size in the human proteome is approximately 50,000. If the average molecular mass of an amino acid in a protein is around 110 (18 subtracted since water is released on amide bond formation), the average number of amino acids in the protein would be around 454. If 10% of the amino acids are Arg and Lys, then on average, there would be approximately 50 Lys and Arg, and hence 50 tryptic peptides of average molecular mass 1000. The fragments are introduced in the MS, where a peptide fragment fingerprint analysis can be performed. The molecular weights of the fragments can be determined and compared with those of known peptide digestion products from known proteins to identify the analyte protein.

Ions with the original N terminus are denoted as a, b, and c, while ions with the original C terminus are denoted as x, y, and z. c and y ions gain an extra proton from the peptide to form positively charged -NH_3^+ groups. Figure **4.5.15** below shows peaks for a 4-amino acid peptide fragmentation pattern

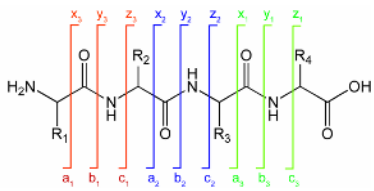


Figure **4.5.15**: 4-amino acid peptide fragmentation peaks. https://commons.wikimedia.org/wiki/F...gmentation.gif. Creative Commons Attribution-Share Alike 3.0 Unported

Ions with the original N terminus are denoted as a, b, and c, while ions with the original C terminus are denoted as x, y, and z. c and y ions gain an extra proton from the peptide to form positively charged -NH_3^+ groups. The ions observed depend on many factors, including the peptide sequence, its original charge, the collision energy used to induce fragmentation, etc. Low energy fragmentation of peptides in ion traps usually produces a, b, and y ions, along with peaks resulting from loss of NH_3 (a*, b*, and y*) or H_2O (ao, bo, and yo). No peaks resulting from the fragmentation of side chains are observed. Fragmentation at two sites in the peptide (usually at the b and y sites in the backbone) forms an internal fragment.

The y1 peak represents the tryptic peptide's C-terminal Lys or Arg (in this example). Peak y2 has one more amino acid than y1, and the molecular mass difference identifies the extra amino acid. Peak y3 is likewise one amino acid larger than y2. All three y fragment peaks have a common Lys/Arg C-terminal, and the charged fragment contains the C-terminal end of the original peptide. All b fragment peaks for a given peptide contain a common N-terminal amino acid, with b1 the smallest. Note that the subscript represents the number of amino acids in the fragment. By identifying b and y peaks, the actual sequence of small peptides can be determined. Usually, spectra are matched to databases to identify the structure of each peptide and, ultimately, that of the protein. The actual m values for fragments can be calculated as follows, where (N is the molecular mass of the neutral N terminal group, (C) is the molecular mass of the neutral c terminal group, and (M) is the molecule mass of the neutral amino acids. (For N-terminal amino acid, add 1 H. For C terminus, add OH.)

- a: (N)+(M)-CHO
- b: (N)+(M)
- y: (C)+(M)+H (note in the figure above that the amino terminus of the y peptides has an extra proton in the +1 charged peptides.)

m/z values can be calculated from the m values by adding the mass of the extra proton to the overall z, if the overall charge is +1, etc.

For example, from these MW values, the human Glu1- fibrinopeptide B sequence can be determined from MS/MS spectra shown in an annotated form in Figure **4.5.16**. Note that most of the b peaks are b*, resulting from the loss of NH_3 from the N terminus.

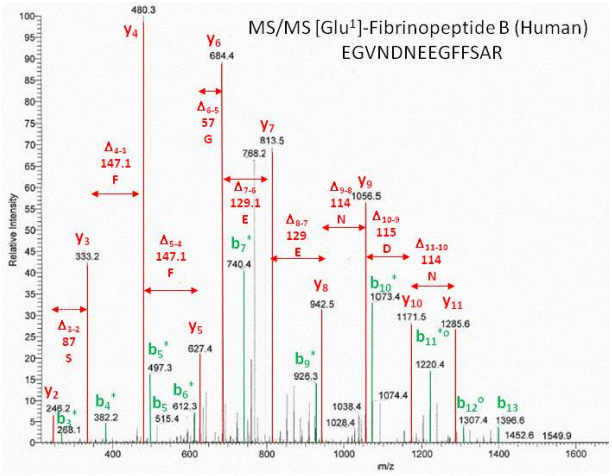


Figure **4.5.16**: Annotated MS/MS spectra of human Glu1- fibrinopeptide B

Now, let's step back and get a broader picture of structure analysis by mass spectrometry. In general, proteins are analyzed either in a "top-down" approach, in which they are analyzed intact, or in a "bottom-up" approach, in which they are first digested into fragments. An intermediate "middle-down" approach, in which larger peptide fragments are analyzed, may also

be used. However, the top-down approach is mostly limited to low-throughput single-protein studies due to issues in handling whole proteins, their heterogeneity, and the complexity of their analyses.

In the second approach, referred to as the "bottom-up" MS, proteins are enzymatically digested into smaller peptides using a protease such as trypsin, which cleaves peptide chains mainly at the carboxyl side of lysine or arginine, except when either is followed by proline. It is used for numerous biotechnological processes. The process is commonly referred to as trypsin proteolysis or trypsinization, and proteins that have been digested/treated with trypsin are said to have been trypsinized.

Subsequently, these peptides are introduced into the mass spectrometer and identified by peptide mass fingerprinting or tandem mass spectrometry. Hence, this approach uses peptide-level identification to infer the presence of proteins pieced back together via *de novo* repeat detection. The smaller and more uniform fragments are easier to analyze than intact proteins and can be also determined with high accuracy, this "bottom-up" approach is therefore the preferred method of studies in proteomic studies. A further approach that is beginning to be useful is the intermediate "middle-down" approach, in which proteolytic peptides larger than the typical tryptic peptides are analyzed.

Proteins of interest are usually part of a complex mixture of multiple proteins and molecules, which coexist in the biological medium. This presents two significant problems. First, the two ionization techniques used for large molecules are effective only when the mixture contains roughly equal amounts of the constituents. At the same time, different proteins tend to be present in widely differing amounts in biological samples. If such a mixture is ionized using electrospray or MALDI, the more abundant species tend to "drown" or suppress signals from less abundant ones. Second, the mass spectrum from a complex mixture is very difficult to interpret due to the overwhelming number of mixture components. This is exacerbated by the fact that a protein's enzymatic digestion yields multiple peptide products.

In light of these problems, methods such as one- and two-dimensional gel electrophoresis and high-performance liquid chromatography are widely used to separate proteins. The first method fractionates whole proteins via two-dimensional gel electrophoresis (Figure 3.31). The first dimension of 2D gel electrophoresis is isoelectric focusing (IEF). In this dimension, the protein is separated by its isoelectric point (pI), and the second dimension is SDS-polyacrylamide gel electrophoresis (SDS-PAGE). This dimension separates proteins based on their molecular weight. Once this step is complete, in-gel digestion proceeds.

In some situations, combining both of these techniques may be necessary. Gel spots identified on a 2D Gel are usually attributable to one protein. If protein identity is desired, the in-gel digestion method is usually used, in which the protein spot of interest is excised and proteolytically digested. The peptide masses resulting from the digestion can be determined by mass spectrometry using peptide mass fingerprinting. If this information does not allow unequivocal protein identification, its peptides can be subject to tandem mass spectrometry for *de novo* sequencing. Small changes in mass and charge can be detected with 2D-PAGE. The disadvantages of this technique are its small dynamic range compared to other methods. Some proteins are still difficult to separate due to their acidity, basicity, hydrophobicity, and size (too large or too small).

The second method, high-performance liquid chromatography (HPLC/MS), is used to fractionate peptides after enzymatic digestion. Characterization of protein mixtures using HPLC/MS is also called shotgun proteomics and MuDPIT (Multi-Dimensional Protein Identification Technology). One or two steps of liquid chromatography fractionate a peptide mixture that results from the digestion of a protein mixture. The chromatographic eluant can be directly introduced into the mass spectrometer via electrospray ionization or deposited onto a series of small spots for later MALDI mass analysis.

Figure **4.5.17** shows the general schema for analyzing proteins by mass spectrometry.

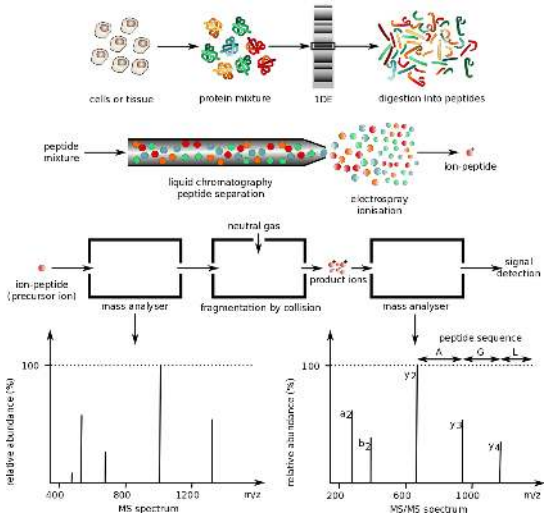


Figure **4.5.17**: Schematic of Protein Fingerprinting by Mass Spectrometry. Philippe Hupé - Emmanuel Barillot, Laurence Calzone, Philippe Hupé, Jean-Philippe Vert, Andrei Zinovyev, Computational Systems Biology of Cancer Chapman & Hall/CRC Mathematical & Computational Biology, 2012

Protein mixtures are prepared from cell culture or tissue samples and separated by gel electrophoresis. Single proteins are isolated and digested using trypsin to produce a peptide mixture. Peptides are separated by liquid chromatography and analyzed by mass spectrometry

4.5.7: 3D Structural Determination

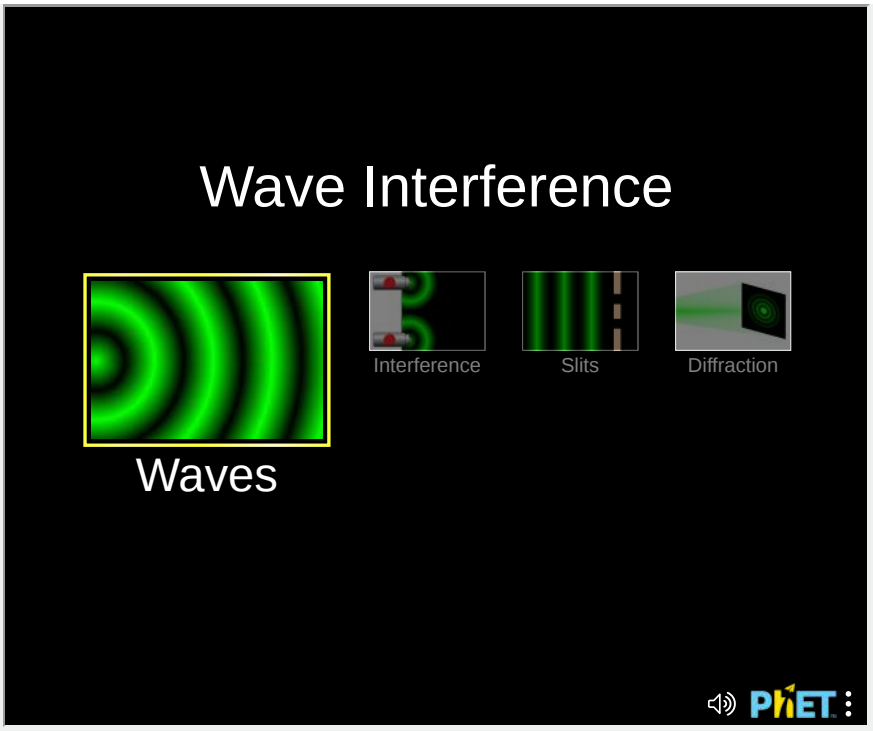
The 3D structure of a protein can be determined in four main ways: X-ray crystallography, multidimensional NMR, cryoelectron microscopy, and computer modeling using artificial intelligence and machine learning.

4.5.7.1: *X-Ray Crystallography*

In this technique, proteins are induced to form solid crystals in which the individual molecules pack in a well-defined crystal lattice. X-rays are aimed at the crystals. The X-rays are scattered off the crystal and collected by a detector. The scattered X-rays undergo constructive and destructive interference (as water waves do), forming a diffraction pattern. The crystal's spacing and types of atoms determine the diffraction pattern. Hence, from a given 3D structure, a specific diffraction pattern is formed. Using sophisticated mathematics (Fourier Transformations), the diffraction pattern can be converted back into the 3D structure of the object —in this case, the protein's atoms.

PHET Simulation - Diffraction

Constructive and destructive interference and the formation of a "diffraction" pattern can be readily visualized in two PHET simulations. Follow the instructions below.



Simulation 1: Slits

- Select Slits to open the simulation and select these choices in order:
- Type of wave: pick one that looks like a bullet
- Frequency (middle green)
- Amplitude: max
- Check Screen
- Choose 2 Slits
- Slit width 200
- Slit separation 400
- Click the Green button.

You will see light/dark patterns moving toward the screen. The light zones arise from constructive interference and dark from destructive interference of the two waves as they emerge from the slits.

Simulation 2: Diffraction

Now, look at the diffraction pattern from light moving through simple and more complex openings and interacting with an object using the PHET animation and the step below.

- First, refresh the browser window.
- Select Diffraction to open the simulation and select these choices in order.
- Pick 450 nm wavelength.
- Choose in succession the four vertical icons (circle, square, circle/square, array of circles, person).
- Observe the diffraction patterns as you change the slit size.

X-ray scattering can also be viewed as light "reflecting" from a series of planes formed by atoms in the crystal, with the planes separated by specific distances (in the Angstrom range). The x-rays that are "reflected" from innumerable planes recombine constructively and destructively to form a diffraction pattern. X rays are used since the size of the "slits", and the distance between these "reflective" planes, must be comparable to the wavelength of the incident light, which for x-rays is 0.5 – 2.5 Å.

A diffraction pattern is mathematically decoded to form an electron density map, since electrons scatter X-rays. Hydrogen atoms don't appear in X-ray crystal structures because they lack sufficient electrons to serve as effective scattering centers. Computer programs are used to fit the electron density map to a 3D arrangement of atoms separated by characteristic bond distances corresponding to the functional groups and side chains found in proteins. The quality/amount of crystals helps determine the quality of the diffraction pattern and the resulting structure. X-ray crystallographers define quality in terms of resolution. A resolution of 5 Å - 10 Å can reveal the structure of polypeptide chains, 3 Å - 4 Å of groups of atoms, and 1 Å - 1.5 Å of individual atoms.

Figure **4.5.18** shows the process from crystal to model.

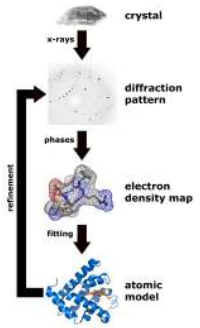


Figure 4.5.18: Overview of structure determination by mass spectrometry. Thomas Splettstoesser (www.scistyle.com), CC BY-SA 3.0 <https://creativecommons.org/licenses/by-sa/3.0>, via Wikimedia Commons

Figure 4.5.19 below shows the electron density map around tyrosine 103 from myoglobin from crystal structures at two different resolutions (left, 1a6m, 1.0 Å resolution and right, 108m, 2.7 Å resolution).

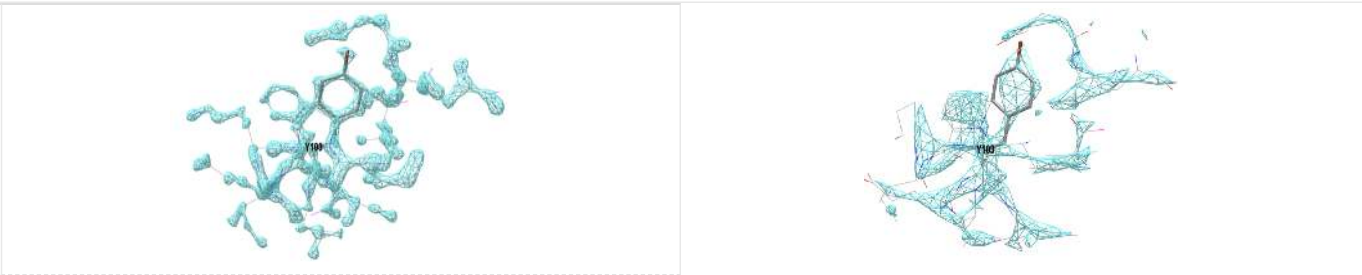


Figure 4.5.19: Electron density map around tyrosine 103 from myoglobin from crystal structures at two different resolutions (left, 1a6m, 1.0 Å resolution and right, 108m, 2.7 Å resolution). Click on the images to see full iCn3D models showing the electron density map around the Y103. (Choice of Y103 from https://pdb101.rcsb.org/learn/guide-...ata/resolution)

This figure shows 2Fo-Fc electron density maps, which use the observed diffraction data (Fo) and the diffraction data calculated from the atomic model (Fc). Proteopedia has an excellent description of electron density maps.

Not all proteins can be readily crystallized. The process is, in many ways, an art as much as it is a science. Membrane proteins fall into this category.

4.5.7.2: Nuclear Magnetic Resonance (NMR)

Many readers have probably performed ^{1}H-NMR on small molecules in introductory and organic chemistry labs. Interpreting spectra of molecules with many hydrogen atoms in straight and branched chains, in rings, and in functional groups is not simple. Imagine doing that to determine the structure of a small protein with 1000s of hydrogen atoms! The spectrum would be essentially indecipherable. Luckily, **multi-dimensional NMR** techniques have allowed the solution (not crystal) structure of small proteins to be determined. These methods are outside of the scope of this book. For those interested in more detail, read A Brief Introduction to NMR Spectroscopy of Proteins by Poulsen.

Let's briefly introduce a 2D NMR peak for a simple molecule, ethyl acetate. The 1D ^{1}H-NMR spectrum for the molecule is shown in Figure 4.5.20.

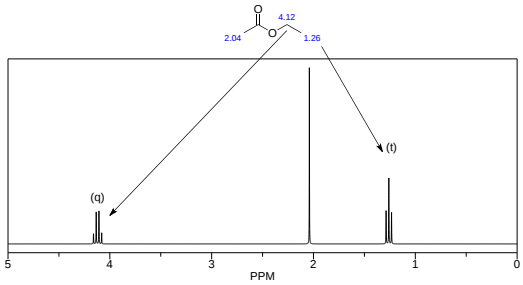


Figure 4.5.20: 1D ^{1}H-NMR spectrum of ethylacetate

Now, let's show a simulated 2D COSY spectrum of the same molecule. The image and explanation below are adapted from Structure & Reactivity in Organic, Biological, and Inorganic Chemistry by Chris Schaller, licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. Structure and Reactivity

In homonuclear correlation spectroscopy (COSY), we can look for hydrogens that are coupled to each other. In ethyl acetate, it's pretty clear where they are. There is a quartet and a triplet; the hydrogens corresponding to those two peaks are probably beside each other in the structure. The COSY spectrum takes that ^{1}H spectrum and spreads it into two dimensions. Instead of being displayed as a row of peaks, the peaks are spread out into a 2D array. Figure 4.5.21 shows an annotated simulated COSY spectrum. The peaks are displayed along one axis, and the same peaks are displayed along the other.

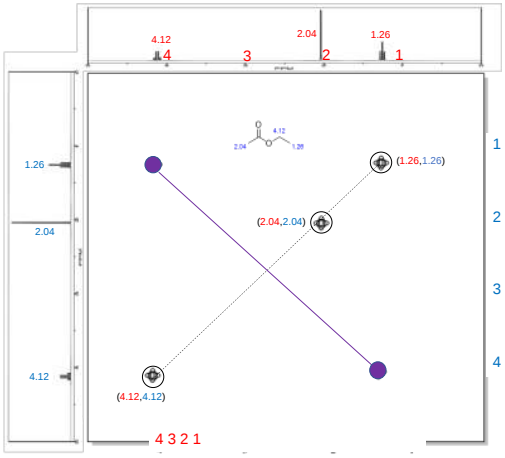


Figure **4.5.21**: Annotated simulated COSY spectrum of ethylacetate

What does it mean to be coupled? It means that magnetic information is transmitted between the atoms. How can we tell? Essentially, we can send a pulse of electromagnetic radiation into one set of hydrogen atoms and look for a response elsewhere. Of course, if we send a pulse of radio waves at a frequency that a particular hydrogen will absorb, we will see a response in that hydrogen itself. That's why we see the peaks on the diagonal (dotted line). The peaks along the diagonal in the spectra (1.26, 1.26; 2.04, 2.04; 4.12, 4.12) don't provide any new information, since they are the main peaks in the 1D NMR.

However, we also see responses from other hydrogen atoms magnetically linked to the original one. They give the peaks (shown as purple circles) that do **not** appear along the diagonal. Those peaks indicate which hydrogens are coupled to which other hydrogens. The hydrogens at 1.26 ppm are coupled to the ones at 4.12 ppm, giving a "**cross-peak**" at (1.26, 4.12). There is also a cross-peak at (4.12, 1.26) because that relationship goes both ways.

This coupling should make sense because the protons that give the signals at 4.12 (-O-CH_2-CH_3) and 1.26 (-CH_2-CH_3) are on adjacent carbons and split each other's signals, as seen in the 1D NMR. This example shows how we can identify which protons in a 2D NMR spectrum are coupled through 3 bonds (H-C-C-H), a critical piece of information for determining protein structures by NMR.

Variants of 2D NMR are also used in protein structure. They use NMR-active nuclei in addition to ^{1}H, including ^{13}C (natural abundance 1%) and ^{15}N (natural abundance 0.37%). Given these low abundances, proteins for NMR structure determination are often purified in cells grown in media enriched in ^{13}C and ^{15}N precursors.

HMBC (Heteronuclear Multiple Bond Correlation) and HMQC (Heteronuclear Multiple Quantum Coherence): Just as COSY spectra show which protons are coupled to each other, HMBC (and the related HMQC) give information about the relative relationships between protons and carbons in a structure. In an HMQC spectrum, a ^{13}C spectrum is displayed on one axis and a ^{1}H spectrum on the other. Cross-peaks show which proton is attached to which carbon. COSY spectra show a 3-bond coupling (from H-C-C-H), whereas HMQC shows a 1-bond coupling (just C-H).

Nuclear Overhauser Effect (NOE) Spectroscopy (NOESY): This technique shows through-space interactions within the molecule rather than the through-bond interactions in COSY and HMBC/HMBQ. This method is especially useful for determining a molecule's stereochemical relationships. In two stereoisomers, all atoms are connected in the same order by the same bonds. A COSY or an HMBC spectrum couldn't distinguish between these isomers.

HNCA: This is an example of 3D NMR. It shows a correlation between an amide proton, the amide nitrogen to which it is attached, and the carbons attached to the amide nitrogen. HNCA data are viewed in slices, where you examine one nitrogen at a time. One axis shows the shift of the proton attached to that nitrogen, and the other axis shows the shifts of the carbons attached to the nitrogen. The abbreviation HNCA comes from the pathway for transferring magnetization (amide **H** to amide **N**, then to the attached **C-a**lpha).

NMR structures in the Protein Data Bank consist of many slightly different structures. This arises from the dynamic behavior of the proteins in water, compared to when the structure is determined from a crystal. Comparative analyses between crystal and NMR structures show that secondary structures are equally accurate, that loops in NMR structures are probably too flexible, and that loops in protein, often on the surface, are too rigid, which makes sense given the packing restraints with a crystal lattice.

4.5.7.3: Cryo-electron Microscopy

Cryogenic-electron microscopy (cryo-EM) has recently emerged as a powerful structural biology technique that delivers high-resolution density maps of macromolecular structures. Figure **4.5.22** shows a cryo-EM and structure determined from it.

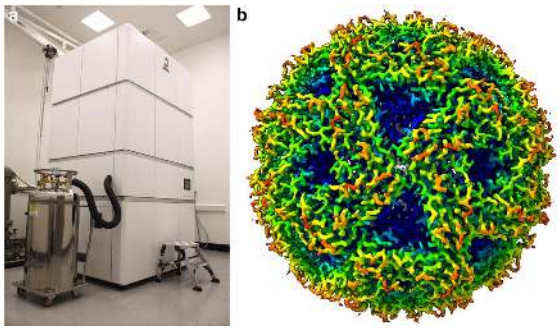


Figure **4.5.22**: **Cryo-Electron Microscopy.** (a) the Scottish Centrel for Macromolecular Imaging JEOL CryoARM 300. (b) High-resolution 2.2-Å resolution structure of lumazine synthase. *Bhella, D. (2019) Biophysical Reviews 11:515-519*. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/)

Resolutions approaching 1.5 Å are now possible, and maps in the 1–4-Å range inform the construction of atomistic models with high confidence. This new capacity for investigators to determine macromolecular structures at high resolution and without the need to form crystals has led to an explosion of interest in adopting cryo-EM.

Protein suspensions are frozen on 3-mm-diameter transmission-electron microscope (TEM) support grids made from a conductive material (e.g., Cu or Au) that are coated with a carbon film with a regular array of perforations 1–2 μm in diameter. A total of 3–5 μl of sample is loaded onto the grid, which is then immediately blotted with filter paper to create a film of

buffer/protein on the grid that, when frozen, will be thin enough for the electron beam to penetrate. Optimizing ice thickness is a vital step in sample preparation, as thicker ice layers increase the likelihood that an incident electron will undergo multiple scattering, thereby reducing image quality. In the case of extreme ice thickness, the electron beam does not penetrate at all. After blotting, the grid is rapidly plunged into a bath of liquid ethane—a very effective cryogen that freezes water rapidly enough to prevent ice crystal formation. Forming a vitreous ice layer is the fundamental step in cryo-EM and preserves the target in a near-native state. The resulting vitreous ice layer with suspended protein molecules must then remain close to liquid nitrogen temperature (− 196 °C) during storage and TEM imaging to prevent phase changes to other types of ice that are not amenable to high-quality imaging and preservation of protein structure.

Figure **4.5.23** summarizes the process and single particle structure determination.

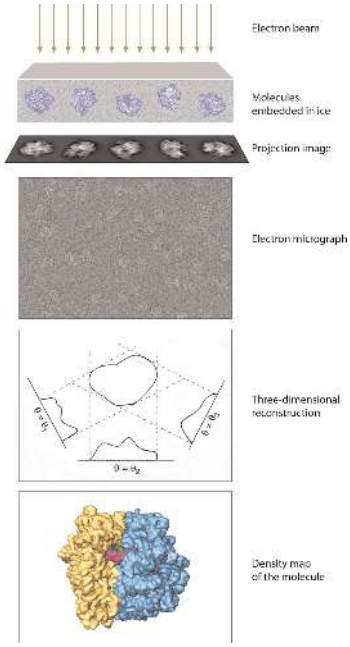


Figure **4.5.23**: Principle of cryo-EM and single-particle reconstruction. Agirrezabala, X., Frank, J., 2010. From DNA to proteins via the ribosome: Structural insights into the workings of the translation machinery. Human Genomics 4, 226. https://doi.org/10.1186/1479-7364-4-4-226. CC BY 4.0

The figure above shows the cryoEM structure-determination process for ribosomes. When frozen, they are found in random orientations embedded in a thin layer of ice. Exposure to a low-dose electron beam in the transmission electron microscope produces a projection image (i.e., the electron micrograph). A typical electron micrograph shows E. coli ribosomes as low-contrast single particles on a noisy background. After the orientations of the particles have been determined, usually by matching them to a reference using computer algorithms, they are used to reconstruct a density map using back-projection or a similar reconstruction algorithm. This density map is segmented into components (subunits, ligands), and these components are displayed in different colors in a surface representation (bottom panel; small and large subunits are shown in yellow and blue, respectively). A- and P-site tRNAs are colored red, pink, and green, respectively.

Given the increasing popularity of this technique, we present Figure **4.5.24** below, which shows another representation of the end-point of a higher-resolution 3D model.

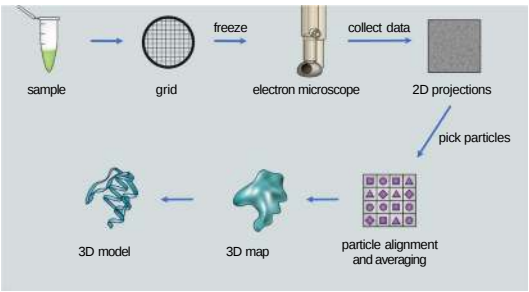


Figure **4.5.24**: A schematic of the single-particle reconstruction cryoEM pipeline. Hey Tony et al., 2020. Machine learning and big scientific data. *Phil. Trans. R. Soc. A*.**378**2019005420190054. http://doi.org/10.1098/rsta.2019.0054. Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/

Motion occurs when the beam strikes the ice dome, where the single structure is found. Motion pictures are captured with fast detectors, and computers are used to compensate for motion, which reduces the resolution of the structure. In addition, as X-rays damage molecules, so can electron beams. Earlier frames in the movie show less damage. The beam's motion and damage effects can now be corrected to increase the resolution.

Here is a YouTube video from PDB 101 that describes the technique.

Methods for determining atomic structure. PDB-101: Educational resources supporting molecular explorations through biology and medicine. Christine Zardecki, Shuchismita Dutta, David S. Goodsell, Robert Lowe, Maria Voigt, Stephen K. Burley. (2022) *Protein Science* **31**: 129-140 https://doi.org/10.1002/pro.4200. CC By 4.0 license.

Prediction of Protein Structure and Protein Function

A new chapter section, 4.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24), has been created to reorganize material on predicting protein structure from sequence and the reverse problem: predicting sequence from protein structure (for designer proteins) and from protein function.

4.5.8: Comparisons of 3D Structure Determination Methods

Obtaining a pure, highly concentrated (mM) protein sample is a major bottleneck for X-ray crystallography and NMR. The high concentration is required because both techniques are insensitive to single-molecule analysis, and a large population of a particular protein is required to overcome the signal-to-noise barrier. Similarly, the sample needs to be very homogeneous, so protein purification is necessary at some point. Cryo-EM requires considerably less protein than the other two methods, but still a lot by any standard. Typically, Cryo-EM requires preparations at a concentration of 1 mg/ml in a volume of at least 50 µl. In contrast, crystal formation might require 500 µl of protein at 5 -10 mg/ml concentration. Cryo-Em requires the protein to be prepared in a low-salt buffer solution with minimal additives to ensure good freezing and image contrast.

Recombinant protein production using *E. coli* is the method of choice when large quantities of protein are required. This process involves cloning the gene (often cDNA) encoding the protein of interest into a suitable inducible vector, transforming the vector into an *E. coli* host, and growing the culture in a rich medium. The bacterial host will multiply during a growth phase, after which it is induced to express the protein of interest. If all goes well, the protein will be soluble and expressed at high levels. Unfortunately, this process is easier said than done. Many eukaryotic proteins do not express well in prokaryotic hosts, and modifications are needed to optimize the bacterial host, codon usage, media, etc., to obtain a decent yield of recombinant protein.

Additionally, expressed proteins are often found as insoluble inclusion bodies and require high concentrations (2M to 8M) of denaturants such as urea or guanidine hydrochloride to solubilize them, followed by stepwise dialysis into an appropriate buffer to refold them. Alternatively, eukaryotic organisms such as *S. cerevisiae* (yeast), insect, and mammalian cell lines can be used, especially if post-translational modifications are studied. However, yield decreases and overall costs increase are common with these organisms.

Protein production is difficult for ^{15}N- and ^{13}C- NMR because proteins must be labeled with these isotopes. Only these isotopes have nuclei with + ½ and -½ spin states, which enable the energy transitions required for a radiofrequency NMR signal. Of course, ^{1}H also has ½ spin states but is highly abundant.

Protein stability is an issue for both crystallography and NMR. Once a protein has been expressed, purified, and concentrated, it must maintain its structural integrity for the experiments. In crystallography, a protein sample is placed in various solutions (often high in polyethylene glycol) that promote crystallization. Crystallization conditions are tested in a high-throughput 96-well screening plate format, and any hits are further optimized using a larger volume of the corresponding solution. While a crystallization condition may eventually be found, the process can take anywhere from a few days to a year or two, making it the rate-limiting step for protein crystallographers. During this time, the protein must stay in solution and maintain its structure to produce a high-quality crystal, a condition that may not be met.

Similarly, a stable, highly concentrated protein sample is required to perform more advanced NMR experiments. These experiments require days or even weeks to run, during which the solution's homogeneity is key to obtaining high-quality spectra. Should the protein unfold or precipitate from solution during an experiment, the resulting chemical change would either not produce a signal or one that could not be used to determine structure/dynamics.

Working with frozen-hydrated specimens for Cryo-EM brings many challenges for manipulations and imaging. Exposure to atmospheric water vapor rapidly leads to frost buildup on the grid when handling cryo-EM grids for loading into the microscope. Under the TEM, these ice crystals on the grid surface appear as huge boulders that completely block the electron beam. Thus, grids are kept under liquid nitrogen whenever possible to minimize frost contamination. Problems with ice conditions are common—insufficient rapid freezing leads to hexagonal ice, while devitrification occurs when samples warm up, forming cubic ice. Various degrees of contamination may occur, and frosting at atmospheric pressure can cause the aforementioned ice crystal deposition. In contrast, contamination within the column or under low-vacuum conditions results in a more subtle artifact.

One of the hallmarks of protein crystallography is that size does not matter. Whether working with a 25 kDa monomeric protein or a 900 kDa multimeric complex, its structure can be determined if it can be crystallized and produce a high-resolution diffraction pattern. This is because, once in crystal form, a protein is in a more-or-less static conformation that can yield a single structural model in X-ray crystallography. Cryo-EM is similar in this regard. Very large structures, including massive nucleoprotein complexes, such as the ribosome, can be elucidated using Cryo-EM. The same cannot be said for NMR.

In NMR, the protein is in a soluble state and, therefore, in constant movement. The molecular tumbling rate is the most important movement governing spectral quality. The tumbling rate decreases significantly for proteins larger than 40 kDa, increasing the transverse relaxation rate (T_2). This results in a weaker, rapidly decaying NMR signal, leading to peak broadening and spectral overlap.

One of the major advantages of NMR is its ability to record small and large-scale protein dynamics, which is suppressed when a protein is crystallized. Although a crystallized protein may exhibit some motion within the lattice, these motions manifest as static or dynamic disorder, resulting in two different conformations of a particular region and, in the latter case, an averaged electron density. In general, crystallization may restrict a protein's natural flexibility and motions. Cryo-EM suffers from this same limitation as samples are frozen and immobile. However, cryo-EM can capture a snapshot of the native structure as freezing is instantaneous and does not require crystal lattice formation.

Crystallography, however, is not left out of dynamic structural analysis. Time-resolved crystallography can monitor protein structural changes upon ligand addition or environmental perturbations. Because all protein crystals are highly hydrated, they can serve as crucibles for some biochemical reactions. The crystal is typically soaked in a solution containing the ligand of interest to initiate the biochemical reaction, then quickly placed on the beamline, and the diffraction pattern is obtained. This can be performed multiple times if necessary to obtain a variety of structural intermediates. The process requires many things to go right: the protein cannot become disordered, the crystal cannot crack during soaking, and a high-powered synchrotron is required to collect high-quality diffraction data over short exposure times.

In the end, protein X-ray crystallography, cryo-EM, and NMR spectroscopy are not mutually exclusive techniques; one can easily fill in where the other falls short. In analyzing NMR dynamics experiments, for example, one can greatly benefit from existing crystal structure data or cryo-EM data onto which the NMR structural data can be superimposed. Similarly, NMR structure data can supplement cryo-EM or crystal structures with additional information on the protein's dynamics, binding interactions, and conformational changes in solution.

4.5.9: Molecular Dynamics Simulations

Introduction to Molecular Mechanics and Molecular Dynamics

Molecular modeling and computational chemistry are important parts of modern biochemistry. Models display, in a meaningful and instructive way, the large amounts of data generated when X-ray crystallography and NMR are used to determine the structures of large biological molecules and complexes. Remember, however, that primary X-ray crystallographic data (in the form of electron density maps) are just that; they must be interpreted like any other type of data. Structures must be refined and energy minimized to produce more realistic structures (without van der Waals overlap or missing atoms, for example). In addition, atoms within any molecule are not static but move as bonds vibrate, angles bend, etc. This implies that large biomolecules can adopt many possible conformations of varying energy. For proteins, some of these conformations might center on an average conformation at a local or global energy minimum, separated by activation energy barriers.

In contrast, small molecules whose structures can be minimized using ab initio or semi-empirical quantum mechanics (e.g., Spartan), large molecular structures (such as DNA, RNA, proteins, and their complexes) must be minimized using molecular mechanics based on Newton's laws. Atoms are treated as masses and bonds as springs with appropriate force constants. A force field containing all the relevant parameters for a given atom (for example, sp^3, sp^2, sp^2 aromatic, and sp C) and bond types is used to solve energy equations that sum all energies over all atoms and bonds in the molecule. These energies include interactions among bonded atoms (stretching, bending, torsion, wagging) and those among nonbonded atoms (electrostatic and van der Waals). In minimization calculations, the positions of atoms within a molecule are systematically or randomly moved, and the energy is recalculated to find a lower energy state and, hence, a more stable molecule. Minimization calculations can not probe all conformational space or easily move a structure from a local minimum to a global minimum if a large energy barrier separates the two. Energy minimization is usually performed in the absence of a solvent. Common force fields used for macromolecules are CHARMM, AMBER, and GROMOS. Parameters for specific atom types in a given bond include atomic mass, van der Waals radius, partial charge for atoms (from quantum mechanics), bond length (from electron diffraction data), angles, and force constants for bonds (modeled as springs, obtained from IR). These parameters are derived from experiments and theoretical (usually quantum mechanical) calculations on small organic molecules. A potential energy equation comprising terms from bond stretching, angle bending, and torsional angle changes (bonded interactions), as well as electrostatic and van der Waals interactions (nonbonded), is then solved (described below).

Molecular dynamics aims to simulate the actual changes in a molecule over time after an energy input (e.g., heat application at a higher temperature) is added to a molecule at equilibrium. The structure is placed in a "bath" of thousands of water molecules to make the simulation realistic. As is described below, if the energies of atoms in a large molecule are known, the forces acting on those atoms can be deduced. From Newton's Second Law (F = ma), the velocity, or change in position, of an atom in the structure over time can be determined. Suppose the dynamic simulation can be run for a long enough time. In that case, alternate conformations (perhaps those centered around a global minimum and those nearby in energy space—local minima) may be sampled. By determining what fraction of the simulated conformations resemble the two alternative conformations, the ΔG for the interconversion of the two states can be calculated. As you can imagine, these calculations require substantial computer time. They give very important information. However, protein conformational changes are often, if not always, associated with binding a biological molecule to a binding partner. *In silico* experiments offer important clues and support for results obtained using other study methods.

Molecular mechanics (MM) and molecular dynamics (MD) have become powerful tools for analyzing and predicting the properties of complex biological structures. The Nobel Prize in Chemistry in 2013 was awarded to Martin Karplus, Michael Levitt, and Arieh Warshel "for the development of multiscale models for complex chemical systems". Karplus, in particular, developed much of the present basis for MD simulations.

Energy (E), Force (F), and Motion

Considering the relationship between force and energy helps make the energy equations for the individual components more understandable. You have studied two general force equations in introductory chemistry and physics courses. Coulomb's Law describes the electrostatic force of attraction, F_C, between two charges, q_1 and q_2, separated by a distance r.

$$F_C = k\frac{q_1 q_2}{r^2} \tag{4.5.8}$$

The other is Hooke's Law, which describes the restorative force on a mass connected to a spring on stretching or compression of the spring.

$$F = -kx \tag{4.5.9}$$

where x is the spring's displacement from an equilibrium (at rest) position.

Our first interest is understanding how these equations might lead to equations describing the potential energy of a two-charge system or a compressed or stretched spring. We can best understand this by considering a simple example: a ball placed at various locations on a hill. If placed on a flat surface at the top and bottom of the hill, no net force is acting on the ball ($F_{net} = 0$), so it will not move. If placed at various locations on the downslope, it will experience a net downward force, shown qualitatively in Figure 4.5.25 below.

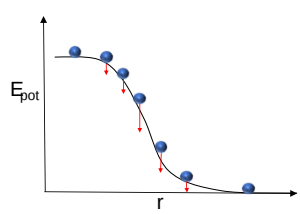


Figure 4.5.25: Potential Energy vs r for a ball on a hill

Astute observers will note that the magnitude of the force vector is proportional to the slope. From this simplistic approach, we come to the following equation relating F to E:

$$F = -\frac{dE}{dr} \tag{4.5.10}$$

The minus sign is required since the force is downward, but the energy increases upward.

This simplified approach can be extended into three dimensions to give the following equation (which will have meaning to those with an advanced calculus background), where F is the negative gradient of the potential energy:

$$F = -\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right) E = -\nabla E \tag{4.5.11}$$

Applying the 1D equation to Hooke's Law gives

$$\begin{gathered} dE = -Fdr = -kxdx \\ \int dE = -k \int xdx \end{gathered} \tag{4.5.12}$$

which gives

$$E = \frac{kx^2}{2} \tag{4.5.13}$$

This gives a parabolic graph of E vs displacement. Figure 4.5.26 below shows an interactive PHET simulation of Hooke's Law. Click Energy and then select Energy to see the parabolic plot. Change the force constant to alter the "steepness" of the resulting parabolic curve.

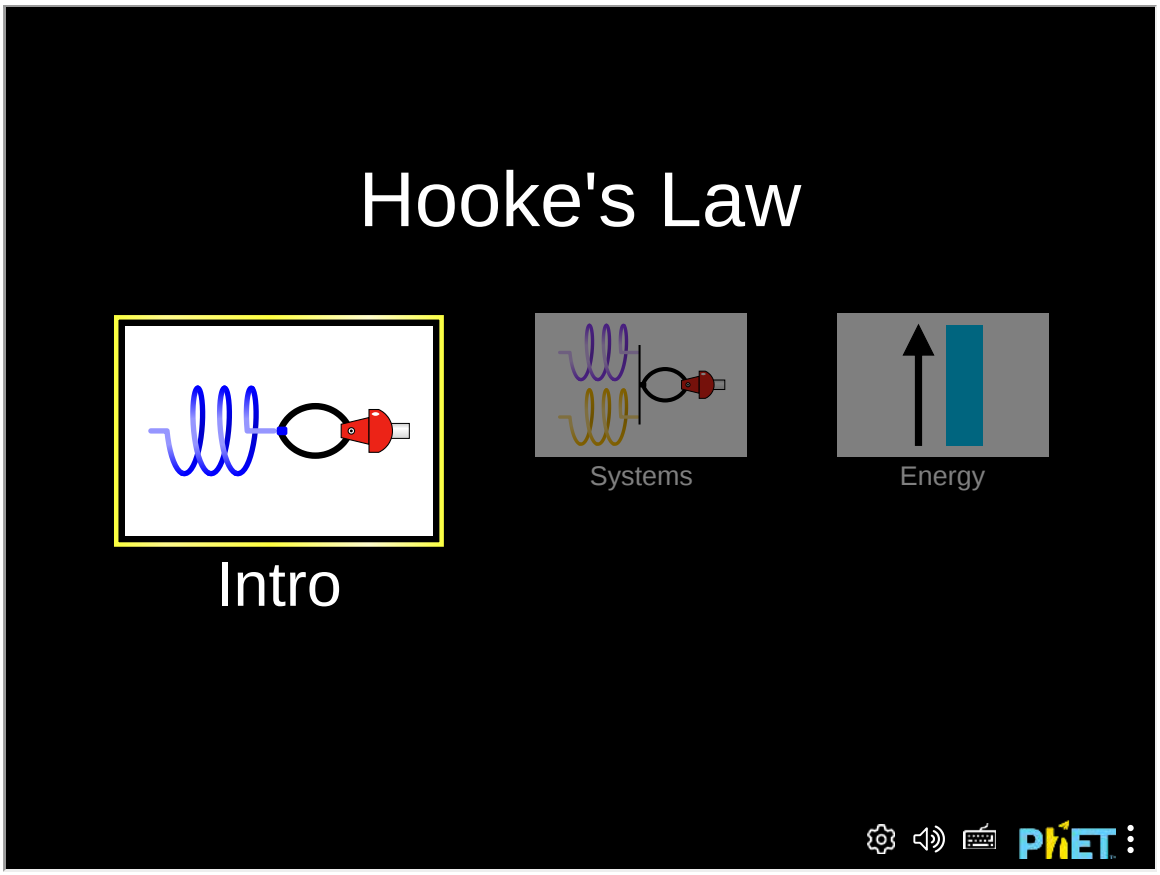


Figure 4.5.26: PHET simulation of Hooke's Law. https://phet.colorado.edu/en/simulations/hookes-law

The same approach can be applied to Coulomb's Law. Notice that the result equation for E results in increasingly negative values as r gets smaller, only when q1 and q2 have opposite charges.

$$\begin{gathered} dE = -Fdr = -k\frac{q_1 q_2}{r^2} dr \\ \int dE = -kq_1 q_2 \int r^{-2} dr \end{gathered} \tag{4.5.14}$$

gives

$$E = \frac{kq_1 q_2}{r} \tag{4.5.15}$$

A graph of E vs r for both attractive and repulsive interactions is shown in Figure 4.5.27 below.

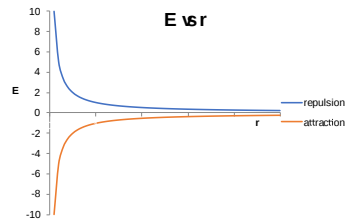


Figure 4.5.27: E vs r for both attractive and repulsive interactions

Molecular Mechanics

Note: The following review is based on an NIH Guide to Molecular Modeling (1996), which, to the best of our knowledge, was removed from the web.

Molecular mechanics uses Newtonian mechanics to calculate the energy of atoms in large molecules, such as proteins. It assumes that nuclei and electrons are single particles with radii and calculated charges. Bonds are treated as springs connecting atoms. Energies are calculated classically (not with quantum mechanics). Parameters, many based on quantum mechanical calculations of small molecules, are assigned to all bonds, angles, dihedrals, etc. Interactions are bonded (local) and nonbonded.

Bonded interactions involve atoms connected by one bond (bond stretch), two bonds (angle bending), and three bonds (dihedral angle change). These three types of bonded interactions are shown with black arrows on the right-hand side of Figure 4.5.28 below.

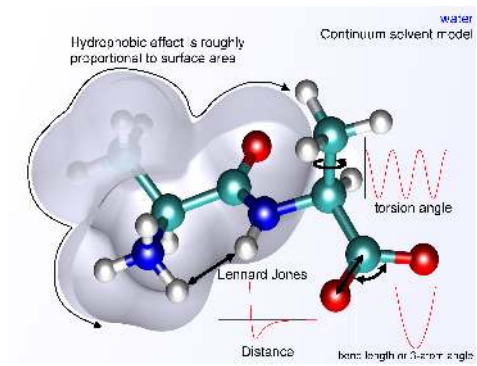


Figure 4.5.28: Bonded and Nonbonded Interactions in Proteins. Force field (chemistry). (2023, July 15). In *Wikipedia*. https://en.wikipedia.org/wiki/Force_field_(chemistry). CC BY-SA 3.0

Non-bonded atoms (greater than two bonds apart) interact through induced dipole-induced dipole interactions (one example of which is steric repulsions) and electrostatic attraction/repulsion. The double black arrow labeled Lennard-Jones in the above figure shows an example of a nonbonded interaction.

All energy terms from these interactions are summed to give the energy of a given conformation. The energy should be considered relative to that of other conformations. Here is the basic energy equation for all of these energy terms:

$$\text{Energy (E)} = E_{\text{Stretch}} + E_{\text{Bending}} + E_{\text{Torsion}} + E_{\text{Non-bonded Interactions}} \tag{4.5.16}$$

The "force field" consists of the energy equations and the parameters for each energy term. Many different commercially available force fields exist.

Bonded Interaction Energies

The mathematical form of the energy terms varies from force field to force field. The more common forms will be described.

Stretching (Vibrational) Energy

$$E_{\text{stretch}} = \Sigma_{\text{bonds}}\, k_b\, (r - r_o)^2 \tag{4.5.17}$$

Figure 4.5.29 illustrates bond stretching or vibration

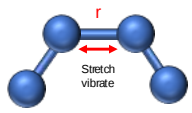


Figure 4.5.29: Bond stretching or vibration

The stretching energy equation is based on Hook's law. The k_b parameter defines the bond spring's stiffness. R_0 is the equilibrium distance between the two atoms. It should make sense that deviations from the equilibrium length would be associated with higher energy. The E vs. r curve is a parabola, as shown in Figure 4.5.30 below for a system with the lowest energy at r=6.

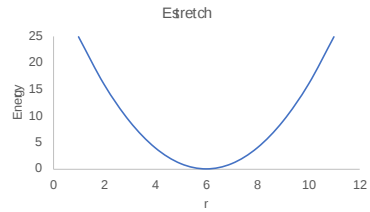


Figure 4.5.30: Energy vs r (r_0=6) for bond stretching (vibration)

Only small changes in r are allowed, as too large an r value would lead to bond breaking.

Bending Energy

$$E_{bending} = \Sigma_{angles}\, k_\Theta\, (\Theta - \Theta_0)^2 \tag{4.5.18}$$

Figure 4.5.31 illustrates bond bending

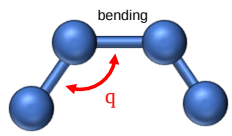


Figure 4.5.31: Bond bending

The bending energy equation is also based on Hooke's law. The kΘ parameter controls the stiffness of the angle spring, while the Θ_0 is the equilibrium angle. As above, the graph of E vs theta is expected to be a parabola as shown in Figure 4.5.32 below for a system when the lowest energy is at θ=45

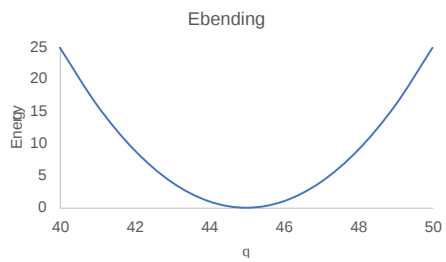


Figure 4.5.32: $E_{bending}$ vs r when lowest energy is at θ=45.

Torsion Energy

$$E_{torsion} = \Sigma_{torsions}\, A[1 + \cos(\text{ntau} - \Theta)] \tag{4.5.19}$$

Figure 4.5.33 illustrates torsion angle rotation

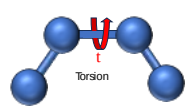


Figure 4.5.33: Torsion angle rotation

The torsion energy is modeled by a periodic function, much as you have seen with energy plots associated with Newman projections sighting down the central C-C bond in butane. Figure 4.5.34 below shows two different torsion energy functions.

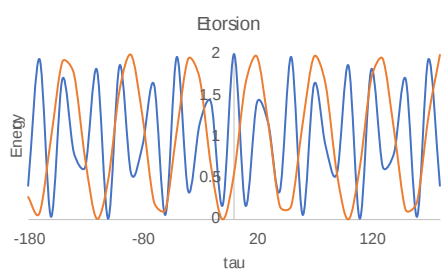


Figure 4.5.34: Torsion energy vs tau for A=1, n=1, phi=0 (blue) and A=1, n=2,phi=90 (orange).

Non-Bonded Interaction Energy

The non-bonded energy is calculated for all possible pairs of nonbonded atoms, i and j:

$E_{nonbonding} = \Sigma_i\, \Sigma_j\, [\, -B_{ij}/r_{ij}^6 + A_{ij}/r_{ij}^{12}\,] + \Sigma_i\, \Sigma_j\, (q_i\, q_j) / r_{ij}$

$$E_{\text{nonbonding}} = \Sigma_i \Sigma_j \left[-B_{ij}/r_{ij}^6 + A_{ij}/r_{ij}^{12}\right] + \Sigma_i \Sigma_j \left(q_i q_j\right)/r_{ij} \mid \tag{4.5.20}$$

The first term represents van der Waals interactions, while the second represents Coloumbic electrostatic interactions. Figure 4.5.35 illustrated nonbonded interactions.

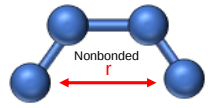


Figure 4.5.35: Nonbonded interactions

You should remember that induced dipole-induced dipole interactions are short-range and occur among all atoms. The 6-12 energy equation, based on the Lennard-Jones potential, shows a negative (attractive) term proportional to $-1/r^6$ and a repulsive term proportional to $+1/r^{12}$. Figure 4.5.36 below shows a graph of the attractive, repulsive, and summation of the energy terms in the Lennard-Jones potential.

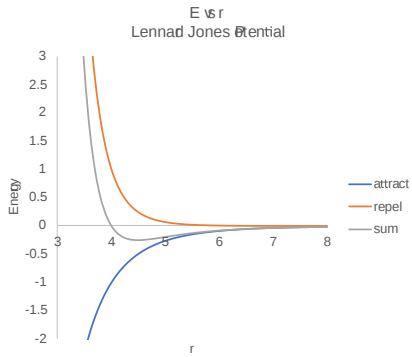


Figure 4.5.36: Lennard-Jones potential vs r

The A and B parameters control the depth and position (interatomic distance) of the potential energy for a pair of non-bonded interacting atoms (e.g., C:C and O:C). In effect, A determines the degree of stickiness of the van der Waals attraction, and B determines the degree of hardness of the atoms (e.g., marshmallow-like, billiard ball-like, etc.).

The B parameter is related to the "stickiness" of the interactions and the polarization of the atoms. B can be obtained from atomic polarizability measurements or calculated using quantum mechanics. The A parameter is empirically derived to fit nonbonded contacts between atoms in crystal structures.

Summary Interactions

Some programs assign charges using rules or templates, especially for macromolecules. The torsional potential is calibrated to a particular charge calculation method in some force fields. Using a different method can invalidate the force-field consistency. An additional bonded interaction term, improper dihedrals, is sometimes added, as illustrated below. The following equation gives the potential for that:

$E_{\text{improper}} = \Sigma_{\text{angles}}\, k\omega\, (\omega - \omega_0)^2$

$$E_{\text{improper}} = \Sigma \text{ angles } k\omega\,(\omega - \omega_0)^2 \tag{4.5.21}$$

Molecular Dynamics

This section comes directly from the NIH tutorial:

"In the broadest sense, MD is concerned with molecular motion. Motion is inherent to all chemical processes. Simple vibrations, like bond stretching and angle bending, give rise to IR spectra. Chemical reactions, hormone-receptor binding, and other complex processes are associated with many intra- and intermolecular motions.

Thermodynamics describes the driving force for chemical processes, while kinetics describes the mechanism by which they occur. Thermodynamics governs the energetic relationships between different chemical states, while kinetics describes the sequence and rates of events that occur as molecules transform between their various possible states.

Molecular dynamics studies usually focus on conformational transitions and local vibrations. MD alters the intermolecular degrees of freedom step-wise, analogous to energy minimization. The individual steps in energy minimization aim only to establish a downhill direction toward a minimum. On the other hand, the steps in MD meaningfully represent the changes in atomic position, ri, over time (i.e., velocity).

Newton's equation ($F_i = m_i a^i$) is used in the MD formalism to simulate atomic motion. As Newton's equation states, the rate and direction of motion (velocity) are governed by the forces that the system's atoms exert on one another. In practice, the atoms are assigned initial velocities that satisfy the total kinetic energy of the system, which, in turn, is determined by the desired simulation temperature. This is done by slowly heating the system (initially at absolute zero) and allowing the energy to equilibrate among the constituent atoms. The basic ingredients of MD are the calculation of the force on each atom and, from that information, the position of each atom through a specified time (typically on the order of picoseconds = 10^{-12} seconds).

The force on an atom can be calculated from the change in energy between its current position and its position a small distance away. This can be recognized as the derivative of the energy with respect to the change in the atom's position: $-dE/dr_i = F_i$.

Energies can be calculated using either MM or quantum mechanics methods. MM energies are limited to applications that do not involve drastic changes in electronic structure, such as bond-making or breaking. Quantum mechanical energies can be used to study dynamic processes involving chemical changes. The latter technique is extremely novel and limited in availability.

Knowledge of atomic forces and masses can then be used to determine the positions of each atom at a series of extremely small time steps (on the order of femtoseconds). A trajectory is the resulting series of snapshots of structural changes over time. The use of this method to compute trajectories can be more easily seen when Newton's equation is expressed in the following form

$-dE/dr_i = m_i a = md^2r_i/dt^2$

$$-dE/dr_i = m_i a = md^2r_i/dt^2 \tag{4.5.22}$$

In practice, trajectories are not obtained directly from Newton's equation because there is no analytical solution. First, the atomic accelerations are computed from the forces and masses acting on each atom. The velocities are next calculated from the accelerations based on the following relationship:
$a_i = dv_i/dt$. Lastly, the positions are calculated from the velocities: $v_i = dr_i/dt$. A trajectory between two states can be subdivided into sub-states separated by a small time step, delta t (e.g., 1 femtosecond).

The initial atomic positions at time t are used to predict the atomic positions at time t = delta t. The positions at t = delta t are used to predict the positions at t = 2Δt, and so on.

The leapfrog method is a common numerical method for calculating trajectories from Newton's equations. The method derives its name from the fact that velocity and position information alternate at ½-time-step intervals. MD has no defined termination point other than the time that can be practically covered.

Figure **4.5.37** below shows a 0.5 ns molecular dynamics simulation of sperm whale myoglobin, a muscle protein that binds O_2 at a macrocyclic ligand called heme. The heme and oxygen were not included in the simulation. The protein (molecular weight 17,331), consist of 8 alpha helices (red) connected by random coils. The backbone and side chains are shown in green sticks.

Figure **4.5.37**: A 0.5 ns molecular dynamics simulation of sperm whale myoglobin.

The simulation was performed in GROMACS using NeuroApp. The input PDB files for the simulations were first processed using Charmm-GUI. The output simulation files (input_protein and md_center.xtc from the output folder) from the Gromacs simulation were opened in Pymol and processed to make the animations shown above (File, Export Movie, as PNG, then animated GIF in Adobe Photoshop). Abraham, J. M. et al., GROMACS: High-performance molecular simulations through multi-level parallelism from laptops to supercomputers, https://www.sciencedirect.com/, September 2015, https://doi.org/10.1016/j.softx.2015.06.001. Neurosnap Inc. - Computational Biology Platform for Research. Wilmington, DE, 2022. https://neurosnap.ai/.

A molecular dynamics simulation of a small protein module (NTL9) folding is shown below.

MD simulations can be used to obtain theoretical values for ΔG and Keq values for conformational changes, binding of small ligands, and changes in protonation states for side chains. This process is based on the idea that the conformations sampled in silico MD simulations reflect those found in vitro (i.e., they are part of the thermodynamically expected and accessible conformations of the molecules during normal conformational shifts). This is called the Ergodic Hypothesis. Given the short time scales of MD simulations (limited by computer power), this hypothesis can't apply to dynamic results unless the sample conformations are close in energy, with no large activation energy barrier between them. If it is, then the following equation could apply:

$$\begin{aligned} \Delta G^0 &= -RT \ln K_{eq} \\ \Delta G^0 &= -RT \ln P_2/P_1 = -RT \ln f_2/f_1 \end{aligned} \tag{4.5.23}$$

where

P_n is the probability of being in a given state, and f_n is the fraction in a given state.

You will note that none of the potential energy functions use quantum mechanical parameters. This is due, in part, to the complexity of the systems studied. This is beginning to change as more effort is devoted to understanding the quantum mechanical aspects of complex bonded systems. New advances in even simple systems can illustrate this point. Take, for example, ethane. The conformational analysis of this simple molecule is discussed in all organic chemistry books. The energy vs dihedral angle (viewing the molecule down the C-C bond and measuring the angle between the C-H bonds on adjacent C atoms) oscillates every 120 °. The energy is maximum when the dihedral angles are 0 °, 120 °, 240 °, and 360°, corresponding to the eclipsed conformations. It reaches minima at the staggered (gauche) conformations at 60 °, 180 °, and 270 °. Why is the eclipsed form higher in energy than the staggered form? All organic books would state that there is greater steric repulsion (of the electron clouds) in the eclipsed forms, which raises their energy compared to the staggered forms. However, Pophristic shows that to be incorrect. For the correct answer, you must turn to quantum mechanics and the phenomena of hyperconjugation. The staggered conformation is energetically favored not only because it is less sterically restricted but also because it is a lower energy form due to resonance-like stabilization of the σ CH molecular orbitals. There is a greater correct phase overlap of σ CH and s* CH molecular orbitals on the adjacent Cs in the staggered conformation than in the eclipsed form.

4.5.10: Proteome Analysis

The proteome is the entire set of proteins produced or modified by an organism or system. Proteomics has enabled the identification of ever-increasing numbers of proteins. This varies over time and with distinct requirements or stresses a cell or organism undergoes. Proteomics is an interdisciplinary field that has benefited significantly from genetic information from various genome projects, including the Human Genome Project. It covers the exploration of proteomes, including overall protein composition, structure, and activity levels. It is an important component of functional genomics.

Proteomics is the next step in studying biological systems after genomics and transcriptomics. It is more complicated than genomics because an organism's genome is relatively constant, whereas proteomes vary from cell to cell and over time. Distinct genes are expressed in different cell types, meaning that even the basic set of proteins produced in a cell needs to be identified.

In the past, this phenomenon was assessed by RNA analysis, but it was found to lack correlation with protein content. It is known that mRNA is not always translated into protein, and the amount of protein produced per mRNA depends on the gene it is transcribed from and the cell's current physiological state. Proteomics confirms the presence of the protein and provides a direct measure of its quantity.

A cell may make different sets of proteins at different times or under different conditions, for example, during development, cellular differentiation, the cell cycle, or carcinogenesis. As mentioned, most proteins can undergo a wide range of post-translational modifications, further increasing proteome complexity.

Therefore, a proteomics study can become complex very quickly, even if the study is restricted. In more ambitious settings, such as when a biomarker for a specific cancer subtype is sought, the proteomics scientist might elect to study multiple blood serum samples from multiple cancer patients to minimize confounding factors and account for experimental noise. Furthermore, many proteins undergo post-translational modifications such as phosphorylation. Many of these post-translational modifications are critical to the protein's function. Thus, complicated experimental designs are sometimes necessary to account for the dynamic complexity of the proteome.

4.5.11: Protein Analyses Web Sites

- Uniprot - the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information
- The Human Proteome Atlas - an open-access resource for human proteins
- The Blood Resource from the Human Proteome Atlas - focus on the individual protein levels in the blood of both healthy individuals and patients diagnosed with various diseases

4.5.12: Summary

This chapter builds upon our previous discussion of protein purification by exploring the array of analytical techniques used to study protein structure and function at varying levels of resolution. After isolating a “pure” protein, the next step is to decipher its structure—from simple measurements like concentration to high-resolution three-dimensional models that reveal intricate details of protein-ligand interactions.

At the low-resolution end, various spectroscopic methods are introduced to determine protein concentration and secondary structure. Techniques such as UV absorbance, dye-binding assays (e.g., Bradford, Lowry, BCA), and quantitative amino acid analysis provide essential data on protein composition and purity. Circular dichroism (CD) spectroscopy is highlighted as a tool to differentiate among α-helices, β-sheets, and random coils based on the protein’s absorption of circularly polarized light.

The chapter then transitions to methods that yield more detailed structural information. Mass spectrometry is described in depth—from the principles of ionization (via ESI or MALDI) and mass analyzers to approaches for protein sequencing, including top-down and bottom-up strategies. These techniques are critical for determining molecular weight, amino acid composition, and mapping post-translational modifications.

High-resolution structural techniques, including X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy (cryo-EM), are examined next. Each method’s strengths, limitations, and sample requirements are discussed, illustrating how they complement each other. X-ray crystallography and cryo-EM provide static snapshots of protein structures at atomic resolution, whereas NMR offers insights into protein dynamics and conformational flexibility in solution.

Finally, the chapter introduces molecular dynamics (MD) simulations and molecular mechanics as computational tools that model protein behavior over time. These in silico approaches help bridge the gap between static experimental structures and the dynamic nature of proteins in biological systems, providing quantitative insights into energy landscapes and conformational transitions.

Overall, this chapter emphasizes the importance of integrating diverse analytical and computational techniques to achieve a comprehensive understanding of protein structure-function relationships, a fundamental goal of advanced studies in biochemistry and molecular biology.

4.5.13: References

Molnar, C. and Gair, J. (2013) Antibodies. Chapter in Concepts in Biology, Published by B.C. Open Textbook Project. Available at: https://opentextbc.ca/biology/chapter/23-3-antibodies/

The Human Atlas Project. (2019) Methods. Available at: https://www.proteinatlas.org/learn/method

Uhlén M et al, 2015. **Tissue-based map of the human proteome.** *Science*
PubMed: 25613900 DOI: 10.1126/science.1260419

Thul PJ et al, 2017. **A subcellular map of the human proteome.** *Science.*
PubMed: 28495876 DOI: 10.1126/science.aal3321

Uhlen M et al, 2017. **A pathology atlas of the human cancer transcriptome.** *Science.*
PubMed: 28818916 DOI: 10.1126/science.aan2507

Ahern, K. and Rajagopal, I. (2019) Biochemistry Free and Easy. Published by Libretexts. Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_and_Easy_(Ahern_and_Rajagopal)/09%3A_Techniques/9.04%3A_Gel_Exclusion_Chromatography.

Magdeldin, S. (2012) Gel Electrophoresis - Principles and Basics. Published by InTech under Creative Commons Attribution 3.0. Available at: https://pdfs.semanticscholar.org/4b93/70ac3946cec6e12c369679c4178a5ef38e61.pdf

Structural Biochemistry/Proteins/X-ray Crystallography. (2018, November 19). *Wikibooks, The Free Textbook Project*. Retrieved 15:40, August 17, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins/X-ray_Crystallography&oldid=3488057.

UCD: Biophysics 200A (2019) "NMR Spectroscopy vs X-ray Crystallography", Chapter published in Current Techniques in Biophysics. Published by Libretexts and available at: https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_200A_-_Current_Techniques_in_Biophysics/NMR_Spectroscopy_vs._X-ray_Crystallography

Wikipedia contributors. (2019, June 27). Protein purification. In *Wikipedia, The Free Encyclopedia*. Retrieved 23:32, July 28, 2019, from en.Wikipedia.org/w/index.php?title=Protein_purification&oldid=903657925

Wikipedia contributors. (2019, February 15). Fast protein liquid chromatography. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:14, August 15, 2019, from en.Wikipedia.org/w/index.php?title=Fast_protein_liquid_chromatography&oldid=883530035

Wikipedia contributors. (2019, July 9). Protein mass spectrometry. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:27, August 16, 2019, from en.Wikipedia.org/w/index.php?title=Protein_mass_spectrometry&oldid=905547289

Wikipedia contributors. (2019, July 8). Peptide synthesis. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:13, August 17, 2019, from en.Wikipedia.org/w/index.php?title=Peptide_synthesis&oldid=905401648

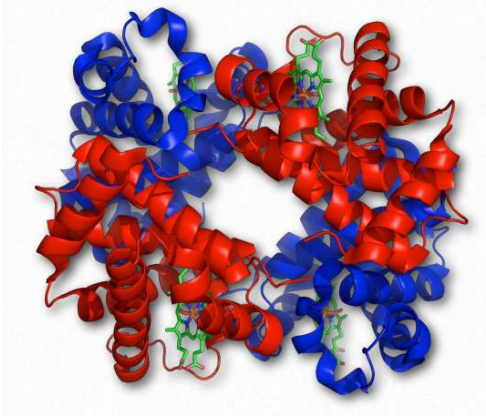
Mass Spectral Analysis

This page titled 4.5: Analyses of Protein Structure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

4.6: Protein function, domains and cooperativity

Source: BiochemFFA_2_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

In this section, we hope to bring to life the connection between structure and function of proteins. So far, we have described notable features of the four elements (primary, secondary, tertiary, and quaternary) of protein structure and discussed example proteins/motifs exhibiting them. In this section, we will examine from a functional perspective a few proteins/domains whose function relies on secondary, tertiary, or quaternary structure. It is, of course a bit of a narrow focus to ascribe protein function to any one component of structure, but our hope is by presenting these examples, we can bring to life the way in which a protein's secondary, tertiary, and quarternary structure lead to the functions it has.



Hemoglobin Wikipedia

Fibrous proteins - secondary structure

Proteins whose cellular or extracellular roles have a strong structural component are composed primarily of primary and second structure, with little folding of the chains. Thus, they have very little tertiary structure and are fibrous in nature. Proteins exhibiting these traits are commonly insoluble in water and are referred to as fibrous proteins (also called scleroproteins). The examples described in this category are found exclusively in animals where they serve roles in flesh, connective tissues and hardened external structures, such as hair. They also contain the three common fibrous protein structures α -helices (keratins), β-strands/sheets (fibroin & elastin) and triple helices (collagen). The fibrous proteins have some commonality of amino acid sequence. Each possesses an abundance of repeating sequences of amino acids with small, non-reactive side groups. Many contain short repeats of sequences, often with glycine.

Keratins

Figure 2.56 - The horns of an impala are composed of keratin Wikipedia

The keratins are a family of related animal proteins that take numerous forms. α-keratins are structural components of the outer layer of human skin and are integral to hair, nails, claws, feathers, beaks, scales, and hooves. Keratins provide strength to tissues, such as the tongue, and over 50 different keratins are encoded in the human genome. At a cellular level, keratins comprise the intermediate filaments of the cytoskeleton. α- keratins primarily contain α-helices, but can also have β-strand/sheet structures. Individual α-helices are often intertwined to form coils of coiled structures and these strands can also be further joined together by disulfide bonds, increasing structural strength considerably. This is particularly relevant for α-keratin in hair, which contains about 14% cysteine. The odor of burned hair and that of the chemicals used to curl/uncurl hair (breaking/re-making disulfide bonds) arise from their sulfurous components. β-keratins are comprised of β-sheets, as their name implies.

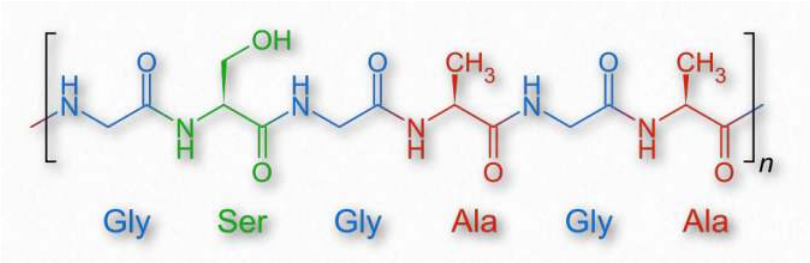


Figure 2.57 - The repeating amino acid sequence of fibroin

Fibroin

An insoluble fibrous protein that is a component of the silk of spiders and the larvae of moths and other insects, fibroin is comprised of antiparallel β-strands tightly packed together to form β- sheets. The primary structure of fibroin is a short repeating sequence with glycine at every other residue (Figure 2.57). The small R-groups of the glycine and alanine in the repeating sequence allows for the tight packing characteristic of the fibers of silk. Wikipedia link HERE Elastin As suggested by its name, elastin is a protein with elastic characteristics that functions in many tissues of the body to allow them to resume their shapes after expanding or contracting. The protein is rich in glycine and proline and can comprise over 50% of the weight of dry, defatted arteries.

Figure 2.58 - Weaving of a silk sari Wikipedia

Elastin

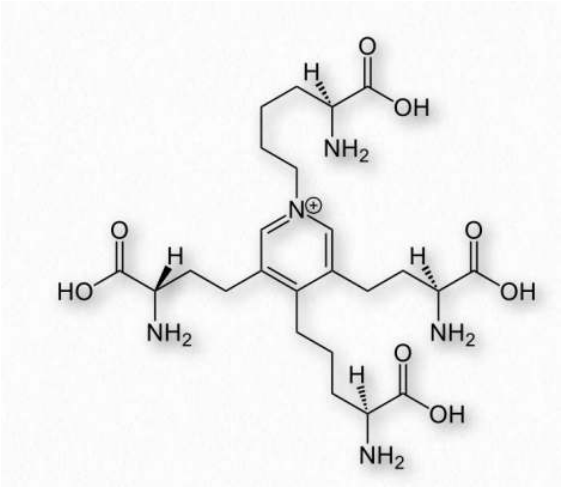


Figure 2.59 - Desmosine Wikipedia

is made by linking tropoelastin proteins together through lysine residues to make a durable complex crosslinked by desmosine. In arteries, elastin helps with pressure wave propagation for facilitating blood flow.

Collagen

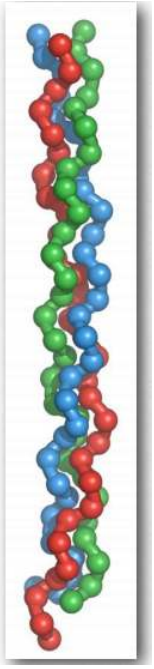


Figure 2.60 - Collagen's triple helix Wikipedia

Collagen is the most abundant protein in mammals, occupying up to a third of the total mass. There are at least 16 types of collagen. Its fibers are a major component of tendons and they are also found abundantly in skin. Collagen is also prominent in cornea, cartilage, bone, blood vessels and the gut.

Collagen's structure is an example of a helix of helices, being composed of three lefthanded helical chains that each are coiled together in a right-handed fashion to make the collagen fiber (Figure 2.60). Each helix is stretched out more than an α-helix, giving it an extended appearance. On the inside of the triple helical structure, only residues of glycine are found, since the side chains of other amino acids are too bulky. Collagen chains have the repeating structure glycinem-n where m is often proline and n is often hydroxyproline (Figure 2.61).

Gly-Pro-Met-Gly-Pro-Ser-Gly-Pro-Arg

Gly-Leu-Hpr-Gly-Pro-Hpr-Gly-Ala-Hpr

Gly-Pro-Gln-Gly-Phe-Gln-Gly-Pro-Hpr

Gly-Glu-Hpr-Gly-Glu-Hpr-Gly-Ala-Ser

Gly-Pro-Met-Gly-Pro-Arg-Gly-Pro-Hpr

Gly-Pro-Hpr-Gly-Lys-Asn-Gly-Asp-Asp

Hpr = Hydroxyproline

Figure 2.61 - Repeating sequences in collagen

Collagen is synthesized in a pre-procollagen form. Processing of the pre-procollagen in the endoplasmic reticulum results in glycosylation, removal of the 'pre' sequence, and hydroxylation of lysine and proline residues (see below). The hydroxides can form covalent cross-links with each other, strengthening the collagen fibers. As pro-collagen is exported out of the cell, proteases trim it, resulting in a final form of collagen called tropocollagen.

Hydroxylation

Hydroxylation of proline and lysine side chains occurs post-translationally in a reaction catalyzed by prolyl-4-hydroxylase and lysyl-hydroxylase (lysyl oxidase), respectively. The reaction requires vitamin C. Since hydroxylation of these residues is essential for formation of stable triple helices at body temperature, vitamin C deficiency results in weak, unstable collagen and, consequently, weakened connective tissues. It is the cause of the disease known as scurvy. Hydrolyzed collagen is used to make gelatin, which is important in the food industry. collagens. Wikipedia link HERE

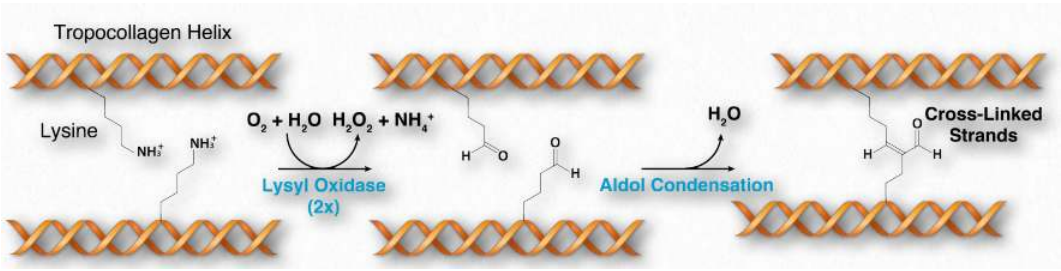


Figure 2.62 - Oxidation and cross-linking of lysine residues in tropocollagen. Only two strands of the triple helix are shown for simplicity Image by Aleia Kim

Lamins

Lamins are fibrous proteins that provide structure in the cell nucleus and play a role in transcription regulation. They are similar to proteins making up the intermediate filaments, but have extra amino acids in one coil of the protein. Lamins help to form the nuclear lamin in the interior of the nuclear envelope and play important roles in assembling and disassembling the latter in the process of mitosis. They also help to position nuclear pores. In the process of mitosis, disassembly of the nuclear envelope is promoted by phosphorylation of lamins by a protein called mitosis promoting factor and assembly is favored by reversing the reaction (dephosphorylation).

Structural domains - tertiary structure

Every globular protein relies on its tertiary structure to perform its function, so rather than trying to find representative proteins for tertiary structure (an almost impossible task!), we focus here on a few elements of tertiary structure that are common to many proteins. These are the structural domains and they differ from the structural motifs of supersecondary structure by being larger (25-500 amino acids), having a conserved amino acid sequence, and a history of evolving and functioning independently of the protein chains they are found in. Structural domains are fundamental units of tertiary structure and are found in more than one protein. A structural domain is selfstabilizing and often folds independently of the rest of the protein chain.

Leucine zipper

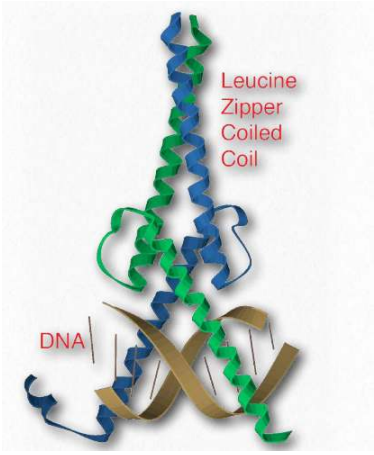


Figure 2.63 - Leucine zipper bound to DNA Wikipedia

A common feature of many eukaryotic DNA binding proteins, leucine zippers are characterized by a repeating set of leucine residues in a protein that interact like a zipper to favor dimerization. Another part of the domain has amino acids (commonly arginine and lysine) that allow it to interact with the DNA double helix (Figure 2.63). Transcription factors that contain leucine zippers include Jun-B, CREB, and AP-1 fos/ jun.

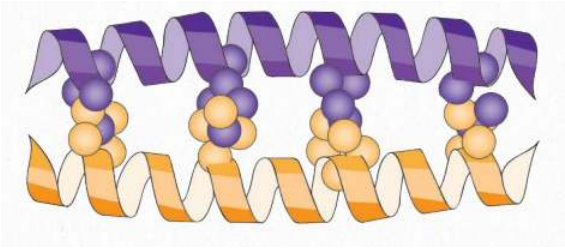


Figure 2.64 - Leucine zipper structure. Leucines are indicated by orange and purple balls. Image by Penelope Irving

Zinc fingers

The shortest structural domains are the zinc fingers, which get their name from the fact that one or more coordinated zinc ions stabilize their finger-like structure. Despite their name, some zinc fingers do not bind zinc. There are many structural domains classified as zinc fingers and these are grouped into different families. Zinc fingers were first identified as components of DNA binding transcription factors, but others are now known to bind RNA, protein, and even lipid structures. Cysteine and histidine side chains commonly play roles in coordinating the zinc.

Src SH_2 domain

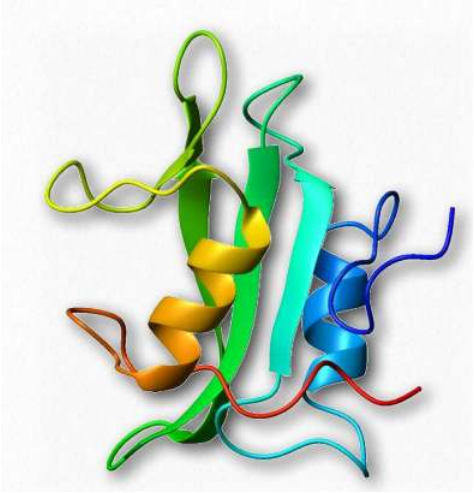


Figure 2.65 - SH2 Domain Wikipedia

The Src oncoprotein contains a conserved SH_2structural domain that recognizes and binds phosphorylated tyrosine side chains in other proteins (Figure 2.65). Phosphorylation is a fundamental activity in signaling and phosphorylation of tyrosine and interaction between proteins carrying signals is critically needed for cellular communication. The SH_2domain is found in over 100 human proteins.

Helix-turn-helix domain

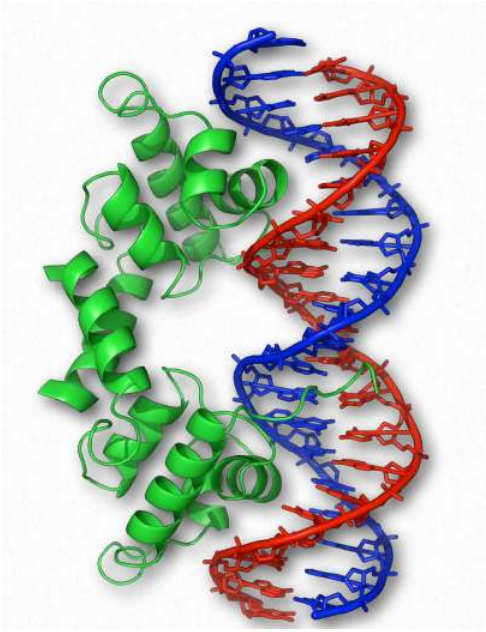


Figure 2.66 - Helix-Turn-Helix Domain of a Protein Bound to DNA Wikipedia

Helix-turn-helix is a common domain found in DNA binding proteins, consisting of two α-helices separated by a small number of amino acids. As seen in Figure 2.66, the helix parts of the structural domain interact with the bases in the major groove of DNA. Individual α-helices in a protein are part of a helix-turn-helix structure, where the turn separates the individual helices.

Pleckstrin homology domain

Pleckstrin Homology (PH) domains are protein domains with important functions in the process of signaling. This arises partly from the affinity for binding phosphorylated inositides, such as PIP2 and PIP3, found in Figure 2.66 - Helix-Turn-Helix Domain of a Protein Bound to DNA Wikipedia Figure 2.65 - SH2 Domain Wikipedia biological membranes. PH domains can also bind to G-proteins and protein kinase C. The domain spans about amino acids and is found in numerous signaling proteins. These include Akt/Rac Serine/ Threonine Protein Kinases, Btk/ltk/Tec tyrosine protein kinases, insulin receptor substrate (IRS-1), Phosphatidylinositolspecific phospholipase C, and several yeast proteins involved in cell cycle regulation.

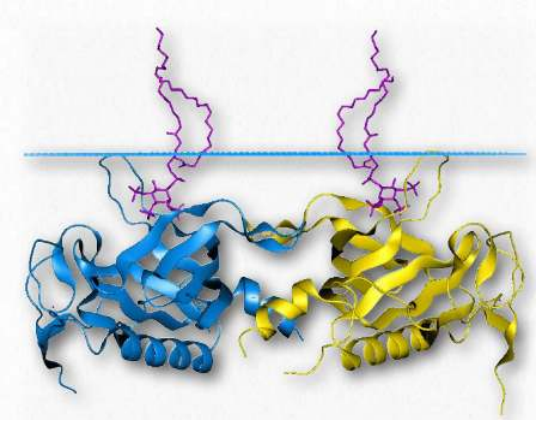


Figure 2.67 - Pleckstrin homology domain of Btk tyrosine protein kinase. The protein is embedded in a membrane (above blue line) Wikipedia

Structural globular proteins

Enzymes catalyze reactions and proteins such as hemoglobin perform important specialized functions. Evolutionary selection has reduced and eliminated waste so that we can be sure every protein in a cell has a function, even though in some cases we may not know what it is. Sometimes the structure of the proFigure 2.68 - Relationship of basement membrane to epithelium, endothelium, and connective tissue tein is its primary function because the structure provides stability, organization, connections other important properties. It is with this in mind that we present the following proteins.

Basement membrane

The basement membrane is a layered extracellular matrix of tissue comprised of protein fibers (type IV collagen) and glycosaminoglycans that separates the epithelium from other tissues (Figure 2.68). More importantly, the basement membrane acts like a glue to hold tissues together. The skin, for example, is anchored to the rest of the body by the basement membrane. Basement membranes provide an interface of interaction between cells and the environment around them, thus facilitating signaling processes. They play roles in differentiation during embryogenesis and also in maintenance of function in adult organisms.

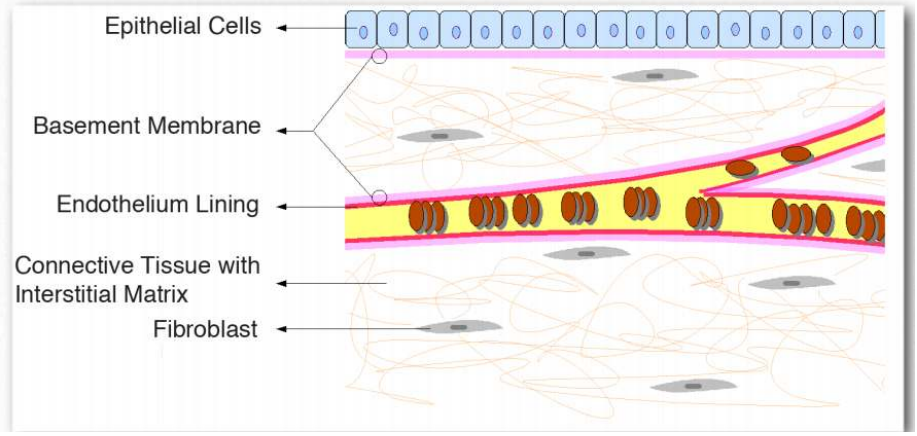


Figure 2.68 - Relationship of basement membrane to epithelium, endothelium, and connective tissue

Actin

Actin is the most abundant globular protein found in most types of eukaryotic cells, comprising as much as 20% of the weight of muscle cells. Similar proteins have been identified in bacteria (MreB) and archaeons (Ta0583). Actin is a monomeric subunit able to polymerize readily into two different types of filaments. Microfilaments are major component of the cytoskeleton and are acted on by myosin in the contraction of muscle cells (See HERE). Actin will be discussed in more detail in the next section HERE.

Intermediate Filaments

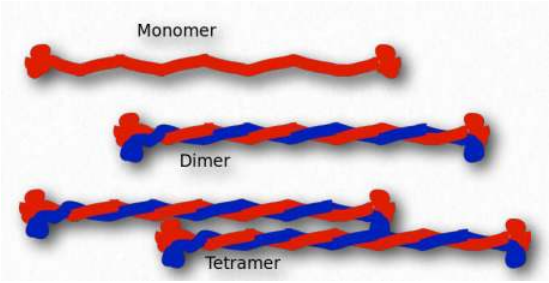


Figure 2.69 - Assembly of intermediate filaments

Intermediate filaments are a part of the cytoskeleton in many animal cells and are comprised of over 70 different proteins. They are called intermediate because their size (average diameter = 10 nm) is between that of the microfilaments (7 nm) and the microtubules (25 nm).

The intermediate filament components include fibrous proteins, such as the keratins and the lamins, which are nuclear, as well as cytoplasmic forms. Intermediate filaments give flexibility to cells because of their own physical properties. They can, for example, be stretched to several times of their original length.

Six types

There are six different types of intermediate filaments. Type I and II are acidic or basic and attract each other to make larger filaments. They include epithelial keratins and trichocytic keratins (hair components). Type III proteins include four structural proteins - desmin, GFAP (glial fibrillary acidic protein), peripherin, and vimentin. Type IV also is a grouping of three proteins and one multiprotein structure (neurofilaments). The three proteins are α-internexin, synemin, and syncoilin. Type V intermediate

filaments encompass the lamins, which give structure to the nucleus. Phosphorylation of lamins leads to their disassembly and this is important in the process of mitosis. The Type VI category includes only a single protein known as nestin.

Tubulin

A third type of filament found in cells is that of the microbutules. Comprised of a polymer of two units of a globular protein called tubulin, microtubules provide “rails” for motor proteins to move organelles and other “cargo” from one part of a cell to another. Microtubules and tubulin are discussed in more detail HERE.

Vimentin

Vimentin (Figure 2.70) is the most widely distributed protein of the intermediate filaments. It is expressed in fibroblasts, leukocytes, and blood vessel endothelial cells. The protein has a significant role maintaining the position of organelles in the cytoplasm, with attachments to the nucleus, mitochondria, and endoplasmic reticulum (Figure 2.70). Vimentin provides elasticity to cells and resilience that does not arise from the microtubules or microfilaments. Wounded mice that lack the vimentin gene survive, but take longer to heal wounds than wild type mice. Vimentin also controls the movement of cholesterol from lysosomes to the site of esterification. The result is a reduction in the amount of cholesterol stored inside of cells and has implications for adrenal cells, which must have esters of cholesterol.

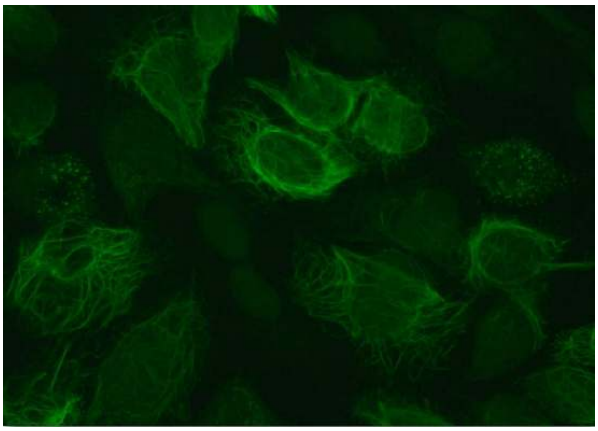


Figure 2.70 - Vimentin in cells Wikipedia

Mucin

Mucins are a group of proteins found in animal epithelial tissue that have many glycosyl residues on them and typically are of high molecular weight (1 to 10 million Da). They are gel-like in their character and are often used for lubrication. Mucus is comprised of mucins. In addition to lubrication, mucins also help to control mineralization, such as bone formation in vertebrate organisms and calcification in echinoderms. They also play roles in the immune system by helping to bind pathogens. Mucins are commonly secreted onto mucosal surfaces (nostrils, eyes, mouth, ears, stomach, genitals, anus) or into fluids, such as saliva. Because of their extensive mucosylation, mucins hold a considerable amount of water (giving them the “slimy” feel) and are resistant to proteolysis.

Vinculin

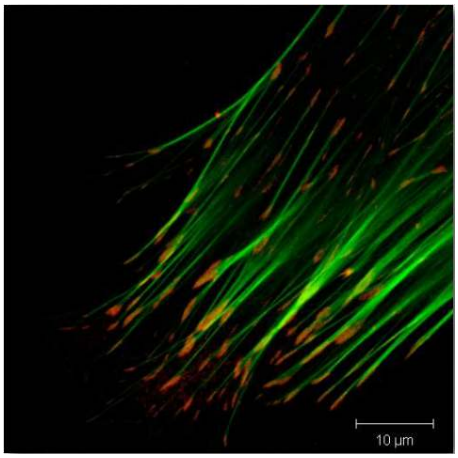


Figure 2.71 - Actin filaments (green) attached to vinculin in focal adhesion (red) Wikipedia

Vinculin (Figure 2.72) is a membrane cytoskeletal protein found in the focal adhesion structures of mammalian cells. It is found at cell-cell and cell-matrix junctions and interacts with integrins, talin, paxillins and F-actin. Vinculin is thought to assist (along with other proteins) in anchoring actin microfilaments to the membrane (Figure 2.71). Binding of vinculin to actin and to talin is regulating by polyphosphoinositides and can be inhibited by acidic phospholipids.

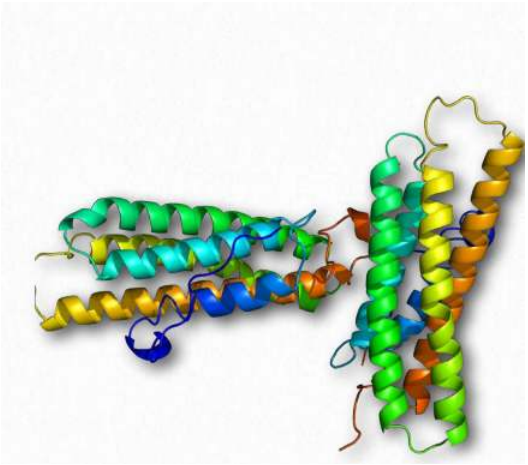


Figure 2.72 - Vinculin Wikipedia

Syndecans

Syndecans are transmembrane proteins that make a single pass with a long amino acid chain (24-25 residues) through plasma membranes and facilitate G proteincoupled receptors' interaction with Figure 2.71 - Actin filaments (green) attached to vinculin in focal adhesion (red) Wikipedia ligands, such as growth factors, fibronectin, collagens (I, III, and IV) and antithrombin-1. Syndecans typically have 3-5 heparan sulfate and chondroitin sulfate chains attached to them.

Heparan sulfate can be cleaved at the site of a wound and stimulate action of fibroblast growth factor in the healing process. The role of syndecans in cell-cell adhesion is shown in mutant cells lacking syndecan I that do not adhere well to each other. Syndecan 4 is also known to adhere to integrin. Syndecans can also inhibit the spread of tumors by the ability of the syndecan 1 ectodomain to suppress growth of tumor cells without affecting normal epithelial cells.

Defensin

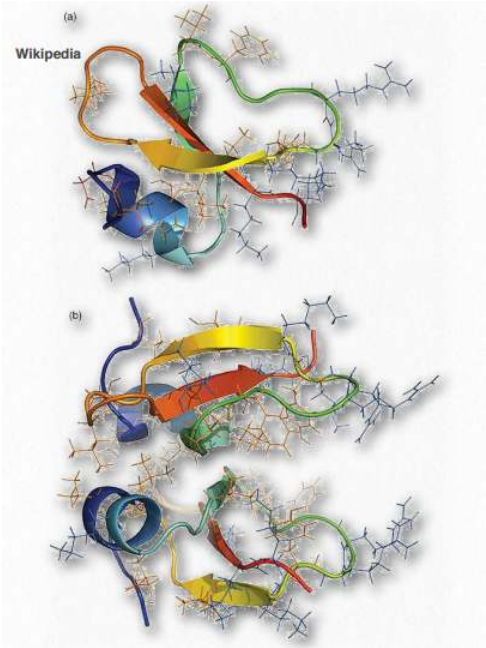


Figure 2.73 - Defensin monomer (top) and dimer (bottom) - Cationic residues in blue, hydrophobic residues in orange, and anionic residues in red Wikipedia

Defensins (Figure 2.73) are a group of small cationic proteins (rich in cysteine residues) that serve as host defense peptides in vertebrate and invertebrate organisms. They protect against infection by various bacteria, fungi, and viruses. Defensins contain between 18 and 45 amino acids with (typically) about 6- 8 cysteine residues. In the immune system, defensins help to kill bacteria engulfed by phagocytosis by epithelial cells and neutrophils. They kill 120 Figure 2.72 - Vinculin Wikipedia bacteria by acting like ionophores - binding the membrane and opening pore-like structures to release ions and nutrients from the cells.

Focal adhesions

In the cell, focal adhesions are structures containing multiple proteins that mechanically link cytoskeletal structures (actin bundles) with the extracellular matrix. They are dynamic, with proteins bringing and leaving with signals regarding the cell cycle, cell motility, and more almost constantly. Focal adhesions serve as anchors and as a signaling hub at cellular locations where integrins bind molecules and where membrane clustering events occur. Over 100 different proteins are found in focal adhesions.

Focal adhesions communicate important messages to cells, acting as sensors to update information about the status of the extracellular matrix, which, in turn, adjusts/ affects their actions. In sedentary cells, they are stabler than in cells in motion because when cells move, focal adhesion contacts are established at the “front” and removed at the rear as motion progresses. This can be very important in white blood cells’ ability to find tissue damage.

Ankyrin

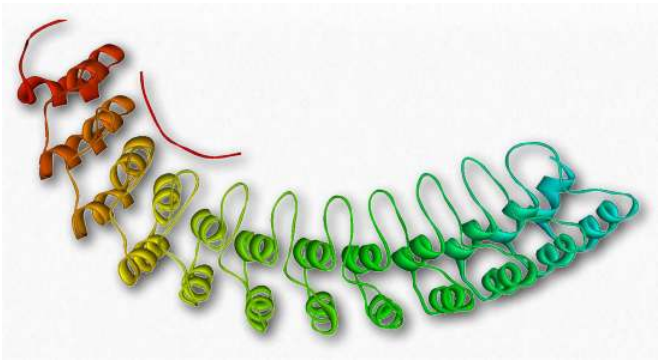


Figure 2.74 - Ankyrin’s membrane-binding domain

Ankyrins (Figure 2.74) are a family of membrane adaptor proteins serving as “anchors” to interconnect integral membrane proteins to the spectrin-actin membrane cytoskeleton. Ankyrins are anchored to the plasma membrane by covalently linked palmitoyl-CoA. They bind to the β subunit of spectrin and at least a dozen groups of integral membrane proteins. The ankyrin proteins contain four functional domains: an N-terminal region with 24 tandem ankyrin repeats, a central spectrin-binding domain, a “death domain” interacting with apoptotic proteins, and a C-terminal regulatory domain that is highly varied significantly among different ankyrins.

Spectrin

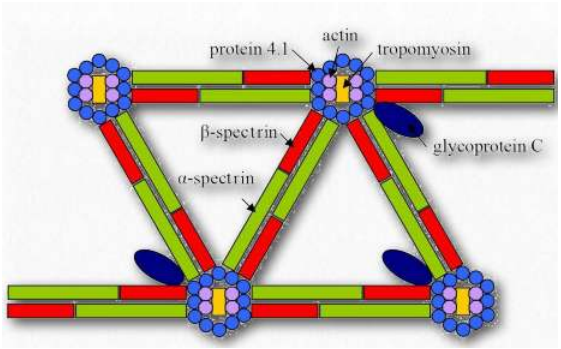


Figure 2.75 - Spectrin and other proteins in the cytoskeleton Wikipedia

Spectrin (Figures 2.75 & 2.76) is a protein of the cellular cytoskeleton that plays an important role in maintaining its structure and the integrity of the plasma membrane. In animals, spectrin gives red blood cells their shape. Spectrin is located inside the inner layer of the eukaryotic plasma membrane where it forms a network of pentagonal or hexagonal arrangements.

Spectrin fibers collect together at junctional complexes of actin and is also attached to ankyrin for stability, as well as numerous integral membrane proteins, such as glycophorin.

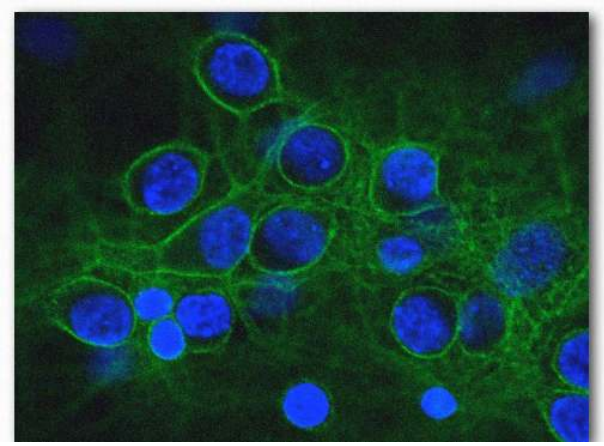


Figure 2.76 - Spectrin (green) and nuclei (Blue) Wikipedia

Integrins

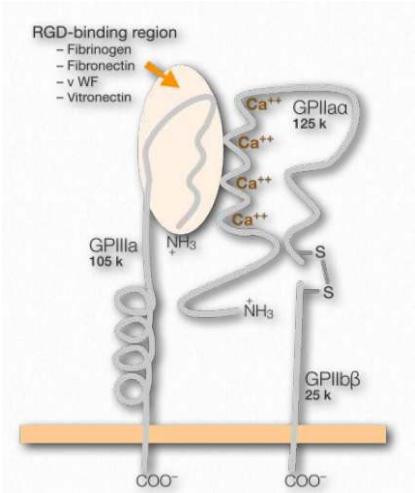


Figure 2.77 - Integrin and its binding site (on top left) Wikipedia

In multicellular organisms, cells need connections, both to each other and to the extracellular matrix. Facilitating these attachments at the cellular end are transmembrane proteins known as integrins (Figure 2.77). Integrins are found in all metazoan cells. Ligands for the integrins include collagen, fibronectin, laminin, and vitronectin. Integrins function not only in attachment, but also in communication, cell migration, virus linkages (adenovirus, for example), and blood clotting. Integrins are able to sense chemical and mechanical signals about the extracellular matrix and move that information to intracellular domains as part of the process of signal transduction. Inside the cells, responses to the signals affect cell shape, regulation of the cell cycle, movement, or changes in other cell receptors in the membrane. The process is dynamic and allows for rapid responses as may be necessary, for example in the process of blood clotting, where the integrin known as GPIbIIIa (on the surface of blood platelets) attaches to fibrin in a clot as it develops.

Integrins work along with other receptors, including immunoglobulins, other cell adhesion molecules, cadherins, selectins, and syndecans. In mammals the proteins have a large number of subunits - 18 α- and 8 β-chains. They are a bridge between its links outside the cell to the extracellular matrix (ECM) and its links inside the cell to the cytoskeleton. Integrins play central role in formation and stability of focal adhesions. These are large molecular complexes arising from clustering of integrin-ECM connections. In the process of cellular movement, integrins at the “front” of the cell (in the direction of the movement), make new attachments to substrate and release connections to substrate in the back of the cell. These latter integrins are then endocytosed and reused.

Integrins also help to modulate signal transduction through tyrosine kinase receptors in the cell membrane by regulating movement of adapters to the plasma membrane. β1c integrin, for example, recruits the Shp2 phosphatase to the insulin growth factor receptor to cause it to become dephosphorylated, thus turning off the signal it communicates. Integrins can also help to recruit signaling molecules inside of the cell to activated tyrosine kinases to help them to communicate their signals.

Cadherins

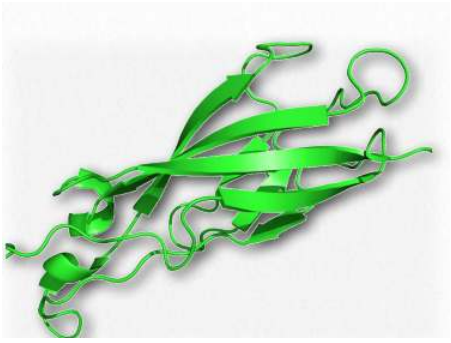


Figure 2.78 - Extracellular ectodomain of a cadherin

Cadherins (Figure 2.78) constitute a type-1 class of transmembrane proteins playing important roles in cell adhesion. They require calcium ions to function, forming adherens junctions that hold tissues together (See Figure 2.69). Cells of a specific cadherin type will preferentially cluster with each other in preference to associating with cells containing a different cadherin type. Caderins are both receptors and places for ligands to attach. They assist in the proper positioning of cells in development, separation of different tissue layers, and cell migration.

Selectins

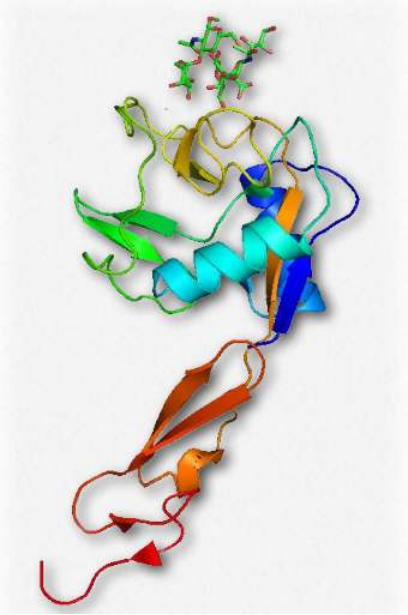


Figure 2.79 - Selectin bound to a sugar Wikipedia

Selectins (Figure 2.79) are cell adhesion glycoproteins that bind to sugar molecules. As such, they are a type of lectin - proteins that bind sugar polymers (see HERE also). All selectins have an N-terminal calcium-dependent lectin domain, a single transmembrane domain, and an intracellular cytoplasmic tail.

There are three different types of selectins, 1) E-selectin (endothelial); 2) L (lymphocytic; and 3) P (platelets and endothelial cells. Selectins function in lymphocyte homing (adhesion of blood lymphocytes to cells in lymphoid organs), in inflammation processes, and in cancer metastasis. Near the site of inflammation, P-selectin on the surface of blood capillary cells interacts with glycoproteins on leukocyte cell surfaces. This has the effect of slowing the movement of the leukocyte. At the target site of inflammation, E- selectin on the endothelial cells of the blood vessel and L-selectin on the surface of the leukocyte bind to their respective carbohydrates, stopping the leukocyte movement. The leukocyte then crosses the wall of the capillary and begins the immune response. Selectins are involved in the inflammatory processes of asthma, psoriasis, multiple scleroris, and rheumatoid arthritis.

Laminins

Laminins are extracellular matrix glycoproteins that a major components of the basal lamina and affect cell differentiation, migration, and adhesion. They are secreted into the extracellular matrix where they are incorporated and are essential for tissue maintenance and survival. When laminins are defective, muscles may not form properly and give rise to muscular dystrophy.

Laminins are associated with fibronectin, entactin, and perlecan proteins in type IV collagen networks and bind to integrin receptors in the plasma membrane. As a consequence, laminins contribute to cellular attachment, differentiation, shape, and movement. The proteins are trimeric in structure, having one α-chain, a β-chain, and a γ-chain. Fifteen combinations of different chains are known.

Vitronectin

Vitronectin is a glycoprotein (75kDa) found in blood serum (platelets), the extracellular matrix, and in bone. It promotes the process of cell adhesion and spreading and binds to several protease inhibitors (serpins). It is secreted from cells and is believed to

play roles in blood clotting and the malignancy of tumors. One domain of vitronectin binds to plasminogen activator inhibitor and acts to stabilize it. Another domain of the protein binds to cellular integrin proteins, such as the vitronectin receptor that anchors cells to the extracellular matrix.

Catenins

Catenins are a family of proteins interacting with cadherin proteins in cell adhesion (Figure 2.69). Four main types of catenins are known, α-, β-, γ-, and δ-catenin. Catenins play roles in cellular organization before development occurs and help to regulate cellular growth. α-catenin and β-catenin are found at adherens junctions with cadherin and help cells to maintain epithelial layers. Cadherins are connected to actin filaments of the cytoskeleton and catenins play the critical role. Catenins are important for the process whereby cellular division is inhibited when cells come in contact with each other (contact inhibition).

When catenin genes are mutated, cadherin cell adhesions can disappear and tumorigenesis may result. Catenins have been found to be associated with colorectal and numerous other forms of cancer.

Glycophorins

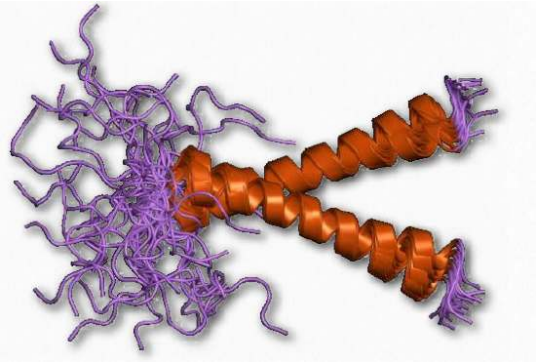


Figure 2.80 - Glycophorin a

All of the membrane proteins described so far are notable for the connections they make to other proteins and cellular structures. Some membrane proteins, though, are designed to reduce cellular connections to proteins of other cells. This is particularly important for blood cells where “stickiness” is undesirable except where clotting is concerned.

Glycophorins (Figure 2.80) are membrane-spanning sialoglycoproteins of red blood cells. They are heavily glycosylated (60%).and rich in sialic acid, giving the cells a very hydrophilic (and negatively charged) coat, which enables them to circulate in the bloodstream without adhering to other cells or the vessel walls.

Five glycophorins have been identified - four (A,B,C,and D) from isolated membranes and a fifth form (E) from coding in the human genome. The proteins are abundant, forming about 2% of the total membrane proteins in these cells. Glycophorins have important roles in regulating RBC membrane mechanical properties and shape. Because some glycophorins can be expressed in various nonerythroid tissues (particularly Glycophorin C), the importance of their interactions with the membrane skeleton may have a considerable biological significance.

Cooperativity and allosterism - quaternary structure

Figure 2.81 - Two polypeptide units of a protein interact in quaternary structure Wikipedia

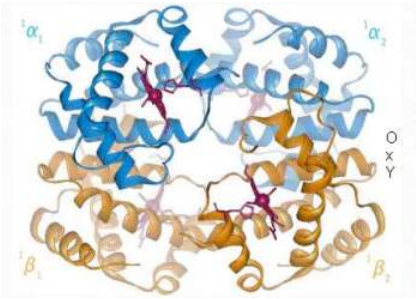
Quaternary structure, of course describes the interactions of individual subunits of a multi-subunit protein (Figure 2.81). The result of these interactions can give rise to important biological phenomena, such as cooperative binding of substrates to a protein and allosteric effects on the action of an enzyme.

Allosteric effects can occur by a series of mechanisms, but a common feature is that binding of an effector to an enzyme subunit causes (or locks) the enzyme in either a Tstate (less activity) or an R-state (more activity). Effectors can be enzyme substrates (homotropic effectors) or non-substrates (heterotropic effectors). Allosterism will be covered in more depth in the Catalysis chapter HERE.

We begin our consideration of quaternary structure with a discussion of cooperativity, how it arises in the multi-subunit protein hemoglobin and how its properties contrast with those of the related, single subunit protein myoglobin.

Cooperativity

Cooperativity is defined as the phenomenon where binding of one ligand molecule by a protein favors the binding of additional molecules of the same type. Hemoglobin, for example, exhibits cooperativity when the binding of an oxygen molecule by the iron of the heme group in one of the four subunits causes a slight conformation change in the subunit. This happens because the heme iron is attached to a histidine side chain and binding of oxygen ‘lifts’ the iron along with the histidine ring (also known as the imidazole ring).



Movie 2.3 - Hemoglobin’s structural changes on binding oxygen Wikipedia

Since each hemoglobin subunit interacts with and influences the other subunits, they too are induced to change shape slightly when the first subunit binds to oxygen (a transition described as going from the T-state to the R-state). These shape changes favor each of the remaining subunits binding oxygen, as well. This is very important in the lungs where oxygen is picked up by hemoglobin, because the binding of the first oxygen molecule facilitates the rapid uptake of more oxygen molecules. In the tissues, where the oxygen concentration is lower, the oxygen leaves hemoglobin and the proteins flips from the R-state back to the Tstate.

CO_2 transport

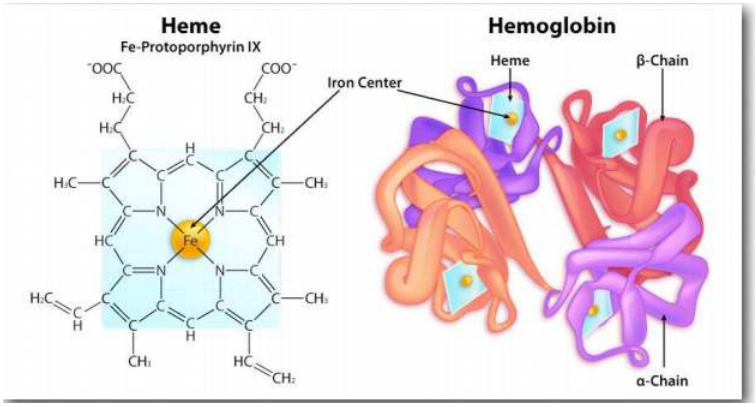


Figure 2.82 - Heme structure within hemoglobin Image by Aleia Kim

Cooperativity is only one of many fascinating structural aspects of hemoglobin that help the body to receive oxygen where it is needed and pick it up where it is abundant. Hemoglobin also assists in the transport of the product of cellular respiration (carbon dioxide) from the tissues producing it to the lungs where it is exhaled. Like the binding of oxygen to hemoglobin, binding of other molecules to hemoglobin affects its affinity for oxygen. The effect is particularly pronounced when comparing the oxygen binding characteristics of hemoglobin's four subunits with the oxygen binding of the related protein myoglobin's single subunit (Figure 2.83).

Different oxygen binding

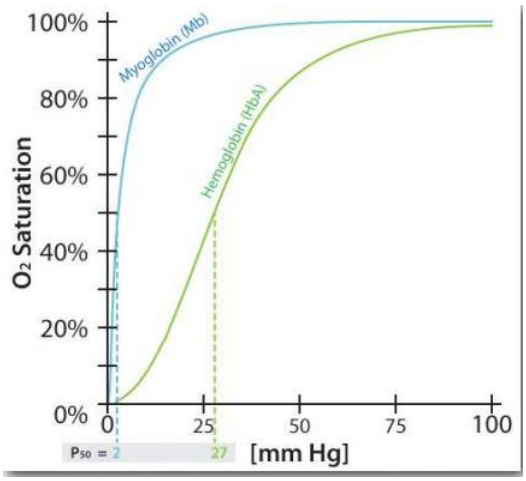


Figure by Aleia Kim

Like hemoglobin, myoglobin contains an iron in a heme group that binds to oxygen. The structure of the globin protein in myoglobin is very similar to the structure of the globins in hemoglobin and hemoglobin is thought to have evolved from myoglobin in evolutionary history. As seen in Figure 2.83, the binding curve of hemoglobin for oxygen is S-shaped (sigmoidal), whereas the binding curve for myoglobin is hyperbolic. What this tells us is that hemoglobin's affinity for oxygen is low at a low concentration oxygen, but increases as the oxygen concentration increases. Since myoglobin very quickly saturates with oxygen, even under low oxygen concentrations, it says that its affinity for oxygen is high and doesn't change.

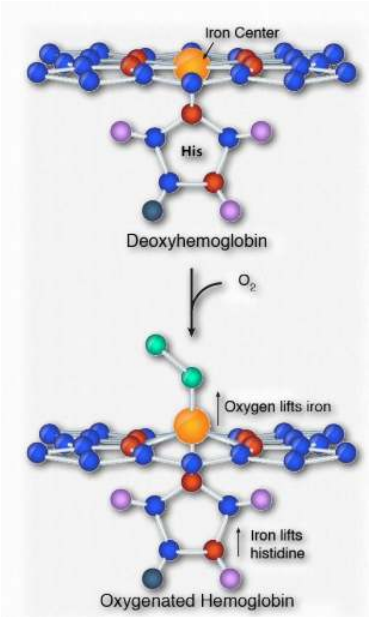
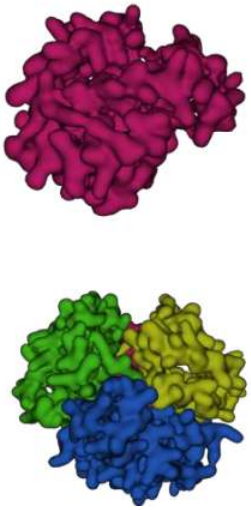


Figure 2.84 - Binding of oxygen at the heme center of hemoglobin Image by Aleia Kim

Because myoglobin has only a single subunit, binding of oxygen by that subunit can't affect any other subunits, since there are no other subunits to affect. Consequently, cooperativity requires more than one subunit. Therefore, hemoglobin can exhibit cooperativity, but myoglobin can't. It is worth noting that simply having multiple subunits does not mean cooperativity will exist. Hemoglobin is one protein that exhibits the characteristic, but many multisubunit proteins do not.



Interactive 2.2 - Hemoglobin in the presence (top) and absence (bottom) of oxygen

Storage vs. delivery

The lack of ability of myoglobin to adjust its affinity for oxygen according to the oxygen concentration (low affinity at low oxygen concentration, such as in tissues and high affinity at high oxygen concentration, such as in the lungs) means it is better suited for storing oxygen than for delivering it according to the varying oxygen needs of and animal body. As we shall see, besides cooperativity, hemoglobin has other structural features that allow it to deliver oxygen precisely where it is needed most in the body.

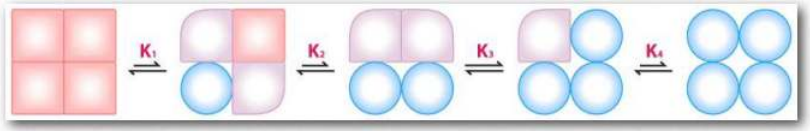


Figure 2.85 - Sequential model of binding. The sequential model is one way to explain hemoglobin's cooperativity. Squares represent no oxygen bound. Circles represent subunits bound with oxygen and rounded subunits correspond to units whose affinity for oxygen increases by interacting with a subunit that has bound oxygen. Image by Aleia Kim 131

Bohr effect

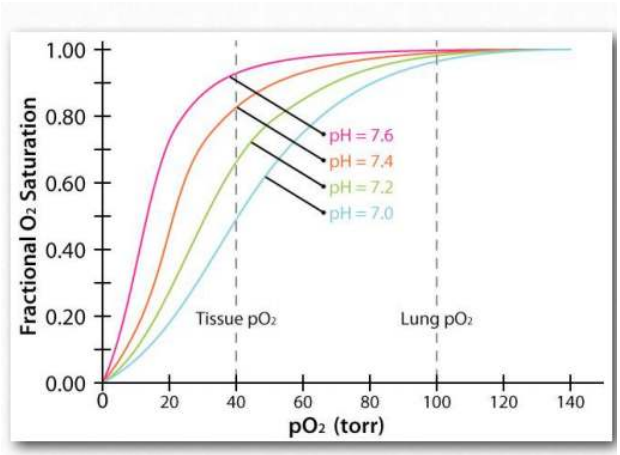


Figure 2.86 - The Bohr effect with respect to pH changes Image by Aleia Kim

The Bohr Effect was first described over 100 years ago by Christian Bohr, father of the famous physicist, Niels Bohr. Shown graphically (Figures 2.86, 2.87, and 2.88), the observed effect is that hemoglobin's affinity for oxygen decreases as the pH decreases and as the concentration of carbon dioxide increases. Binding of the protons and carbon dioxide by amino Figure 2.85 - Sequential model of binding. The sequential model is one way to explain hemoglobin's cooperativity. Squares represent no oxygen bound. Circles represent subunits bound with oxygen and rounded subunits correspond to units whose affinity for oxygen increases by interacting with a subunit that has bound oxygen. Image by Aleia Kim acid side chains in the globin proteins helps to facilitate structural changes in them. Most commonly, the amino acid affected by protons is histidine #146 of the β strands. When this happens, the ionized histidine can form an ionic bond with the side chain of aspartic acid #94, which has the effect of stabilizing the T-state (reduced oxygen binding state) and releasing oxygen. Other histidines and the amine of the amino terminal amino acids in the α-chains are also binding sites for protons.

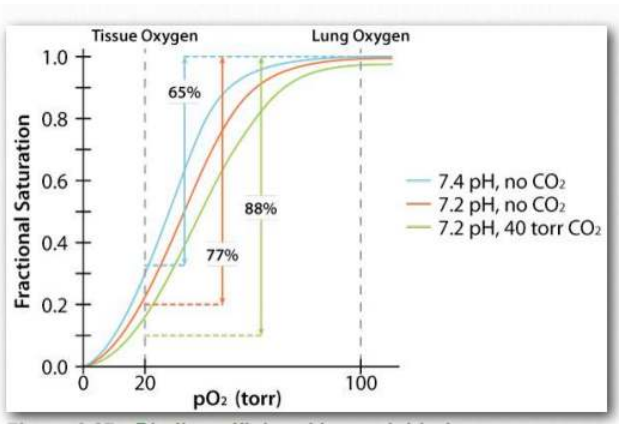


Figure 2.87 - Binding affinity of hemoglobin for oxygen under different conditions Image by Aleia Kim

2,3-BPG

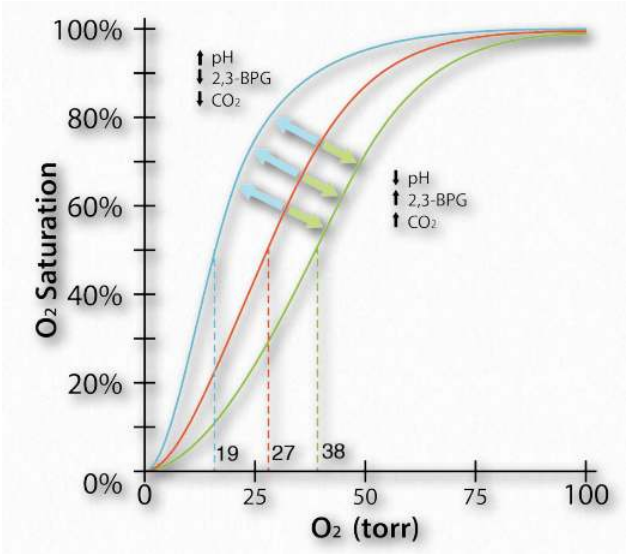


Figure 2.88 - The Bohr effect physiologically - oxygen binding curves for resting muscle (blue), active muscle (green) and reference muscle (orange) with respect to pH, 2,3-BPG, and CO2 Image by Aleia Kim

Another molecule favoring the release of oxygen by hemoglobin is 2,3- bisphosphoglycerate (also called 2,3-BPG or just BPG - Figure 2.89). Like protons and carbon dioxide, 2,3-BPG is produced by actively respiring tissues, as a byproduct of glucose metabolism. The 2,3-BPG mole cule fits into the 'hole of the donut' of adult hemoglobin (Figure 2.89). Such binding of 2,3-BPG favors the T-state (tight - low oxygen binding) of hemoglobin, which has a reduced affinity for oxygen. In the absence of 2,3-BPG, hemoglobin can more easily exist in the R-state (relaxed - higher oxygen binding), which has a high affinity for oxygen.

Smokers

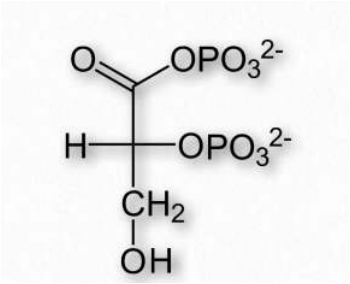


Figure 2.89 - The structure of 2,3 bisphosphoglycerate (2,3-BPG)

Notably, the blood of smokers is higher in the concentration of 2,3-BPG than non-smokers, so more of their hemoglobin remains in the T-state and thus the oxygen carrying capacity of smokers is lower than non-smokers.Another reason why smokers' oxygen carrying capacity is lower than that of non-smokers is that cigarette smoke contains carbon monoxide and this molecule, which has almost identical dimensions to molecular oxygen, effectively outcompetes with oxygen for binding to the iron atom of heme (Figure 2.90). Part of carbon monoxide's toxicity is due to its ability to bind hemoglobin and prevent oxygen from binding.

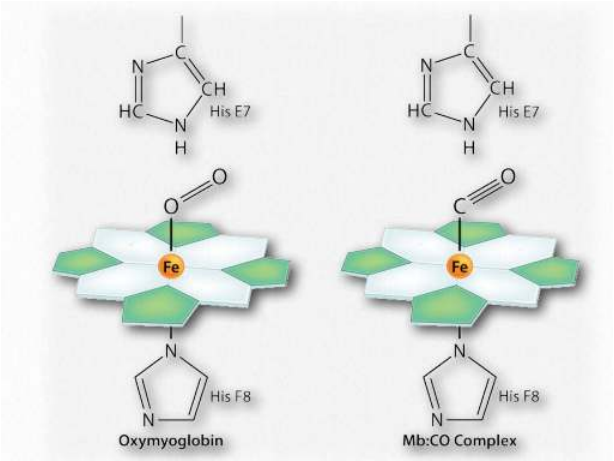


Figure 2.90 - Binding of oxygen (left) and carbon monoxide (right) by a heme group of hemoglobin Image by Aleia Kim

Carbon dioxide

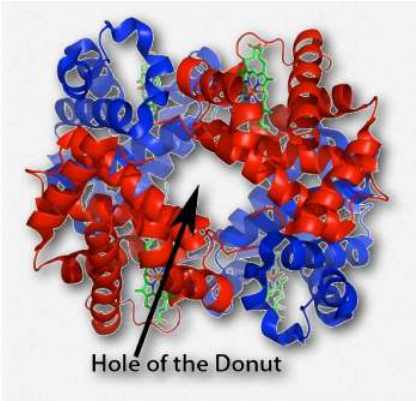


Figure 2.91 - Hemoglobin's hole of the donut for binding 2,3-BPG Wikipedia

Carbon dioxide binds to form a carbamate when binding the α-amine of each globin chain. The process of forming this structure releases a proton, which helps to further enhance the Bohr effect. Physiologically, the binding of CO2 and H+ has significance because actively respiring tissues (such as contracting muscles) require oxygen and release protons and carbon dioxide. The higher the concentration of protons and carbon dioxide, the more oxygen is released to feed the tissues that need it most.

About 40% of the released protons and about 20% of the carbon dioxide are carried back to the lungs by hemoglobin. The remainder travel as part of the bicarbonate buffering system or as dissolved CO2. In the lungs, the process reverses itself. The lungs have a higher pH than respiring tissues, so protons are released from hemoglobin and CO2 too is freed to be exhaled.

Fetal hemoglobin

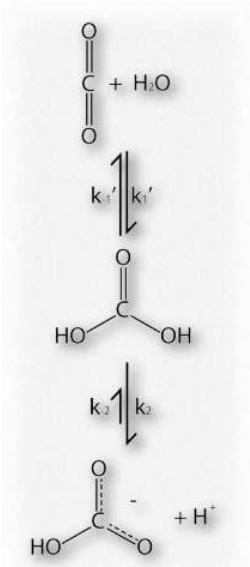


Figure 2.92 - Formation of bicarbonate from CO2 in blood

Adult hemoglobin releases oxygen when it binds 2,3- BPG. This is in contrast to fetal hemoglobin, which has a slightly different configuration (α2γ2) than adult hemoglobin (α2β2). Fetal hemoglobin has a greater affinity for oxygen than maternal hemoglobin, allowing the fetus to obtain oxygen effectively from the mother's blood. Part of the reason for fetal hemoglobin's greater affinity for oxygen is that it doesn't bind 2,3-BPG. Consequently, fetal hemoglobin remains in the R-state much more than adult hemoglobin and because of this, fetal hemoglobin has greater affinity for oxygen than adult hemoglobin and can take oxygen away from adult hemoglobin. Thus, the fetus can get oxygen from the mother.

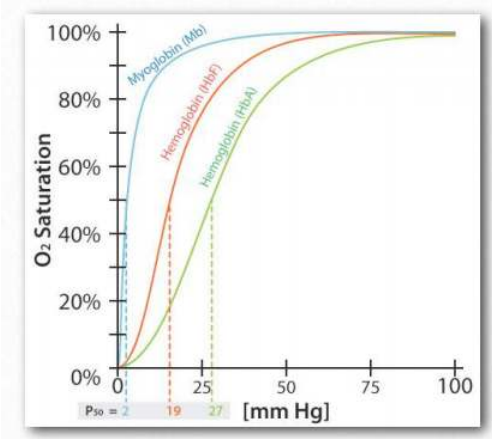


Figure 2.93 - Comparison of oxygen binding of myoglobin (blue), fetal hemoglobin (orange), and adult hemoglobin (green) Image by Aleia Kim

Sickle cell disease

Figure 2.94 - Four normal red blood cells (right) and one sickled red blood cell (left) Wikipedia

Mutations to the globin genes coding for hemoglobin can sometimes have deleterious consequences. Sickle cell disease (also called sickle cell anemia) is a genetically transmitted disease that arises from such mutations. There are different forms of the disease. It is a recessive trait, meaning that to be afflicted with it, an individual must inherit two copies of the mutated gene.

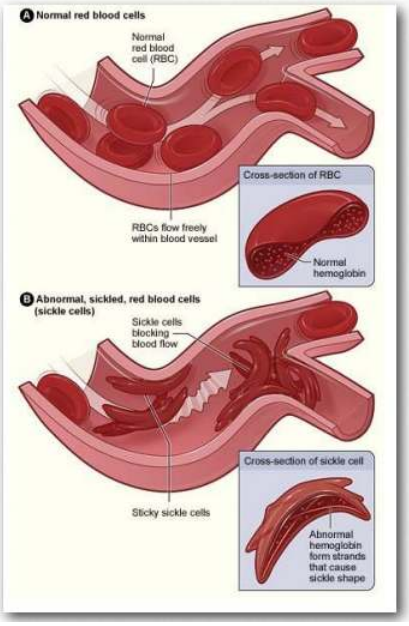


Figure 2.95 - Movement of blood in capillaries. Top - normal red blood cells. Bottom - sickled red blood cells

The predominant form of hemoglobin in adults is hemoglobin A, designated HbA (two α chains and two β chains). The mutant form is known as HbS. The most common mutation is an A to T mutation in the middle of the codon for the seventh amino acid (some counting schemes call it the sixth amino acid) of the β-chain. This results in conversion of a GAG codon to GTG and thus changes the amino acid specified at that position from a glutamic acid to a valine. This minor change places a small hydrophobic patch of amino acids on the surface of the β-globin chains.

Polymerization

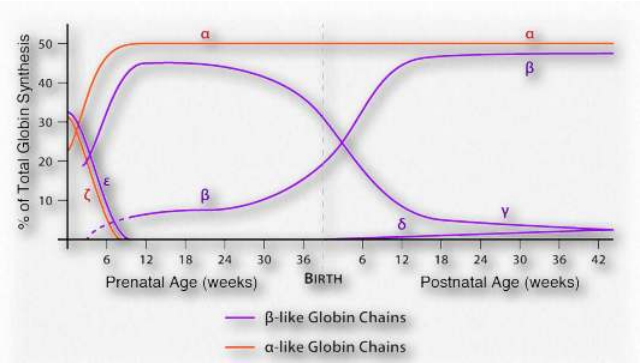


Figure 2.96 - Pattern of expression of six different globins of hemoglobin - α,β,γ,ε,δ, and ζ Image by Aleia Kim

Under conditions of low oxygen, these hydrophobic patches will associate with each other to make long polymers of hemoglobin molecules. The result is that the red blood cells containing them will change shape from being rounded to forming the shape of a sickle (Figure 2.94). Rounded red blood cells readily make it through tiny capillaries, but sickleshaped cells do not.

Worse, they block the flow of other blood cells. Tissues where these blockages occur are already low in oxygen, so stopping the flow of blood through them causes them to go quickly anaerobic, causing pain and, in some cases, death of tissue. In severe circumstances, sickled red blood cells death may result. The disease is referred to as an anemia because the sickling of the red blood cells targets them for removal by the blood monitoring system of the body, so a person with the disease has chronically reduced numbers of red blood cells.

Heterozygote advantage

Interestingly, there appears to be a selective advantage to people who are heterozygous for the disease in areas where malaria is prominent. Heterozygotes do not suffer obvious ill effects of the disease, but their red blood cells appear to be more susceptible to rupture when infected. As a consequence, the parasite gets less of a chance to reproduce and the infected person has a greater chance of survival.

The protective effect of the mutant gene, though, does not extend to people who suffer the full blown disease (homozygotes for the mutant gene). Treatments for the disease include transfusion, pain management, and avoidance of heavy exertion. The drug hydroxyurea has been linked to reduction in number and severity of attacks, as well as an increase in survival time1,2. It appears to work by reactivating expression of the fetal hemoglobin gene, which typically is not synthesized to any significant extent normally after about 6 weeks of age.

Oxygen binding

Animals have needs for oxygen that differ from all other organisms. Oxygen, of course, is the terminal electron acceptor in animals and is necessary for electron transport to work. When electron transport is functioning, ATP generation by cells is many times more efficient than when it is absent. Since abundant ATP is essential for muscular contraction and animals move around a lot - to catch prey, to exercise, to escape danger, etc., having an abundant supply of oxygen is important.

This is particularly a concern deep inside tissues where diffusion of oxygen alone (as occurs in insects) does not deliver sufficient quantities necessary for long term survival. The issue is not a problem for plants since, for the most part, their motions are largely related to growth and thus don’t have rapidly changing needs/demands for oxygen that animals have. Unicellular organisms have a variety of mechanisms for obtaining oxygen and surviving without it. Two other important oxygen binding proteins besides hemoglobin are myoglobin and hemocyanin.

Myoglobin

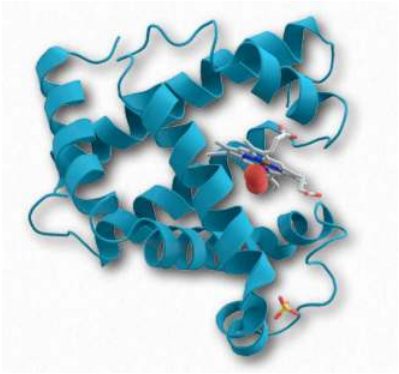


Figure 2.97 - Myoglobin bound to oxygen

Myoglobin is the primary oxygen-storage protein found in animal muscle tissues. In contrast to hemoglobin, which circulates throughout the body, myoglobin protein is only found in muscle tissue and appears in the blood only after injury. Like hemoglobin, myoglobin binds oxygen at a prosthetic heme group it contains.

The red color of meat arises from the heme of myoglobin and the browning of meat by cooking it comes from oxidation of the ferrous (Fe++) ion of myoglobin's heme to the ferric (Fe+++) ion via oxidation in the cooking process. As meat sits in our atmosphere (an oxygen-rich environment), oxidation of Fe++ to Fe+++ occurs, leaving the brown color noted above. If meat is stored in a carbon monoxide (CO) environment, CO binds to the heme group and reduces the amount of oxidation, keeping meat looking red for a longer period of time.

High affinity

Myoglobin (Figure 2.97) displays higher affinity for oxygen at low oxygen concentrations than hemoglobin and is therefore able to absorb oxygen delivered by hemoglobin under these conditions. Myoglobin's high affinity for oxygen makes it better suited for oxygen storage than delivery. The protein exists as a single subunit of globin (in contrast to hemoglobin, which contains four subunits) and is related to the subunits found in hemoglobin. Mammals that dive deeply in the ocean, such as whales and seals, have muscles with particularly high abundance of myoglobin. When oxygen concentration in muscles falls to low levels, myoglobin releases its oxygen, thus functioning as an oxygen "battery" that delivers oxygen fuel when needed and holding onto it under all other conditions. Myoglobin holds the distinction of being the first protein for which the 3D structure was determined by X-ray crystallography by John Kendrew in 1958, an achievement for which he later won the Nobel Prize.

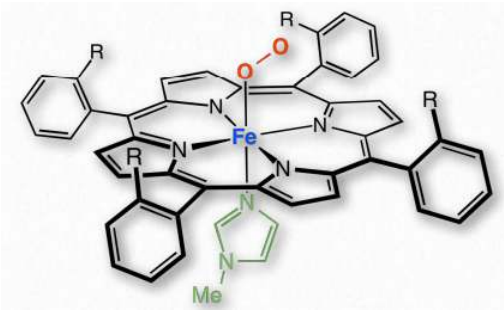


Figure 2.98 - Oxygen bound at heme of myoglobin Wikipedia

Hemocyanin

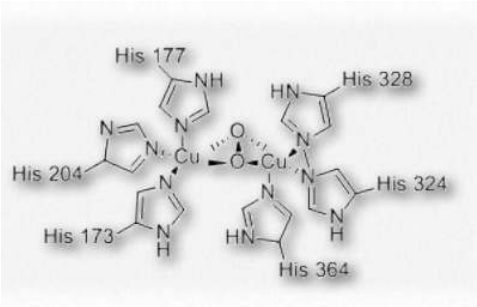


Figure 2.99 - Oxygen binding in hemocyanin Wikipedia

Hemocyanin is the protein transporting oxygen in the bodies of molluscs and arthropods. It is a coppercontaining protein found not within blood cells of these organisms, but rather is suspended in the circulating hemolymph they possess. The oxygen binding site of hemocyanin contains a pair of copper(I) cations directly coordinated to the protein by the imidazole rings of six histidine side chains.

Figure 2.100 - Hemocyanin (purple) in a red rock crab Wikipedia

Most, but not all hemocyanins bind oxygen non-cooperatively and are less efficient than hemoglobin at transporting oxygen. Notably, the hemocyanins of horseshoe crabs and some other arthropods do, in fact, bind oxygen cooperatively. Hemocyanin contains many subunit proteins, each with two copper atoms that can bind one oxygen molecule (O2). Subunit proteins have atomic masses of about 75 kilodaltons (kDa). These may be arranged in dimers or hexamers depending on species. Superstructures comprised of dimer or hexamer complexes are arranged in chains or clusters and have molecular weights of over 1500 kDa.

This page titled 4.6: Protein function, domains and cooperativity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **2.4: Structure and Function- Proteins II** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **4.5: Protein function, domains and cooperativity** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

SECTION OVERVIEW

4.7: The Three-Dimensional Structure of Proteins

4.7.1: Main Chain Conformations

4.7.2: Secondary Structure and Loops

4.7.3: Tertiary, Quaternary, and Symmetrical Structures

4.7.4: Secondary Structural Motifs and Domains

4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure

4.7.6: Intrinsically Disordered Proteins

4.7.7: Fibrillar Proteins

4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics

4.7.9: Protein Stability - Thermodynamics

4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules

4.7.11: Biomolecular Visualization - Conceptions and Misconceptions

4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation

4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools

4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24)

4.7.15: Chapter 4 Questions

This page titled 4.7: The Three-Dimensional Structure of Proteins is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

4.7.1: Main Chain Conformations

Search Fundamentals of Biochemistry

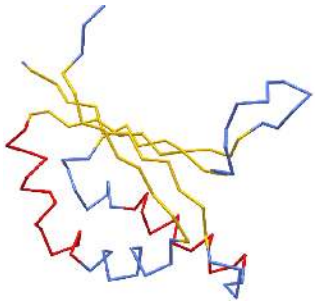
Learning Goals (ChaptGPT o1, 1/25/25)

- **Understand Basic Conformational Principles:**
  - Explain the concept of conformers using simple molecules (e.g., butane) and interpret Newman projections to compare energy differences between staggered (anti) and eclipsed (syn) conformations.
  - Relate the idea of dihedral angles in small hydrocarbons to the rotation around bonds in longer chains, such as in dodecanoic acid and peptide backbones.
- **Connect Simple Organic Conformations to Peptide Structure:**
  - Compare the extended, low-energy "all-trans" (anti) conformation in nonpolar chains with the typical zig-zag arrangement found in protein backbones.
  - Describe how dihedral angles (phi and psi) dictate the conformation of the peptide main chain, and distinguish between the conformations typical of beta sheets (near 180°) and alpha helices (approximately 60°).
- **Explore Peptide Bond Isomerism and Its Consequences:**
  - Differentiate between trans and cis peptide bonds and explain why the trans configuration is highly favored (except in X-Pro bonds) due to reduced steric hindrance and torsional strain.
  - Analyze how the presence of a cyclic amino acid like proline affects the cis/trans ratio in peptide bonds.
- **Master Backbone Dihedral Angles and Ramachandran Plots:**
  - Define the phi (φ) and psi (ψ) dihedral angles and explain their significance in determining the overall backbone conformation.
  - Interpret Ramachandran plots to identify allowed and disallowed regions of φ/ψ space, and relate these regions to common secondary structures (alpha helices, beta sheets, and turns).
  - Discuss how the planarity of the peptide bond (due to resonance) restricts the conformational freedom of the protein backbone.
- **Apply Computational and Experimental Approaches:**
  - Understand how experimental models (e.g., interactive iCn3D models) and computational tools (e.g., force fields using Lennard-Jones potentials) aid in visualizing and predicting backbone conformations.
  - Recognize the importance of considering both steric and electronic factors in modeling peptide conformations and their implications for protein folding and function.

These learning goals will enable students to integrate fundamental organic chemistry concepts with protein structure analysis, providing a strong foundation for understanding how backbone conformation influences the secondary, tertiary, and quaternary structures of proteins.

4.7.1.1: Introduction

In Chapter 3.2, we discussed primary, secondary, tertiary, and quaternary structures of proteins. We used the protein **hydroxynitrile lyase** (5Y02) as an example. An interactive iCn3D model of hydroxynitrile lyase is shown in Figure **4.7.1.1**.



NCBI iCn3D Figure 4.7.1.1: hydroxynitrile lyase (5Y02). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?gGY8QBkCP8dPmfoH6 (Copyright; author via source)

Blue indicates random coil, red indicates alpha helices, and yellow shows beta strands. As discussed in Chapter 3.2, the beta-strand backbones are quite extended compared to the alpha helices, which are more compressed. Atomic force microscopy can be used to pull the helix into alternative extended shapes. Let's explore main chain conformations in greater detail.

If we wish to understand protein structure, it is essential first to understand the possible conformations of the backbone before we consider the effects of side chains and the overall structure of the protein. Where do we start? It's a highly complex topic, but we can simplify it by considering the structures of short hydrocarbons and other organic groups, and by drawing on ideas you learned in earlier chemistry courses. So, let's start with butane, a 4C hydrocarbon.

You will remember from previous chemistry courses that different rotational conformations (conformers) of butane have different energies. These conformers are often visualized using Newman projections, which are viewed down the internal C2-C3 carbon bond in butane. The figure below shows ball and stick representations and Newman projections for three different conformers formed by rotation around the C2-C3 bond. You are familiar with bond angles —the angles between two bonds connecting three atoms. For this discussion, we will consider the dihedral angle, which describes the rotation angle around the middle bond of a 4-atom, 3-bond connected system. It's easiest to use Newman projections to understand dihedrals, which represent the angle between the groups bonded to the front and back (circle) atoms, viewed from the front-to-back direction. Figure 4.7.1.2 shows multiple conformers of butane, represented in ball and stick form or as Newman projections.

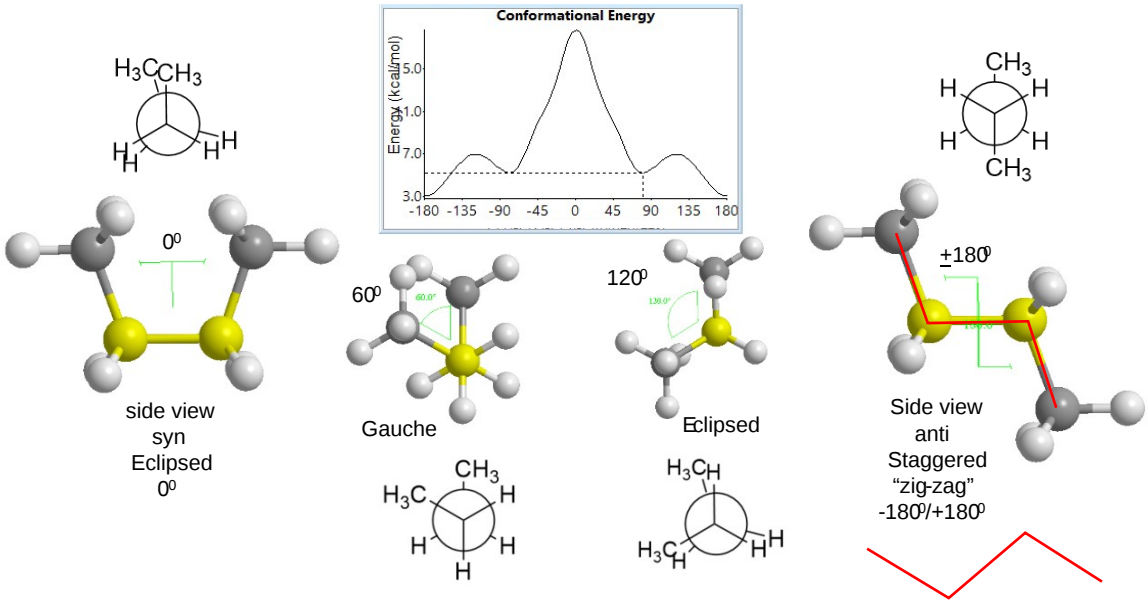


Figure 4.7.1.2: Newman projections for butane

The syn and eclipsed conformers with a dihedral angle of 0° have the highest energy, since the two eclipsed bonds have high torsional energy and the two closest-approaching H atoms on the methyl C atoms have steric (1,6) interactions. Rotating the front C 60° gives a gauche conformation of lower energy. Rotating to 120° results in higher energy due to the torsional strain between the

eclipsed bonds. Rotating fully to 180° gives the conformer of lowest energy and torsional (bond-bond) strain. The steric strain between the two closest H atoms on the methyl groups also raises the energy.

A red line has been drawn between the four carbon atoms connected by three bonds in the anti or trans-conformer. Notice that the red line is a zig-zag, as shown in the rotated red line drawing beneath the ball and stick conformer. All the conformers are available, but the anti or trans-conformer is the most stable and abundant.

Let's try this approach on a more complicated molecule, a 12-carbon atom carboxylic acid, dodecanoic acid. Figure 4.7.1.3 shows the molecule in the all- trans conformation (where all the carbon atoms in the chain are arranged in an anti, zig-zag manner), and a second gauche conformation for the bond highlighted by a red line. Note that rotation around that bond to produce the gauche conformer, a higher-energy form, introduces a kink in the chain. We will see this again when we discuss lipid structure.

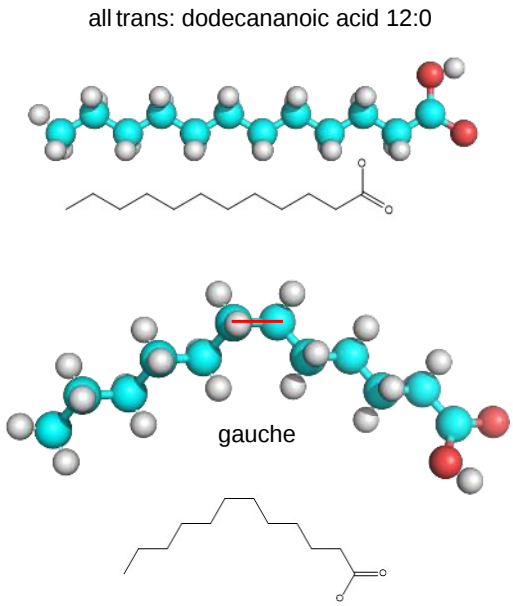


Figure 4.7.1.3: Conformations of dodecanoic acid

Now let's compare the all-trans dodecanoic acid structure to one largely extended conformer of a tetrapeptide, Gly-Gly-Ala-Gly (GGAG), as shown in Figure 4.7.1.4.

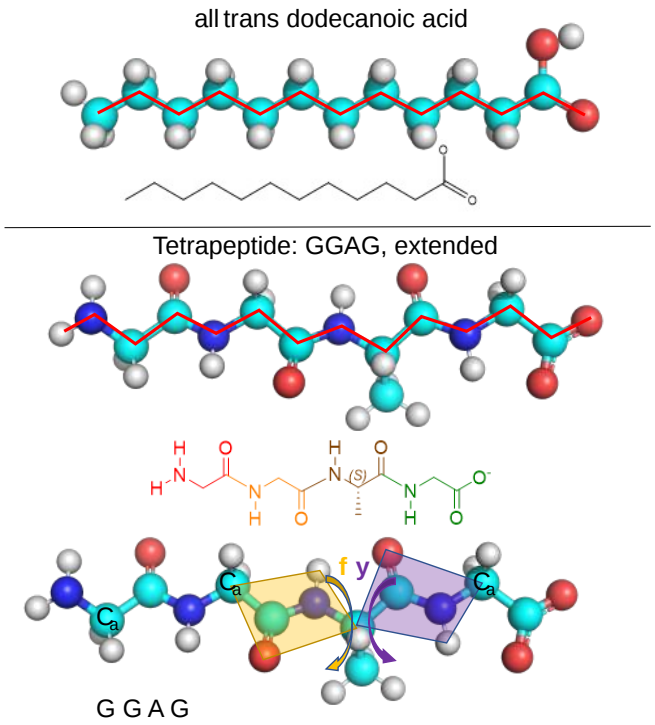


Figure 4.7.1.4: Extended conformations of dodecanoic acids and a tetrapeptide

Notice the similarities in these two structures. The atoms in the main chain of the peptide, N, C_α, C (the carbonyl C), and the carbonyl O, are arranged in the familiar zig-zag fashion, characteristic of the lowest energy conformer. The dihedral angles along the backbone would be, to a first approximation, around $\pm$ 180° depending on whether the rotation is clockwise or counterclockwise from the syn conformer. However, the actual conformation of the GGAG peptide, if found as part of the primary structure of a protein, could adopt a whole range of other conformers with potentially different dihedral angles for each bond in the main chain. Look back at the iCn3D model for **hydroxynitrile lyase** above, and it should be clear that if the stretch of amino acids Gly-Gly-Ala-Gly is in a beta sheet, its dihedral angles for the main chain would be closer to $\pm 180^0$ than if they were in an alpha helix, which would have dihedral angles much less than 180°, perhaps closer to $\pm 60^0$

It might seem like all dihedral angles are possible for the main chain atoms of a protein, which would make understanding protein structure too complicated for our minds to comprehend. However, two significant factors simplify conformational analyses of proteins.

4.7.1.2: Trans and Cis Peptide Bonds

In the extended peptide example above, the alpha Cs of adjacent amino acids are on opposite sides of the peptide bond between the carbonyl carbon atom and the amide nitrogen atom. They are **trans** to that bond. A rectangle was drawn on the above figure centered on the C-N peptide bond for two adjacent amino acids to show the trans orientation of the alpha Cs.

Rotation around single bonds can occur, so the trans peptide bond can isomerize to the cis form. Both isomeric forms of the dipeptide Ala-Ala are shown in the top part of Figure 4.7.1.5.

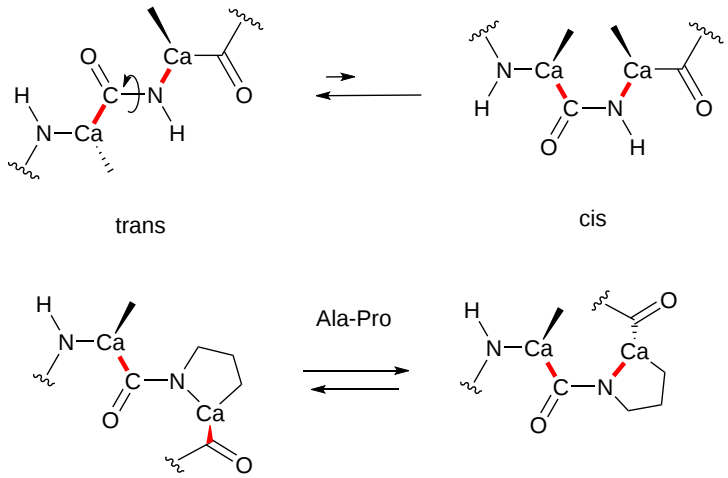


Figure 4.7.1.5: Cis/Trans peptide bonds

This trans arrangement of the alpha Cs in Ala-Ala and most internal X-Y amino acid pairs in a protein is sterically favored by a factor of 1000/1 over the cis form, clearly showing torsional strain between the aligned (red) bonds and steric strain between amino acid side-chain atoms.

The case is different for X-Y amino acid pairs when Y is proline, as shown in the bottom part of the above figure. Proline is a cyclic amino acid and, as such, would be expected to be more sterically restricted in a protein sequence. In X-Pro peptide bonds in proteins, the trans/cis ratio is about 4/1. The torsional (between bonds) and steric strain (between atoms) are similar in both isomers. You would expect to find both trans and cis X-Pro peptide bonds in proteins, given the constraints placed on protein conformation in the tertiary structure of the protein.

Figure 4.7.1.6 is an interactive iCn3D model showing two pairs of X-Y peptide bonds in the protein carboxypeptidase. Answer the questions below about the two peptide bond pairs, shown in stick form and CPK colors.

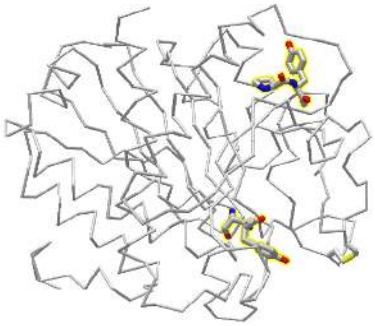


Figure 4.7.1.6: 2 peptide bonds in carboxypeptidase. Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3ARF97MgN4427A (Copyright; author via source)

Exercise 4.7.1.1

Two X-Y amino acids are shown in stick form in the protein carboxypeptidase A

1. What is the sequence of each pair shown from N to the C terminus

2. Are the X-Y peptide bonds cis or pro?

Answer

Add texts here. Do not delete this text first.

4.7.1.3: Backbone dihedral angles and Ramachandran Plots

Let's focus on possible dihedral angles for the peptide bond in protein chains. Just as saturated fatty acid chains have preferred conformations (all anti or trans), peptide chains also have preferred conformations. The complexity is much greater, however. In our dodecanoic acid example, we considered only torsional or dihedral angles around the methylene ($-CH_2-$) carbons. For proteins, we must consider the covalent bonds that link amino acids and the rotations possible for the 20 amino acids. The peptide bond connects the carbonyl C of the ith amino acid to the alpha amine N of the i+1th amino acid. The resulting bond is an amide link. X-ray analysis shows that the C-N bond has double-bond character. This can be accounted for by delocalizing the nonbonding electron pair of the N to the carbonyl C, forming a double bond, with the pi-bonded electrons of the carbonyl C-O bond moving to the O. These resonance structures lead to a planar arrangement of the peptide carbonyl C and amide N and the two other atoms connected to each, since the hybridization of the C and N has sp^2 character, with 120° bond angles, as illustrated in Figure **4.7.1.7**.

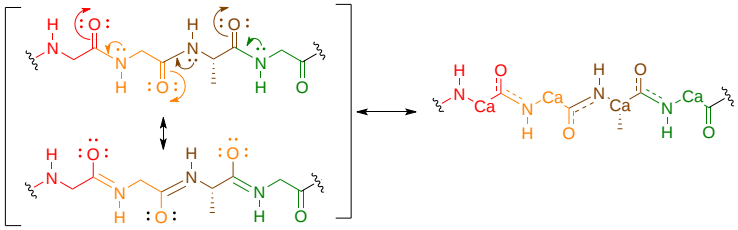


Figure **4.7.1.7**: Resonance forms of peptide bond

This greatly simplifies the number of conformations a protein can adopt since these six atoms can be considered to reside and move in a plane. The alpha C is the corner attachment point of two different planes, which can rotate independently of each other. The two planes can twist around the alpha carbon. The rotation angles for the two planes are called phi (φ), and psi (ψ) are analogous to the torsion angles in the acyl chain of dodecanoic acid. They can vary from -180 to +180 °. The R group substituent attached to the alpha C can also rotate around the side chain's alpha C and the beta C. This angle is defined as chi. Other rotations also occur within the side chain. We will concentrate on phi (φ) and psi(ψ) angles.

Figure **4.7.1.8** shows the peptide Gly-Gly-Ala-Gly (at high pH to give the protonation state shown) with the six atoms around the C-N peptide bonds shown in rectangles. Phi (φ) and psi (ψ) angles are also shown.

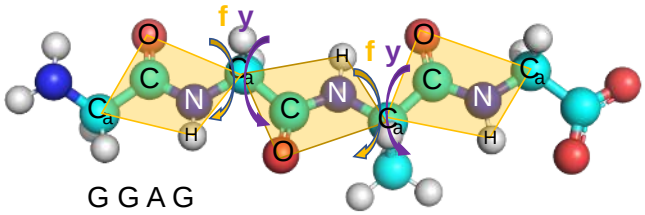


Figure **4.7.1.8**: Phi/Psi angles for Gly-Gly-Ala-Gly

Using these ideas, we can consider a protein backbone a series of linked sequences of rigid, planar peptide units that rotate around phi/psi angles. When the chain is fully extended (as shown in the links above), φ/ψ are ±180 °, with the syn eclipse form defined as having 0^0 φ/ψ angles. When φ and ψ equal 0°, the two peptide bonds flanking the alpha Cs are in the same plane. This conformation is prohibited since the O of the C=O on one plane and the H of the H-N on the other overlap - i.e., they approach closer than their van der Waals radii.

The Proteopedia page below gives a phenomenal explanation and visual representation of φ/ψ angles in proteins.

Proteopedia

:

Ramachandran principle and phi psi angles

Figure **4.7.1.9** is an animated gif showing rotation of phi and psi angles (made from Proteopedia link above). You can see that when either φ or ψ approaches 0^0, steric clashes are most significant.

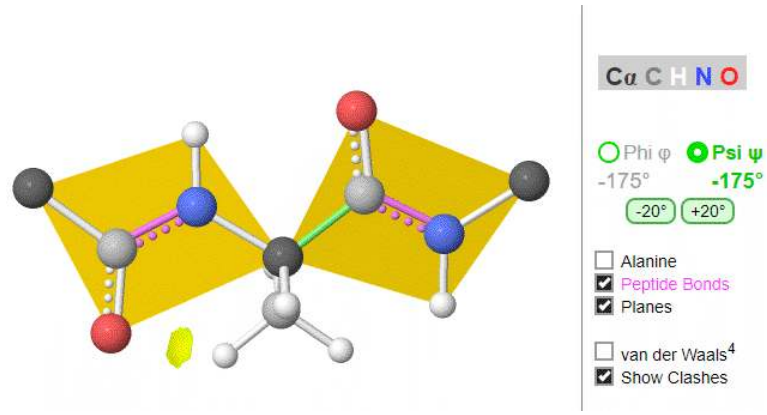


Figure 4.7.1.9: Animation of rotation around Phi/Psi angles f

Ramachandran was the first to calculate which φ/ψ angles are allowed. He modeled the angles permitted to a tripeptide, assuming the atoms were hard spheres. The angles allowed depended in part on the limiting distance chosen for interatomic contacts. (i.e., the usual H -- H distance is 2.0 Å and 3.0 Å for C--C bonds.) The plot below shows the allowed regions in red. Only three small regions of conformational space are available. If you allow a closer approach by 0.1 angstroms, more conformation space is available, but only one new area is available, shown in yellow in Figure 4.7.1.10.

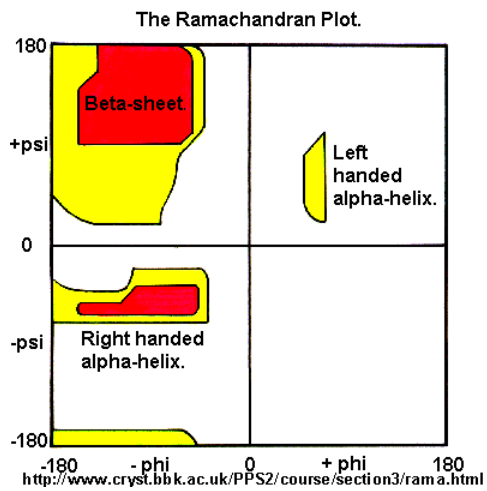


Figure 4.7.1.10: Ramachandran plot. http://www.cryst.bbk.ac.uk/PPS95/cou...etry/rama.html

Right-hand alpha helices fall at -570,-470, while left-hand alpha helices fall at +570,+470. (Notice these are not mirror images of each other. The mirror image of a right-handed alpha helix would be a left-handed helix made of D-amino acids.) Parallel beta sheets are at -119^0, -113^0, while antiparallel sheets fall at -139^0, +135^0. Note that beta sheets have φ/ψ angles closer to the fully extended + 180^0 of the all anti (trans) conformation of dodecanoic acid than the more compact alpha helix. Other types of helices are also found. The 3 10 helix, a sharper helix with three amino acids/turn, falls at -49^0,-26^0. All of these examples of secondary structure fall in allowed regions.

Modern Ramachandran plots do not model atoms as hard spheres but use the Lennard-Jones (6-12) potential to model induced dipole-induced dipole interactions. We discussed this potential function in Chapter 2.4.

A Ramachandran plot of Ala-Ala-Ala is nearly identical to the plot for Phe-Phe-Phe (which is unbranched at the beta carbon, the first methylene C in the side chain). The plot for Thr-Thr-Thr, which has a branch at the beta C (with OH and CH_3 attached), shows somewhat less room than the other plots. Pro-Pro-Pro is most restricted for obvious reasons.

For a longer chain than a tripeptide, there are more restrictions than for $(Ala)_3$, since the chain can't assume a conformation when it passes through itself. The plots for the actual proteins show many points falling in forbidden regions. However, these points would be allowed if the peptide bonds twist a few degrees. Gly bonds also fall outside the allowed areas. This is understandable since the side chain of Gly is H, and it is used in proteins where sharp turns of the chain are necessary.

Figure 4.7.1.11 shows Ramachandran plots for Gly, Ala, Tyr, and Pro in actual proteins (made years OK but can't find reference).

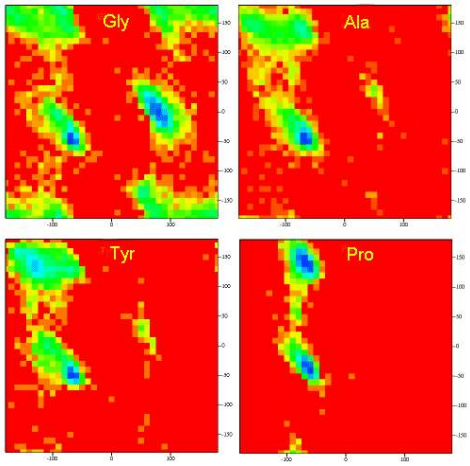


Figure 4.7.1.11: Ramachandran plots for Gly, Ala, Tyr, and Pro from actual proteins

Figure 4.7.1.12, taken from the Proteopedia page above, shows the Ramachandran plots for 100,000 proteins from the Protein Data Bank.

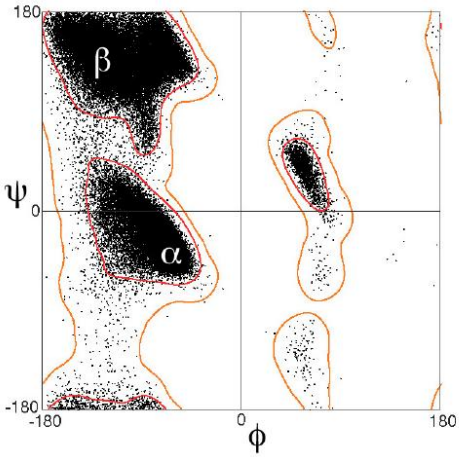


Figure 4.7.1.12: Ramachandran plots for 100,000 proteins. https://proteopedia.org/wiki/index.p...neral_100K.jpg

It's time to explore secondary, tertiary, and quaternary structures in more detail. There are many ways to organize the protein world. For the rest of this chapter, we will adopt the approach of Carl Branden and John Tooze in their seminal book, "Introduction to Protein Structure."

4.7.1.4: Summary

Chapter Summary

This chapter delves into the fundamental principles that govern the conformational properties of protein backbones—a critical aspect of understanding protein folding and structure. It begins by establishing the basic concepts of molecular conformation using simple organic molecules such as butane. By utilizing Newman projections, the chapter illustrates how rotational isomers—such as the high-energy eclipsed conformer and the low-energy anti (trans) conformer—result from variations in dihedral angles.

Building on this foundation, the discussion extends to more complex molecules, such as dodecanoic acid, where the concept of "all-trans" conformations is introduced and kinks created by gauche rotations are highlighted. These ideas are then applied to peptide backbones, exemplified by a tetrapeptide (Gly-Gly-Ala-Gly), demonstrating that the peptide main chain adopts a similar zig-zag, extended conformation in its low-energy state.

The chapter further explains that, due to their partial double-bond character from resonance, peptide bonds are planar and predominantly adopt a trans configuration—except in cases involving proline, where the cis form is relatively more common. This

planarity restricts the degrees of freedom for backbone rotations to two key dihedral angles, phi (φ) and psi (ψ), which define the orientation of adjacent peptide planes.

The chapter introduces Ramachandran plots, which map the allowed and disallowed regions of φ/ψ angles based on steric hindrance and energy considerations, to visualize and analyze these constraints. The plots reveal distinct clusters corresponding to common secondary structures —alpha helices, beta sheets, and random coils —and underscore the unique conformational flexibility of residues like glycine and the restrictions imposed by proline.

Finally, the chapter sets the stage for a deeper exploration of protein secondary, tertiary, and quaternary structures by adopting the framework outlined in classic texts like "Introduction to Protein Structure" by Branden and Tooze. Overall, this chapter highlights the importance of understanding the backbone conformational landscape for deciphering the complex three-dimensional architecture and function of proteins.

This page titled 4.7.1: Main Chain Conformations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.1: Main Chain Conformations** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.2: Secondary Structure and Loops

Learning Goals (ChaptGPT o1, 1/25/25)

- **Understand Helical Structures:**
  - Explain the general concept of protein helices as repetitive structures stabilized by hydrogen bonds between backbone amide Hs and carbonyl Os.
  - Describe the idealized geometries of different helical types—alpha (α), 3_{10}, and pi (π) helices—including parameters such as residues per turn (n), pitch, and the characteristic phi/psi angles.
  - Compare and contrast the alpha helix with the 3_{10} and pi helices in terms of stability, dimensions, and side chain orientation.
  - Interpret helical wheel projections to evaluate amphipathic characteristics and the distribution of hydrophobic and hydrophilic residues along the helix.
- **Examine Beta Structures:**
  - Distinguish between parallel and antiparallel beta strands, and describe how hydrogen bonding patterns and dihedral angles contribute to the formation of beta sheets.
  - Explain how beta strands form "pleated sheet" structures and how the alternating orientation of side chains contributes to sheet stability and overall protein function.
  - Evaluate structural models and interactive 3D representations (e.g., using iCn3D) to visualize beta sheets and beta barrels.
- **Analyze Turns, Loops, and Linkers:**
  - Define and differentiate between various types of reverse turns (beta turns, gamma turns, and hairpins) that connect elements of regular secondary structure.
  - Describe the characteristics of short connector regions (loops, linkers, and bends) and explain how they influence the overall folding and dynamics of a protein.
  - Recognize how specific amino acid sequences, particularly those rich in glycine and proline, facilitate tight turns and loops in protein structures.
- **Explore Backbone Dihedral Angles and Ramachandran Plots:**
  - Define the phi (φ) and psi (ψ) dihedral angles of the peptide backbone and explain how these angles determine the conformational space available to a protein.
  - Interpret Ramachandran plots to identify the allowed and disallowed regions of φ/ψ space, and correlate these with common secondary structures (α-helices, β-sheets, and turns).
  - Discuss how deviations in dihedral angles, due to side chain sterics or flexibility (as in glycine), affect the overall protein structure.
- **Assess Amino Acid Propensities in Secondary Structures:**
  - Describe the factors that determine the propensity of different amino acids to form specific secondary structures, using examples such as the Chou-Fasman parameters.
  - Compare the roles of amino acids with or without beta-carbon branching in stabilizing or destabilizing helical conformations.
  - Explain how the intrinsic properties of amino acids influence the overall stability and folding of the protein, and how these properties can be used to predict structural elements from sequence.
- **Integrate Structural Models with Functional Implications:**
  - Relate the specific secondary structure elements (helices, beta sheets, turns) to the overall 3D architecture and functional properties of proteins.
  - Discuss how secondary structures contribute to the formation of tertiary and quaternary structures and ultimately influence protein activity and interactions.

- Utilize interactive modeling tools (e.g., iCn3D) to explore protein structures such as hydroxynitrile lyase (5Y02) and assess the spatial arrangement of secondary structure elements.

These learning goals are intended to provide students with both a conceptual and a practical understanding of protein secondary structure, enabling them to analyze conformational features from both theoretical and structural perspectives.

Secondary structures are those repetitive structures involving hydrogen bonds between amide Hs and carbonyl Os in the protein backbone. These include

- helices (alpha - α, 3_{10}, and pi - π), in which the hydrogen bonds are within a short continuous stretch of amino acids (a strand),
- beta strands (β sheets) in which the hydrogen bonds are between backbone atoms (again amide Hs and carbonyl Os) on noncontinuous stretches of the protein, and
- reverse turns occur within a very short, continuous stretch of amino acids.

4.7.2.1: Helices

A schematic showing idealized geometries of helices, with amino acids shown as dots for simplicity, is shown in Figure **4.7.2.1**.

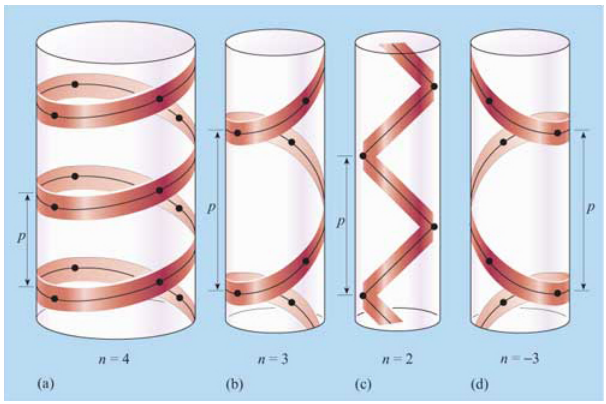


Figure **4.7.2.1**: Helix geometries Open University. https://www.open.edu/openlearn/scien...-section-1.3.1.

The pitch (*p*) represents the spacing between the chain on one side of the helix, the number of amino acids per turn (*n*), and handedness (plus = right-handed, minus = left-handed) are shown in the figure. There are three major types of helices in proteins: the alpha helix (n = 3.6), 3_{10} helix (n = 3), and the pi helix (n = 4.4). Note that n is most commonly not an integer.

4.7.2.1.1: Alpha

The alpha helix is the most common type of helix. They are formed when the carbonyl O of the ith amino acid forms hydrogen bonds to the amide H of the $i+4^{th}$ amino acid (4 amino acids away). Figure **4.7.2.2** shows a short section of an alpha helix running from the N-terminal (bottom) to the C-terminal (top) with the sequence DTASDAA. The amino acids i, i+1, ... i+4 are labeled at their alpha carbons. The red oval highlights the intrastrand H bond between the C=O of the ith amino acid (Asp) and the amide H of the $i+4^{th}$ amino acid (Ser).

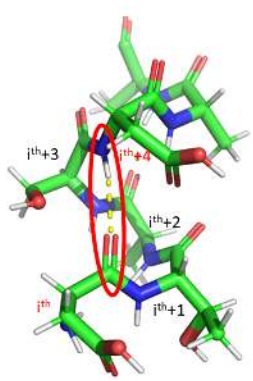


Figure **4.7.2.2**: A short alpha helix

Figure **4.7.2.3** shows a cartoon image showing a longer helix and a schematic showing hydrogen bonding partners.

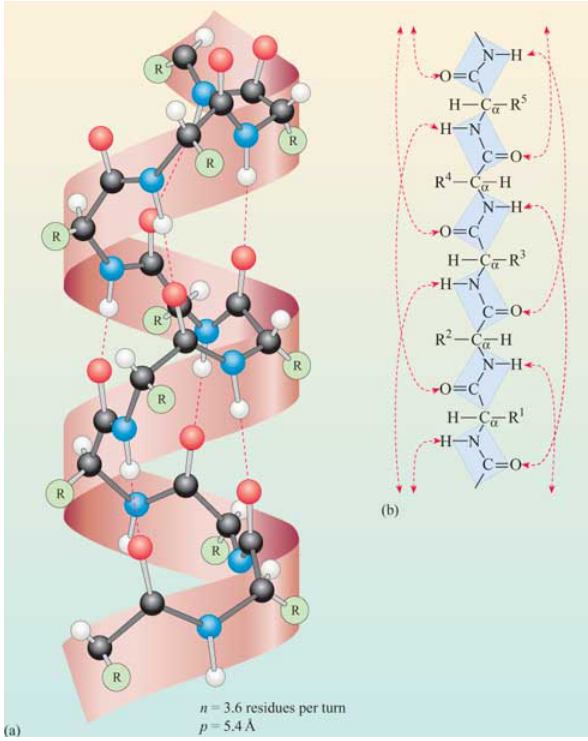


Figure 4.7.2.3: A longer alpha helix Open University. https://www.open.edu/openlearn/scien...-section-1.3.1.

The phi/psi angles for amino acids in the alpha helix are approximately - 57, -47, which emphasize the regular repeating nature of the structure. It can also be characterized by n (the number of amino acid units/turn = 3.6) and the pitch (the helix rise/turn = 5.4 angstroms = 0.54 nm). Since there are 3.6 amino acids per turn, and a full circle or turn is 360^0, each amino acid is staggered at 100^0 increments looking down on the helix axis. To refresh your mind, the phi/psi diagram for a fully extended polypeptide chain (phi 180^0, psi 180^0) is shown below in Figure 4.7.2.4.

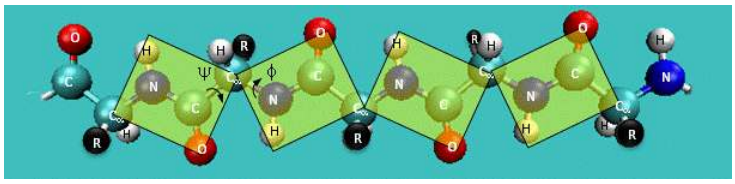


Figure 4.7.2.4: Extended polypeptide with phi/psi planes

Figure 4.7.2.5 shows the side and end-on view of a helix from the antifreeze protein (1wfa) from the winter flounder. The green coil (often shown in red when displaying alpha helices in full proteins) highlights the backbone's repetitive nature. Note that side chains are pointing away from the helix axis. H-bonds are shown as yellow dotted lines within the backbone (one is also shown between two side chains on the top). The spacefill rendering is shown in colors optimized for colorblind people. The end-on view shows that the helix's center is fully packed with atoms from the helix and is **NOT** open (a common misconception among students).

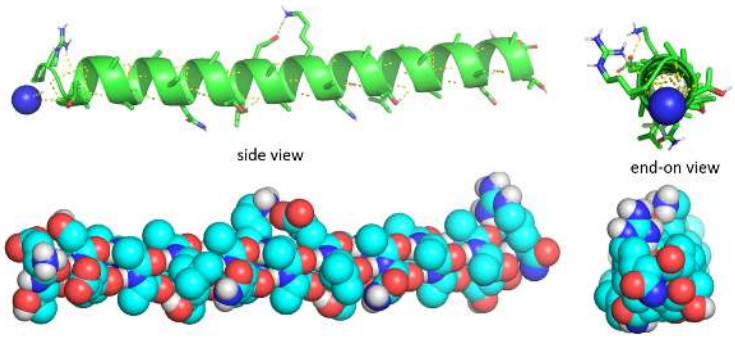


Figure 4.7.2.5: Multiple views and renders of an alpha helix

Some facts:

- The alpha helix is more compact than the fully extended polypeptide chain with phi/psi angles of 180 º.
- In proteins, the average number of amino acids per helix is 11, corresponding to three turns.
- The left-handed alpha helix, although allowed by Ramachandran plot analyses, is never observed since the side chains are too close to the backbone.
- The core of the helix is packed tightly. No central cavities or pores are present in the helix.
- All the R-groups extend backward and away from the helix axis.
- Some amino acids are more commonly found in alpha helices. Amino acids can be divided into two kinds: those with branches at the beta C and those with none. Consider first those that aren't branched. Gly is too conformationally flexible to be found with high frequency in alpha helices, while Pro is too rigid. The amino acids with side chains that can H-bond (Ser, Asp, and Asn) and aren't too long appear to act as competitors of main chain H bond donors and acceptors and destabilize alpha helices. The rest with no branches at the beta C can form helices. Those with branches at the beta carbon (Val, Ile) destabilize the alpha helix due to steric interactions of the bulky side chains with the helix backbone. (Remember, left-handed alpha helices are not naturally found for similar reasons.)
- Alpha keratins, the major component of hair, skin, fur, beaks, and fingernails, are almost all alpha helices.

Figure 4.7.2.6 shows an interactive iCn3D model of an alpha helix from bacteriophage T4 lysozyme (1DYG). Side chains, which are not involved in helix-stabilizing hydrogen bonds, are shown in cyan. H-bonds are shown as green dotted lines.

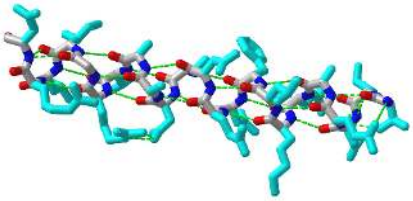


Figure 4.7.2.6: An alpha helix from bacteriophage T4 lysozyme (1DYG). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qhkPT6cMqxvp88 (Copyright; author via source)

The amino acid side chain R groups can be hydrophilic or hydrophobic. They can be localized in specific positions on the helix, forming amphipathic regions on the protein, or fully hydrophobic helices may extend through the plasma membrane as shown in Figure 4.7.2.7.

Figure modified from: Khara, J.S., et al. (2017) *Acta Biomat* 57:103-114 and Ryu, H., et al. (2019) *Int J Mol Sci* 20 (6) 1437

In amphipathic helices, hydrophilic residues are positioned on one side of the helix and hydrophobic residues on the other, as shown in the side view (A) or top-down views (B & C). R-groups may also be entirely hydrophobic within alpha helices that span the plasma membrane, as shown in (D).

Helical wheel projections can show the polarity of the helix's faces, looking down the helix's axis. Here are two such helical wheel projections:

For the sequence MLQSMVSLLQSLVSLIIQ, Figure 4.7.2.8 shows that the helix is amphiphilic.

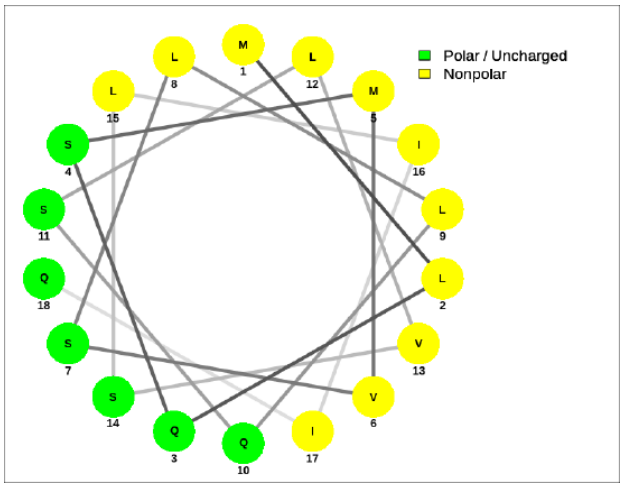


Figure 4.7.2.8: Helical wheel for MLQSMVSLLQSLVSLIIQ http://lbqp.unb.br/NetWheels/

A helical wheel for the membrane-crossing section of the human receptor-type tyrosine-protein phosphatase C protein, ALIAFLAFLIIVTSIALLVVL, is shown in Figure 4.7.2.9.

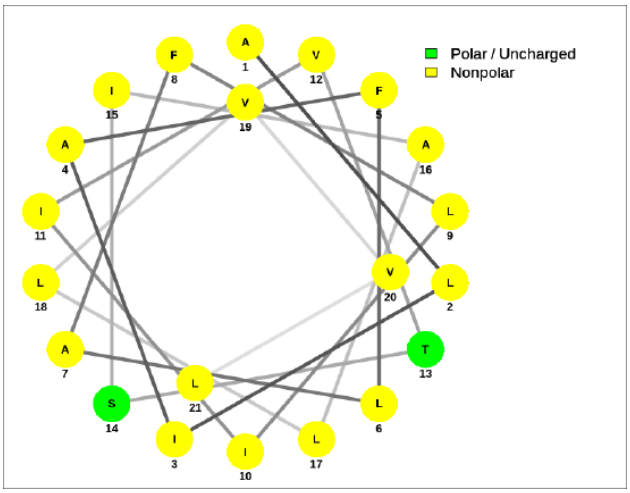


Figure 4.7.2.9: Helical wheel for a membrane helix

The amide bond in the peptide has a significant permanent dipole moment. Since the dipoles of individual amino acids are oriented in the same direction in an alpha helix, the whole helix has a significant dipole moment. This is illustrated in Figure 4.7.2.10.

recreate this image:

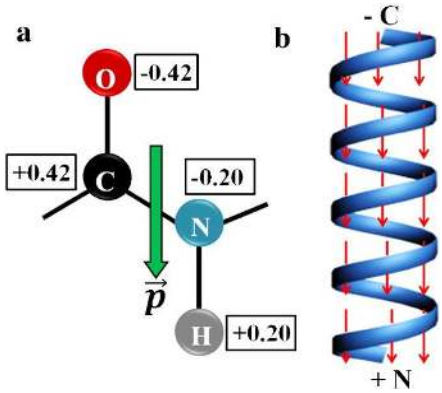


Figure 4.7.2.10: Helix Dipole (https://www.sciencedirect.com/scienc...05273612003860; reproduce image)

In the alpha helix, all the amides in the backbone point in one direction from the C-terminus to the N-terminus. This leads effectively to one long dipole of magnitude n(3.5) Debyes, where n is the number of amino acids in the helix. A long helix can

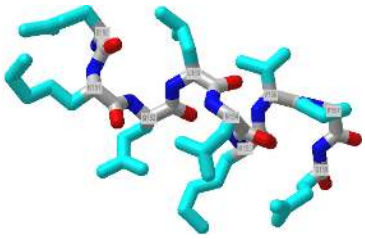
produce a significant electric field and affect protein binding properties.

4.7.2.1.1.1: 3_{10} helices

The 3_{10} (or 3 10)helix is stabilized by hydrogen bonds between the carbonyl O of the ith amino acid and the amide H of the $i+3^{th}$ aa (3 amino acids away). It has three residues/turn and a pitch (rise per turn) of 6 angstroms, with a rise of 1.3-2 angstroms/residue. Typical phi/psi angles are -50^0,-26°. As with the alternative description of the alpha helix, the $\mathbf{3_{10}}$ helix has **three** amino acids per turn and **10** atoms in the main chain/turn (counting Cα-N-C-Cα atoms). It is longer (for the same number of amino acids) and thinner. The amino acid side chains are staggered at 120^0 increments as you look around the helix axis. Although not very prevalent (about 3 percent of protein amino acids are in 3_{10} helix with an average of 3.3 amino acids in the helix), they presumably serve some function. 3_{10} helices as long as 11 residues have been found.

A helix will be stable within a protein if the packing and noncovalent interactions favor it over other conformations. Many more alpha-helix side chain interactions with surrounding protein are likely, given a 100^0 staggering of side chains compared to a 120^0 staggering in a 3_{10} helix, which line up in three ridges looking down the helical axis. Molecular dynamics studies suggest that parts of a 3_{10} helix might reversibly interconvert to an alpha helix, allowing conformational and binding flexibility. The S4 helix in some voltage-sensitive potassium ion channels with a canonical $R_1xxR_2xxR_3xxR_4xxK_5xxR_6$ (where R and K are Arg and Lys) has been shown to adopt a 3_{10} helical conformation.

Figure **4.7.2.11** shows an interactive iCn3D model of a nine amino acids (150-158) 3_{10} helix from dienelactone hydrolase (1DIN)



NCBI iCn3D Figure **4.7.2.11**: A 3_{10} helix from from dienelactone hydrolase (1DIN). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Kufi7gmpSBdjYA (Copyright; author via source)

4.7.2.1.1.2: π (pi) helices

This helix has 4.4 residues/turn and a helix pitch (rise) of about 4.1 nm. It has hydrogen bonds between the carbonyl O of the ith amino acid and the amide H of the $i+5^{th}$ amino acid (5 amino acids away). Its rise is about 1.2 angstroms/residue and has approximate phi/psi angles of -55^0, -70^0. An alternative designation for the pi helix is 4.4_{16}, with 16 main chain atoms in 1 full turn (see above for the alpha helix). Some reasons for its low abundance include minimal contact between the main chain atoms, given the larger radius and phi-psi angles close to disallowed values. It might also not form as quickly as the other helices, since its nucleation would be more difficult. At the same time, molecular dynamics simulations show alpha helices can interconvert reversibly with pi helices. They are often found between two alpha helices, again suggesting that dynamic interconversions between the two forms are likely.

About 55% of characterized pi helices contain five amino acids. Each side chain is staggered by 85^0, with a rise of about 1.3 Angstroms. Figure 12 below shows an interactive iCn3D model of a short pi helix (aa 265-276) from barley beta-D-glucan glucohydrolase (1x38).

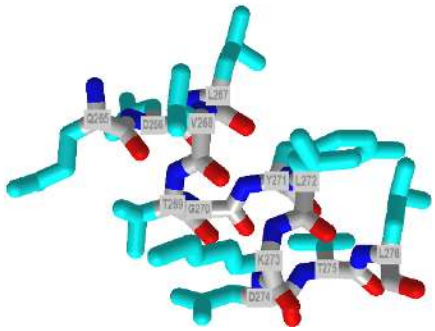


Figure 4.7.2.12: A short pi helix (aa 265-276) from barley beta-D-glucan glucohydrolase (1x38). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...U7H6xjzHd95nbA (Copyright; author via source)

PROTEOPEDIA — LIFE IN 3D — Helices in Proteins: Comparison of alpha, 3_{10}, and pi helices

4.7.2.2: Beta Structure

Beta Structure: **Parallel** and **antiparallel** beta strands are much more extended than alpha helices (phi/psi of -57,-47) but not as extended as a fully extended polypeptide chain (with phi/psi angles of +/- 180), as shown in the figure below. Parallel beta strands have phi/psi angles of -119, +113, while the antiparallel angles are -139, +135. Figure 4.7.2.13

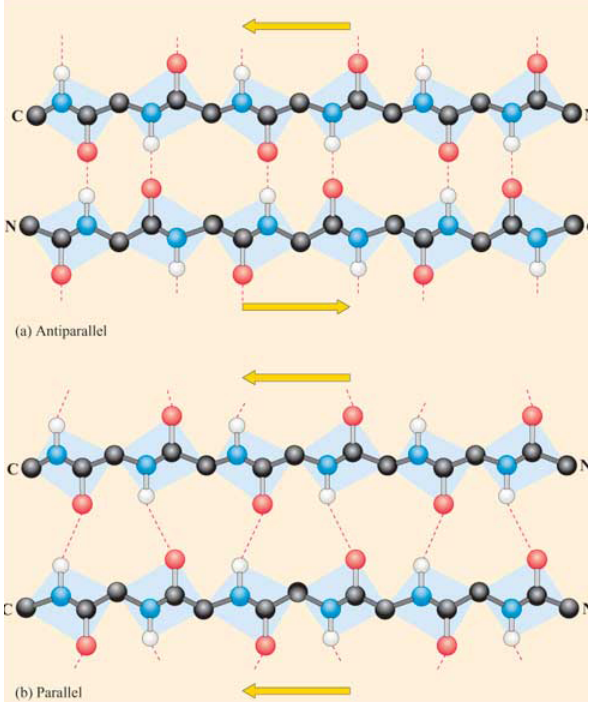


Figure 4.7.2.13Open University. https://www.open.edu/openlearn/scien...-section-1.3.2

Each single strand of the beta-sheet can be pictured as a twofold helix, i.e., a helix with two residues/turn. The arrangement of each successive peptide plane is pleated due to the tetrahedral nature of the alpha C. Hydrogen bonds are inter-strand, not intra-strand, as in the alpha helix.

The figure below shows how the "pleats" in a sheet containing parallel beta strands can be envisioned as rippled sheets. Figure 4.7.2.14

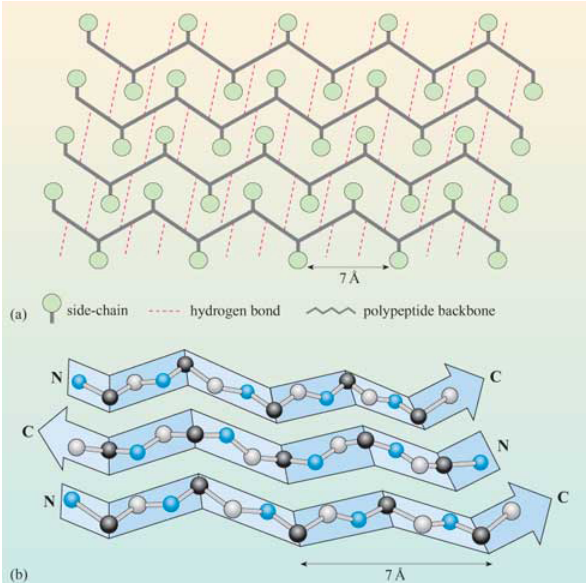


Figure 4.7.2.14

They can be visualized by folding a sheet of paper into narrow folds or pleats to form a "pleated sheet" of paper. Each strip of paper can be pictured as a single peptide strand, with the peptide backbone zigzagging along the strip, with the alpha carbons lying at the folds of the pleats. The R groups are attached to the carbons and extend above and below the pleat folds in the *trans* conformation.

Consider a strand as a continuous, contiguous polypeptide backbone propagating in one direction. Hence, using this definition, a helix consists of a single strand, and all the H-bonds are within the strand (or intrastrand). A beta sheet would then consist of multiple strands since each "strand" is separated from other "strands" by an intervening contiguous stretch of amino acid, which bends within the protein in a way that allows the next section of the peptide backbone, the next "strand," to H-bond with the first "strand." But remember, even in this case, all the H-bonds holding the alpha and beta structures together are intramolecular.

In a parallel beta sheet structure, the optimal H bond pattern leads to a less extended structure (phi/psi of -119, +113) than the optimal arrangement of the H bonds in the antiparallel structure (phi/psi of -139, +135). Also, the H bonds in the parallel sheet are bent significantly. (i.e., the carbonyl O on one strand is not exactly opposite the amide H on the adjacent strand, as it is in the antiparallel sheet.) Hence, antiparallel beta strands are presumably more stable, although both are abundant in nature. Short parallel beta sheets with 4 strands or fewer are uncommon, which may reflect their lower stability.

The side chains in the beta sheet are perpendicular to the sheet's plane, extending out on alternating sides. Parallel sheets characteristically distribute hydrophobic side chains on both sides of the sheet, while antiparallel sheets are usually arranged with all the hydrophobic residues on one side. This requires alternating hydrophilic and hydrophobic amino acid side chains in the primary sequence. Antiparallel sheets are found in silk, running parallel to the silk fibers. The following repeat is found in the primary sequence: $(\text{Ser-Gly-Ala-Gly})_n$, with Gly pointing out from one face and Ser or Ala from the other.

Unfortunately, no PDB structure of the silk "amyloid" protein shows this repetitive structure. The monomers and aggregates of this protein are quite insoluble, so few X-ray structures for proteins like this are available. Figure 4.7.2.15 shows an interactive iCn3D model of the N-terminal part (domain) of the *Bombyx mori* fibroin silk protein (pdb = **3UA0),** which gives an excellent example of antiparallel beta sheets. Notice that two chains align to form a face of the curved antiparallel beta sheet.

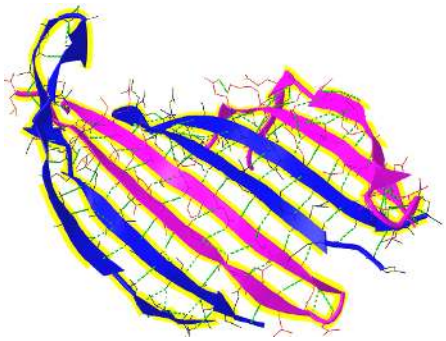


Figure 4.7.2.15: N-terminal Domain of the *Bombyx mori* fibroin silk protein (3UA0). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...TuDiMoi64TaAx9 (Copyright; author via source)

Figure 4.7.2.16 shows a more detailed static image of the antiparallel beta sheets in 3UA0. Note the yellow sticks between the strands representing the H-bonds.

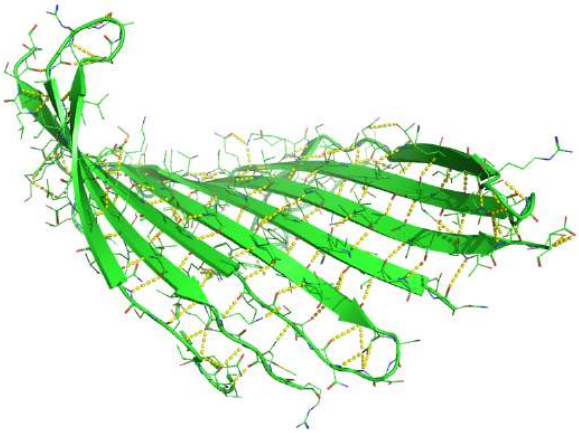


Figure 4.7.2.16:

Beta strands tend to twist in the right-hand direction. This has important consequences for how the beta strands are connected. Parallel strands can form twisted sheets, saddles, and beta barrels.

Figure 4.7.2.17 shows an interactive iCn3D model of the parallel beta sheet structure from the arabinose binding protein (1ABE).

Figure 4.7.2.17: Parallel beta sheet structure from the arabinose binding protein (1ABE). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7A2akqvuCpavu8 (Copyright; author via source)

Figure 4.7.2.18 shows a static image of the parallel beta sheets in 1ABE. Note the yellow sticks between the strands representing the H-bonds.

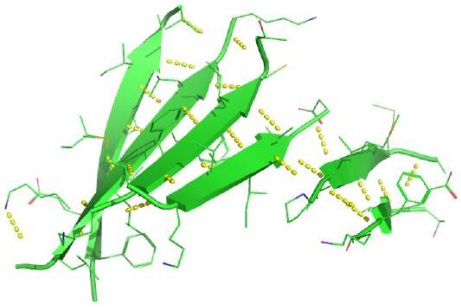


Figure 4.7.2.18: Parallel beta sheet structure from the arabinose binding protein (1ABE)

Figure 4.7.2.19 shows an interactive iCn3D model of a parallel beta barrel from the triose phosphate isomerase (1WYI).

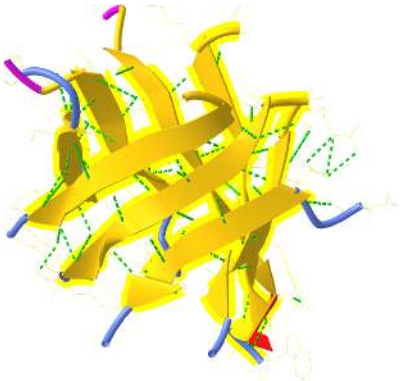


Figure 4.7.2.19: Parallel beta barrel from the triose phosphate isomerase (1WYI) (Copyright; author via source)). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7dki3WjRDtiaRA

Figure 4.7.2.20 shows a static image of the parallel beta sheets in triose phosphate isomerase. Note the yellow sticks between the strands representing the H-bonds. Also, note that the barrel is not hollow but is filled with side chains.

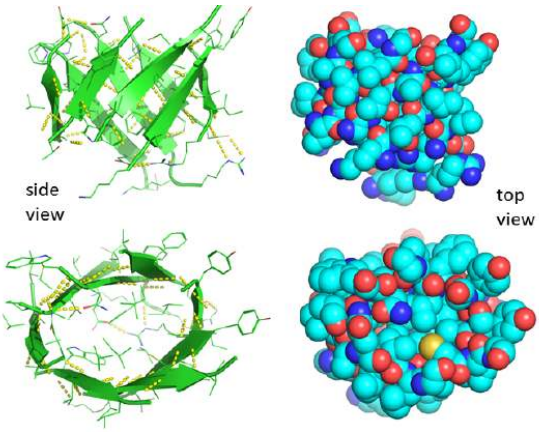


Figure 4.7.2.20: Parallel beta sheets in triose phosphate isomerase

Some facts on parallel beta structure:

- In parallel strands, right-handed connectivity is common.
- In a protein with parallel strands in a register, and given the inherent twist in the strands, the strands arrange to have the H bonds stretched equally at the ends of the chains, giving rise to a twisted saddle shape (top structure above).
- In a protein with parallel strands out of register, and given the inherent twist in the strands, the strands arrange to have the H bonds stretched equally at the ends of the chains, giving rise to a beta barrel (bottom structure above).

4.7.2.3: Connectors, Loops, Linkers, and Bends

About 50% of the amino acids in a globular (spherical) protein are in regular secondary structure (alpha or beta). The amino acids in helices and beta strands are connected by stretches of amino acids that still retain order. Still, that order is less regular than those

found in helices and beta strands, characterized by stretches of amino acids with the same phi/psi angles. Some bear the hallmarks of secondary structure - intrachain hydrogen bonds hold them together. We will consider a few here.

Turns, Reverse Turns, and Hairpins

One example of a connector involving secondary structure (i.e. hydrogen bonds between amide Hs and carbonyl Os of the backbone), is a reverse turn called the **beta bend or beta turn**. These turns often connect successive antiparallel beta strands and are then called beta hairpins. A **hairpin** is a special case of a turn in which the direction of the protein backbone reverses, and the flanking secondary structure elements interact. For example, the **beta hairpin** connects two antiparallel β-strands via hydrogen bonds. The word *beta* can be confusing. It does not mean the structure has hydrogen-bonded amino acids with the same phi-psi angles as beta strands. It's easy to remember the name beta as the beta bend connects two antiparallel beta strands. The term beta comes from the fact that it is a member of a class of turns named with Greek letters, including alpha-, gamma-, delta-, pi-, and beta-turns.

They are almost always on the surface and usually contain four amino acids. However, several types of beta turns and different ways to classify them exist. One involves the number of residues (n) *between the two hydrogen-bonded residues*.

n=2: These contain four amino acids. Amino acids 1 and 4 form hydrogen bonds with n=2 amino acids in between. Another way to describe them is that the hydrogen bond between residues 1 and 4 is between the backbone carbonyl O of the ith amino acid and the amide H of the $i+3^{th}$ amino acid (three amino acids away), so the structure contains four amino acids (i^{th}, $i+1^{th}$, $i+2^{th}$, and $i+3^{th}$). There are two common types:

- Type I: phi 2 = -60, psi 2 =-30; phi 3 = -90, psi 3 = 0; The first amino acid in the actual turn (i^{th} + 1) is in a left-handed alpha helix conformation. Glycine, asparagine, or aspartate are stable at this position since glycine is small and the side chains of Asp and Asn can form hydrogen bonds to the main chain. Glycine is usually found in the second position of the actual turn ($i+2^{th}$).
- Type II: phi 2 = -60, psi 2=120; phi 3 = 90, psi 3 = 0; The first residue of the actual turn is typically Gly, while the second often is a polar amino acid such as Ser and Thr

Here are 2D line drawings showing Type I and Type II beta turns. Figure 4.7.2.21

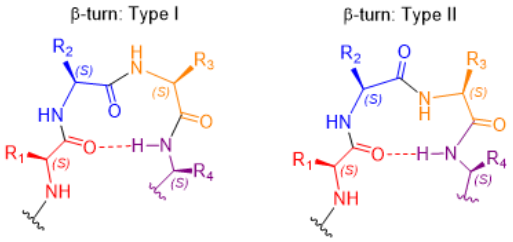


Figure 4.7.2.21

The figure below shows Type I (left) and Type 2 (right) from human egg white lysozyme. Figure 4.7.2.22

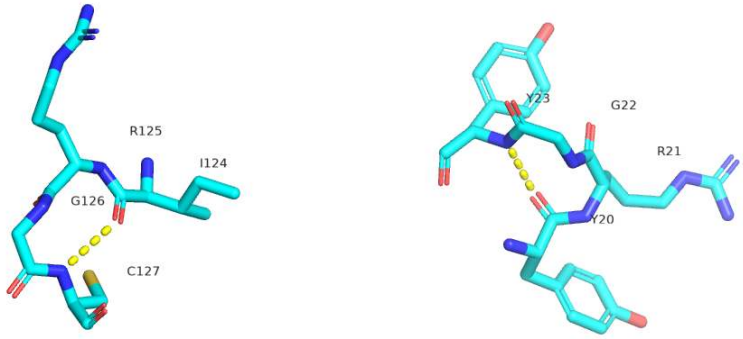


Figure 4.7.2.22

Figure 4.7.2.23 shows an interactive iCn3D model showing two reverse turns in hen egg white lysozyme (1dpx). Notice the tightness of the reverse turn and the presence of proline and glycine.

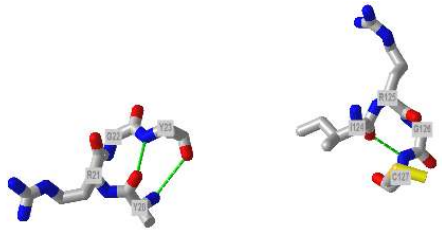


Figure 4.7.2.23: Two reverse turns in hen egg white lysozyme (1dpx) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...U2BzPt98cPYHFA

n=3: These contain five amino acids. Amino acids 1 and 5 form hydrogen bonds with n=3 amino acids between them. Another way to describe them is that the hydrogen bond between residues 1 and 4 is between the backbone carbonyl O of the ith amino acid and the amide H of the $i+4^{th}$ amino acid (four amino acids away).

Shifting back to the Greek letter naming system, the **gamma turn**, the second most common turn, has just three total residues (i^{th}, $i+1^{th}$, and $i+2^{th}$) with the hydrogen bond between the backbone carbonyl of the ith amino acid and the backbone amide H of the $i+2^{th}$ amino acid.

An ***ω-loop*** is a catch-all term for a longer, extended, or irregular loop **without** fixed internal hydrogen bonding. Turns are sometimes found within flexible linkers or loops connecting protein domains. Linker sequences vary in length and are typically rich in polar uncharged amino acids. ***Flexible linkers*** between domains allow twisting and rotation to recruit their binding partners via protein domain dynamics.

Most modeling programs display the linear sequence of a protein, and a linear cartoon rendering shows the alpha structure as squiggles or helices and the beta structure as yellow arrows. The figure below shows a 1D connectivity diagram for part of the protein alpha-lactalbumin (1a4v). Figure 4.7.2.24

Figure 4.7.2.24:

4.7.2.4: Amino Acid Propensities for Secondary Structures

Finally, what types of amino acids are most likely found in different secondary structures? Some rationales for the "propensity" for secondary structure are shown in the figure below. Figure 4.7.2.25

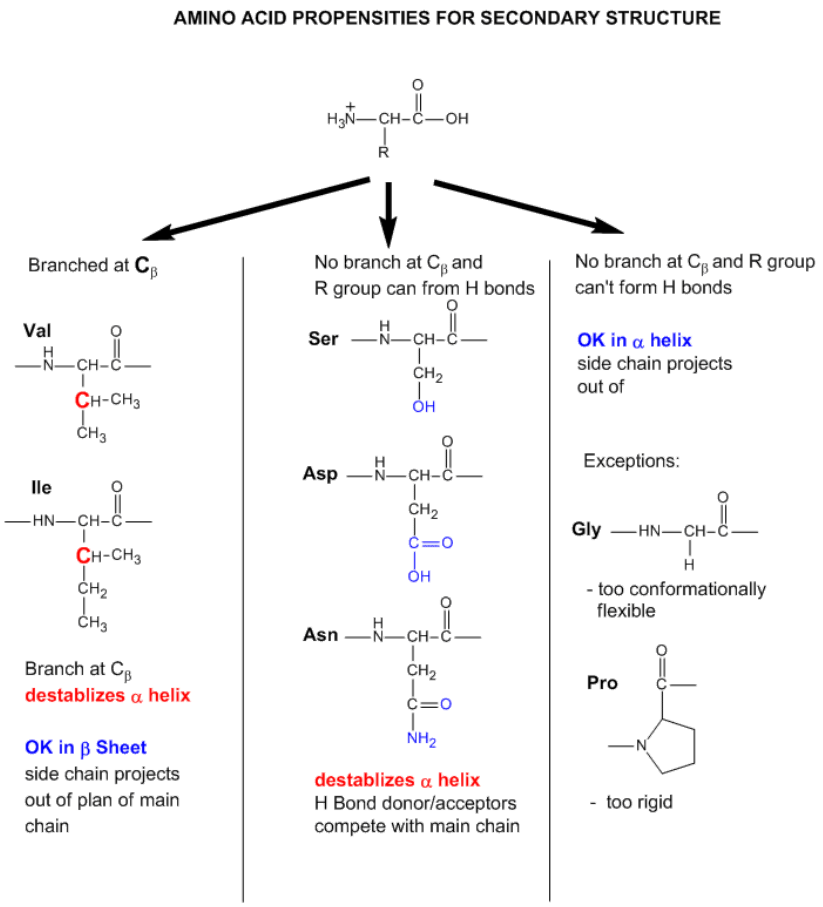


Figure 4.7.2.25:Amino Acid Propensities for Secondary Structure

Chou-Fasman derived a list of propensities for each amino acid for secondary structure, based on their occurrence in determined protein structures.. These values are shown in Figure 4.7.2.26 below.

Helical Residues	Pα	β-Sheet residues	Pβ
Glu$^{(-)}$	1.53	Met	1.67
Ala	1.45	Val	1.65
Leu	1.34	Ile	1.60
His$^{(+)}$	1.24	Cys	1.30
Met	1.20	Tyr	1.29
Gln	1.17	Phe	1.28
Trp	1.14	Gln	1.23
Val	1.14	Leu	1.22
Phe	1.12	Thr	1.20
Lys$^{(+)}$	1.07	Trp	1.19
Ile	1.00	Ala	0.97
Asp$^{(-)}$	0.98	Arg$^{(+)}$	0.90
Thr	0.82	Gly	0.81
Ser	0.79	Asp$^{(-)}$	0.80
Arg	0.79	Lys$^{(+)}$	0.74
Cys	0.77	Ser	0.72
Asn	0.73	His$^{(+)}$	0.71
Tyr	0.61	Asn	0.65
Pro	0.59	Pro	0.62
Gly	0.53	Glu$^{(-)}$	0.26

Figure **4.7.2.26**: Chou-Fasman propensity values. Higher + values represent higher propensities. Jain, Vikas & Tu, Raymond. (2011). Coupled Folding and Specific Binding: Fishing for Amphiphilicity. International journal of molecular sciences. 12. 865-89. 10.3390/ijms12031431. Creative Commons Attribution 3.0 Unported

4.7.2.5: Summary

This chapter explores the critical elements of protein secondary structure, emphasizing how the repetitive arrangement of backbone hydrogen bonds defines key structural motifs and influences overall protein conformation. The discussion begins with an in-depth look at helical structures, including the most common alpha helix and the less prevalent 3_{10} and pi helices. These helices are characterized by their specific geometric parameters—such as the number of residues per turn, pitch, and characteristic phi/psi angles—and by the way their side chains project outward, often creating amphipathic surfaces or contributing to membrane-spanning segments.

Next, the chapter examines beta structures, distinguishing between parallel and antiparallel beta sheets. It explains how beta strands, which are more extended than helices, form pleated sheets via inter-strand hydrogen bonds. The orientation of side chains—alternating above and below the sheet plane—and the differences in hydrogen bond geometry between parallel and antiparallel arrangements are discussed, highlighting factors that affect stability and the overall twist of the sheet.

The chapter also covers the connectors that join these regular secondary structures, such as reverse turns, hairpins, loops, and flexible linkers. These regions, often rich in specific amino acids such as glycine and proline, allow the polypeptide chain to reverse direction and form compact, globular tertiary structures.

To rationalize the conformational preferences of protein backbones, the text introduces the concept of dihedral angles—phi (φ) and psi (ψ)—and describes how they are restricted by the planar nature of the peptide bond, as revealed by resonance stabilization. Ramachandran plots are powerful tools for visualizing the allowed and disallowed regions of the backbone dihedral space and correlating them with common secondary structure elements.

Finally, the chapter reviews amino acid propensities for forming various secondary structures, using approaches like the Chou-Fasman analysis to explain why certain residues favor alpha helices, beta sheets, or turns. This integrated overview provides a foundation for understanding how the recurring patterns of hydrogen bonding and the intrinsic properties of amino acids combine to dictate the folding and functional architecture of proteins.

Overall, the chapter lays the groundwork for a deeper exploration of protein tertiary and quaternary structure by elucidating the fundamental principles governing the formation and stability of protein secondary structure elements.

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

This page titled 4.7.2: Secondary Structure and Loops is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.2: Secondary Structure and Loops** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.3: Tertiary, Quaternary, and Symmetrical Structures

Learning Goals (ChatGPT o1, 1/26/25)

- **Comprehend Protein Tertiary Structure Fundamentals:**
  - Define tertiary structure as the overall three-dimensional folding of a single polypeptide chain, integrating its embedded secondary structures, super-secondary motifs, and domains.
  - Explain how the distribution of hydrophobic, polar, and charged side chains influences the folding process, leading to the burial of nonpolar groups and exposure of polar/charged groups on the protein surface.
- **Analyze Noncovalent Interactions in Protein Folding:**
  - Identify the key noncovalent forces (e.g., hydrogen bonds, salt bridges, hydrophobic interactions, pi–cation, and aromatic interactions) that stabilize the tertiary structure.
  - Evaluate examples of how these interactions, such as those observed in human low molecular weight protein tyrosyl phosphatase or the mouse immunoglobulin G light chain, contribute to the protein's stability and function.
- **Interpret Conformational Models and Micelle Analogies:**
  - Compare simplified models (e.g., micelle-like aggregation) with the more complex intramolecular folding patterns of proteins.
  - Discuss why a strictly micellar model is inadequate for proteins, considering that the entire backbone is polar and contributes to folding dynamics.
- **Explore Tertiary Structure and pKa Modulation:**
  - Explain how the microenvironment within a folded protein can alter the pKa values of ionizable side chains via dehydration (Born effect), Coulombic interactions, and hydrogen bonding.
  - Assess how these modulated pKa values influence the protein's overall charge distribution and stability.
- **Understand Quaternary Structure and Protein Symmetry:**
  - Define quaternary structure as the assembly of two or more protein subunits (homo- or hetero-oligomers) and describe how noncovalent interactions and disulfide bonds contribute to complex formation.
  - Differentiate between various symmetry types (cyclic C_n, dihedral D_n, and cubic symmetries) and describe how these principles apply to protein oligomers such as hemoglobin, ferritin, and chaperonin complexes.
- **Apply Symmetry Concepts to Protein Oligomerization:**
  - Describe how symmetry operations (e.g., proper rotation axes) govern the arrangement of identical subunits in homomeric complexes, while noting that chirality restricts the use of inversion and mirror symmetry in proteins.
  - Explain how evolutionary selection and minimal mutational changes can promote symmetric oligomer formation and influence protein function.
- **Utilize Structural Models and Interactive Tools:**
  - Interpret interactive 3D models (e.g., via iCn3D) to visualize tertiary structure elements, analyze noncovalent interactions, and understand the spatial arrangement of subunits in quaternary assemblies.
  - Correlate structural features observed in models with experimental findings from techniques such as X-ray crystallography and cryo-electron microscopy.
- **Evaluate the Impact of Mutations on Oligomerization:**
  - Discuss how single amino acid substitutions can alter the interfaces between subunits, potentially leading to changes in oligomeric state (e.g., dimer vs. higher order assemblies) and symmetry.
  - Assess the implications of engineered mutations (as demonstrated by Grueninger et al.) for designing proteins with altered oligomerization properties.

These goals are designed to integrate theoretical concepts with practical observations, enabling students to critically analyze how the principles of noncovalent interactions and molecular symmetry govern the complex architecture and function of proteins.

4.7.3.1: Tertiary Structure

The tertiary structure of a single-chain protein is the overall 3D structure of the protein. A protein of a given primary structure folds to form a 3D structure with embedded secondary structures, supersecondary structures, and domains. Folded proteins can have a variety of shapes, from a roughly spherical or "globular" to a more extended "fibrillar" form. Let's consider the more globular one first. How a protein folds will be discussed in greater detail later, but a more descriptive and straightforward view will help us understand the structural features of the folded protein in its tertiary structure.

Start with an unfolded protein. It has a polar backbone with dangling polar charged, polar uncharged, and nonpolar sides attached to the alpha carbon of each amino acid in the primary sequence. On folding, where do these side chains of varying polarity end up? To a first approximation, you may think that a globular protein might fold such that all the hydrophobic amino acid side chains are buried in the protein's interior, surrounded by other hydrophobic side chains. Similarly, you might expect the polar and charged side chains to be on the surface, exposed to water.

Such a model would be analogous to a **micelle,** which has an almost "perfect" separation of polar (on the surface) and nonpolar atoms (buried). Figure **4.7.3.1** shows an interactive iCn3D model of a micelle below, which consists of 54 self-associated molecules of dodecylphosphocholine fatty acids.

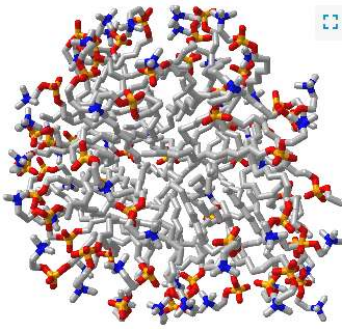


Figure **4.7.3.1**: A Micelle (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GkJxUQJKdHNfJ7

If protein folding were only that simple! Topologically, a protein can't fold in an intramolecular fashion in a strict analogy to the intermolecular aggregation of single-chain amphiphiles into a micelle. Also, consider that the entire backbone is polar! To a first approximation, we would expect the bulk of nonpolar groups to be buried and surrounded by other nonpolar groups. Likewise, we expect the bulk of polar and charged groups to be on the surface. Figure **4.7.3.2** shows an interactive iCn3D model of part of human low molecular weight protein tyrosyl phosphatase (1xww). It shows one buried nonpolar side chain (Phe 10) surrounded by essentially all nonpolar side chains of other amino acids.

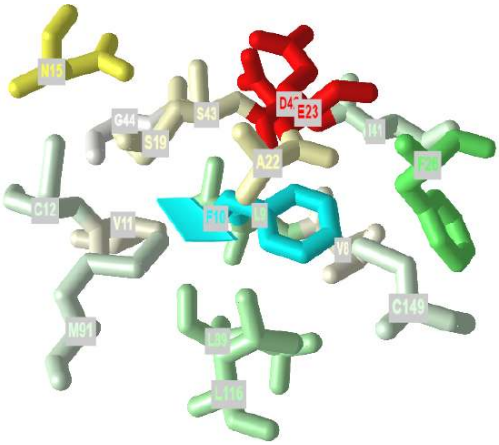


Figure 4.7.3.2: A buried phenylalanine in low molecular weight protein tyrosyl phosphatase (1xww) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LC53qmf4FyszT7

The completely nonpolar Phe 10 side chain is shown in cyan. The atoms with 5 Angstroms are color-coded using the Wimley–White whole residue hydrophobicity scale, which uses not only side chain but peptide bond contributions that are determined experimentally by determining the free energy of transfers of groups to nonpolar environments. The color scale runs from cyan (nonpolar) to red (polar/charged. Like a micelle, the protein is roughly spherical.

Noncovalent interactions between atoms within a protein chain help drive protein folding. The noncovalent interactions (also termed intermolecular forces in traditional introductory chemistry classes) include ion-ion, ion-dipole, hydrogen bonds, dipole-dipole, induced dipole-induced dipole (often called London Dispersion Forces), and variants including ion-induced dipole, etc. We generally use the same terms when describing these interactions within and between proteins. Ion-ion interactions are usually called salt bridges, and induced dipole-induced dipole interactions are often called hydrophobic interactions.

Let's look at specific interactions with a given protein chain (the light chain of the mouse immunoglobulin G, PDB 1D = 4hdi). Manipulate the iCn3D model in the exercises below and answer the following questions. Use your mouse to hover over amino acids to help in identification.

Exercise 4.7.3.1

Name the types of interactions between the following side chains in the iCn3D image above. A larger image can be seen by using this link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?DrdnVubujRaihTcr9

a. Leu 52 and Ile 53, Val 63, Phe 67 and Leu 78

b. Arg 24 and Asp 75

c. Asp 170 and Lys 108

d. Tyr 178 and Lys 147

e. Glu 190 and Arg 160

f. Tyr 191 and Phe 214

Answer

a. hydrophobic interactions, induced dipole-induced dipole

b. salt bridge, ion-ion

c. salt bridge, ion-ion

d. pi-cation

e. salt bridge, ion-ion

f. Aromatic-aromatic, induced dipole-induced dipole

You can analyze the noncovalent interactions within and between a protein using PIC- protein interactions calculator

In the next exercise, identify the likely hydrogen donors and acceptors in the pairs shown

Exercise 4.7.3.1

Are these hydrogen bonds

- side chain to side chain?
- main chain to main chain?
- side chain to main chain?

a. Ile 111: Gln 171

b. Gln 6: Thr 107

c. Ile 53: Trp 40

d. Try 37:Thr 97

Answer

a. Ile 111: Gln 171 - side chain to side chain

b. Gln 6: Thr 107 - side chain to side chain

c Ile 53: Trp 40 - main chain to main chain

d. Try 37:Thr 97 - main chain to main chain

4.7.3.2: A more realistic understanding of noncovalent interactions

But are all the nonpolar side chains buried? How about the polar uncharged and polar charged side chains? What are the preferred dispositions of side chains in proteins as derived from the crystal structure of thousands of proteins? Here are some conclusions from a paper by Pace (Biochemistry. 40, pg 310 (2001).

- About 50% of the amino acids are in secondary structures. On average, there is about 27% alpha helix and 23% beta structure. Of course, some proteins are almost entirely alpha-helical, and some are almost entirely beta-sheet, but most are a mixture.
- The location of side chains varies with polarity. Nonpolar side chains, such as Val, Leu, Ile, Met, and Phe, are predominantly (83%) in the protein's interior.
- Charged polar side chains are almost equally distributed between buried and surface-exposed states. (54% - Asp, Glu, His, Arg, Lys are buried away from water, a bit startling!)
- Uncharged polar groups such as Asn, Gln, Ser, Thr, and Tyr are mostly (63%) buried and not on the surface (a bit startling).
- Globular proteins are quite compact, with water excluded from their interiors. Their packing density ($Vol_{van\ der\ Waals}/Vol_{tot}$) is about 0.75, which is similar to that of the NaCl crystal and equals the closest packing density of 0.74. Compare this to the density of organic liquids, which is about 0.6-0.7.

4.7.3.3: Tertiary structure and pKa Values

If a charged side chain is buried in a protein, you would expect that it would be surrounded, in general, by either oppositely charged side chains, to which it could form an internal salt bridge (ion-ion interaction), or a polar uncharged group with which it could interact through dipole-dipole or, more specifically, H-bond interactions. You would also expect that it would be buried in an uncharged state if it were not near an oppositely charged side chain.

Hence, the pKa of side chains would be dramatically affected by the nature of their microenvironment (as we have already seen with the pKa of acetic acid in solvents of different polarity). NMR spectroscopy has been used to determine the pKa values of specific side chains in proteins whose crystal structure is known. Pace et al. (2009) summarize data on the properties of ionizable side chains across a series of proteins whose structures have been determined. The intrinsic pKa, pKa_{int}, or prototypical pKa value for a side chain exposed to water can be determined using a pentapeptide containing the target amino acid X, flanked by 2 Ala on either side, with both the N- and C-termini of the peptide blocked so they are uncharged. Table **4.7.3.1** below shows the pKa values of ionizable side chains in a series of proteins compared to the control pentapeptide.

Group	Content %	Buried %	pKa int in AAXAA	pKa avg	low pKa	high pKa	# measurements
Asp	5.2	56	3.9	3.5 + 1.2	0.5	9.2	139
Glu	6.5	48	4.3	4.2 + 0.9	2.1	8.8	153
His	2.2	72	6.5	6.6 + 1.0	2.4	9.2	131
Cys	1.2	90	8.6	6.8 + 2.7	2.5	11.1	25
Tyr	3.2	67	9.8	10.3 + 1.2	6.1	12.1	20
Lys	5.9	34	10.4	10.5 + 1.1	5.7	12.1	35
Arg	5.1	56	12.3				
C term			3.7	3.3 + 0.8	2.4	5.9	22
N term			8.0	7.7 + 0.5	6.8	9.1	16

Table 4.7.3.1: pKa values of side chains in actual proteins

A glance at the table shows a large variation in the pKas of ionizable side chains in proteins, with the pKa of Asp spanning 8.7 pH units, indicating that it can act at physiological pH as either a strong acid or a moderate base. Three major effects can perturb the pKa of ionizable side chains:

1. Dehydration of the side chain as it is buried in a protein (Born Effect): The stability of a charged group depends on the polarity of the medium in which it exists. Ions are more stable in water than in nonpolar solvents, as the water molecules can reorient and interact with the ion through ion-dipole or ion-H bond interactions, effectively shielding the ion from other counterions. The shielding effect of water is related to the dielectric constant, ε, of the solvent. Coulomb's law can be written as:

$$F = \frac{kQ_1Q_2}{r^2} = \frac{Q_1Q_2}{4\pi\varepsilon r^2}$$

Epsilon is the dielectric constant of the solvent. Water has a higher dielectric constant (80) than nonpolar solvents (4-10), and it shields opposing charges more effectively, stabilizing them. Hence, the pKa values of side chains of amino acids whose deprotonated states are charged will be raised (so they are less acidic) in nonpolar environments. The reverse holds for side chains whose protonated form is charged. Pace cites, as an example, two mutants of staphylococcal nuclease in which a buried Val 66 is changed to either Asp or Lys. The buried Asp has a pKa of 8.9 compared to 5.5 for the buried Lys. New charge-charge interactions did not compensate for these changes, so the change can be attributed to the dehydration (or Born) effect.

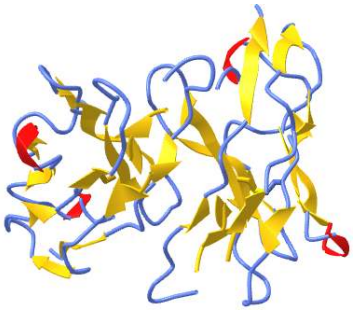
2. Ion-Ion interactions with another charged side chain through Coulombic forces: This effect can be readily observed at the protein's surface. Pace cites a study of RNase S lacking Lys, with a pI of 3.5. Five Asp and Glu residues were replaced on the surface using site-specific mutagenesis with Lys, which lowered the protein's pI to 10.2. At pH 7, the protein without Lys had a charge of -7, while the protein with 5 Lys had a charge of +3. The crystal structures were similar, so Coulombic interactions accounted for the differences in the pKa values of the 11 common side chains. On average, the mutant pKas were 0.75 pH units higher, which makes sense given the mutant's high pI. Calculated pKa values were similar to those determined by NMR. These data are consistent with the idea that Coulombic interactions are the chief cause of pKa changes in surface side chains.

3. Charge-dipole interactions and H bonds: It should be obvious that charge states of ionizable side chains would be adjusted to optimize H bond (and, more generally, charge-dipole) interactions in proteins. If the interactions are optimal in the charged state, pKa values for His and Lys would be increased, and for Asp, Glu, Cys, and Tyr, they would be decreased. Pace cites the buried Asp 76 in RNase T1, in which the Asp is charged but does not form an internal salt bridge. It has a depressed pKa of 0.6 and has 3 H bonds to the side chains of Asn 9, Tyr 11, and Thr 91. Mutants were made to remove the H bonds to see the effect on the pKa of Asp 76. Removing 1, 2, or 3 H bonds changed the pKa to 3.3, 5.1, and 6.4, respectively. The 6.4 value is much higher than the pK_{int}, which can be attributed to the Born effect.

4.7.3.4: Quaternary Structure

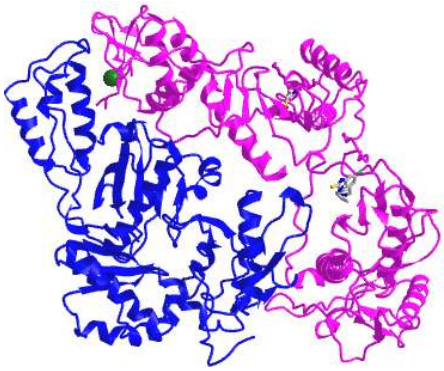
The primary structure is the linear sequence of the protein. The secondary structure is the repetitive structure formed from H-bonds among backbone amide H and carbonyl O atoms. The tertiary structure is the protein's overall 3D structure. Quaternary structure is the overall structure that arises when separate protein chains aggregate with themselves to form homodimers, homotrimers, or homopolymers, or aggregate with different proteins to form heteropolymers. Most protein subunits in a larger protein displaying quaternary structure are held together by noncovalent interactions (intermolecular forces). However, in some cases, disulfide bonds also hold them together (an example includes immunoglobulins).

Figure **4.7.3.3** shows an interactive iCn3D model of a homodimer, the variable domain of the T cell receptor delta chain (1tvd). Carefully rotate the model to see the two identical chains held together by noncovalent interactions.



NCBI iCn3D Figure **4.7.3.3**: variable domain of the T cell receptor delta chain (1tvd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yN6B43P7tvHcR7

Figure **4.7.3.4** shows an interactive iCn3D model of a heterodimer, reverse transcriptase (1rev). Carefully rotate the model to see the two identical chains held together by noncovalent interactions.



NCBI iCn3D Figure **4.7.3.4**: variable domain of the T cell receptor delta chain (1tvd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...2mmkgdR9Z2dTQ6

4.7.3.5: Macromolecule Oligomer Formation and Symmetry

Many proteins are found in aggregated states and have quaternary structures. Hemoglobin consists of two alpha and two beta monomers (or protomers), which assemble to produce the biologically relevant heterotetrameric protein. A given monomer can self-aggregate to form homo-oligomers such as dimers (M_2), trimers (M_3), tetramers(M_4), or higher oligomers (M_n). Oligomers often exhibit symmetry in the geometric arrangement of their subunits. Symmetry is an important component of many kinetic models for catalysis.

Most oligomeric proteins contain protomers arranged symmetrically. What mechanism determines whether a monomeric protein forms a homo-oligomer? Why do they stop at a certain n value? Can proteins be engineered to do so? If a mutation can induce oligomer formation, fewer mutations would be required to produce a symmetric oligomer from subunits. Why? A single mutation in a single monomer would be represented n times in a single oligomer of n monomers. This fact probably underlies oligomers' exquisite symmetry. Hence, a basic understanding of protein oligomer symmetry is necessary.

When studying small molecules, chemists describe symmetry using mathematical symmetry operations and elements, which are useful for structure analysis and molecular spectroscopy. These concepts are usually first encountered in physical and inorganic chemistry classes. They are a bit complicated, so we will offer a limited introduction.

A symmetry operation is a movement of an object, such as a molecule, that results in an identical, superimposable molecule. Each operation has a symmetry element (point, line, or plane) about which the motion occurs. Some examples are shown in Table 4.7.3.2 below:

Element (with Jmol link)	Operation
inversion center (i)	projection through center (point) of symmetry of point x,y,z to point -x,-y,-z
proper rotation axis (Cn)	rotation around a Cn axis by 360°/n where C denotes Cyclic
horizontal (σ_h) and vertical (σ_v) symmetry plane (reflection)	reflection across a horizontal (h) or vertical (v)plane
improper rotation axis (Sn)	rotation around a Sn axis by 360°/n followed by reflection in a plane perpendicular to the axis.

Table 4.7.3.2: Symmetry Elements and Operations

Luckily for students trying to apply these rules to protein oligomers, biomolecules made up of chiral monomers containing L-amino acids cannot be converted to identical structures by inversion or reflection, since the chirality of the monomer would change. For proteins, this would entail an L-to-D amino acid change. That excludes all but proper rotation axes (Cn) from the list above.

A **point group** is a collection of symmetry operations that define the symmetry about a point. The four types of symmetries around a point are those described above: rotational symmetry, inversion symmetry, mirror symmetry, and improper rotation. We'll consider two versions of rotational symmetry.

4.7.3.5.1: Cyclic (Cn) - Single Cn rotation axis.

These are very common in proteins that form dimers, trimers, tetramers, etc, of identical monomers. These are called **homo n-mers**. In this point group, note that the n in Cn is equal to the number of monomers, and the rotation angle is 360°/n. Figure 4.7.3.5, adapted from Voet and Voet, shows a cartoon model for C_2 symmetry.

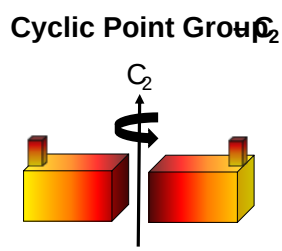
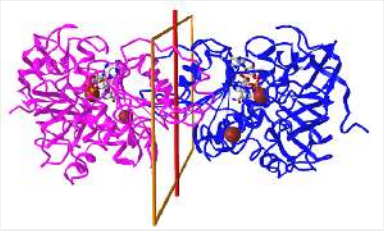
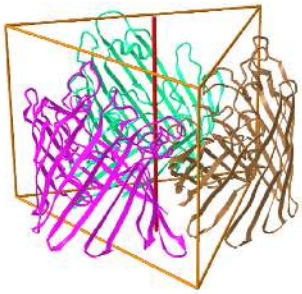
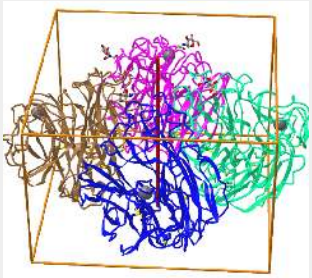
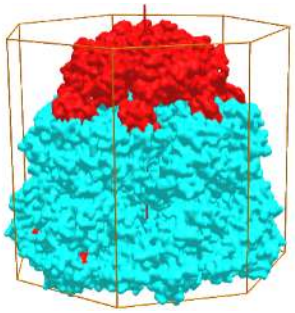


Figure 4.7.3.5: Cartoon model for C_2 symmetry (adapted from Voet and Voet)

Here are some protein examples with C_n symmetry, with the symmetry axis shown as a red vertical line.

Symmetry/ homo n-mer **Protein (**click the link for new window)	NCBI iCn3D **Symmetry** click for the popup model
C2: homo 2-mer alcohol dehydrogenase (1HSO)	
C3: homo 3-mer porin (2POR)	
C4: homo 4-mer neuramindase (2HTY)	

Symmetrical complexes can also have empty interior volumes designed for specific functions. Take, for example, the human mitochondrial Hsp60-Hsp10 chaperonin complex (6MRD), which assists in mitochondrial protein folding. It is a hetero 14-mer (A_7B_7) and displays C_7 symmetry, as shown in Figure **4.7.3.6**, an interactive iCn3D model below. The cyan surface is from 7 A monomers, while the red is from 7 B monomers.



NCBI iCn3D Figure **4.7.3.6**: Hsp60-Hsp10 chaperonin complex (6MRD) with C_7 symmetry (Copyright; author via source). Click the image for a popup (long load) or use this external link:(long load) https://structure.ncbi.nlm.nih.gov/i...jEEY5tu19nXcp6

4.7.3.5.2: Dihedral (Dn) - Mutually perpendicular rotation axes.

These display higher symmetry as they contain (a) C2 axis(es) perpendicular to a single Cn axis. The minimal number of subunits is n. Most protein oligomers fall into this category. The packing (or asymmetric) unit does not have to be a single monomer; it can

be a heterodimer. D_n symmetries are more challenging to visualize, but structures in the PDB conveniently provide the type of global symmetry and stoichiometry when symmetry is present.

A D2 point group has one C2 axis, two perpendicular C2 axes, and four monomers (like Hb). These proteins can dissociate into two dimers (e.g., two α/β dimers in Hb). Note that a different arrangement of 4 monomers could produce an oligomer with C4 symmetry instead of D2. Also note the hemoglobin, a tetramer ($\alpha_2\beta_2$) displays **pseudo** D2 symmetry since the α and β subunits are slightly different in sequence but their folds are almost identical. It also displays C_2 symmetry.

Figure **4.7.3.7**, adapted from Voet and Voet, shows a cartoon model for D_2 symmetry.

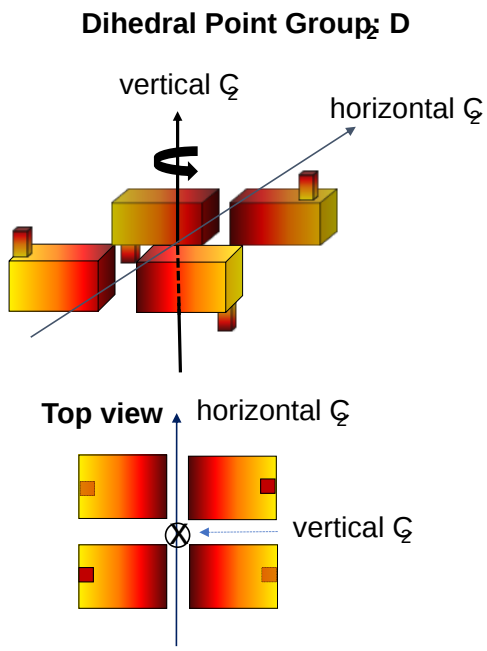
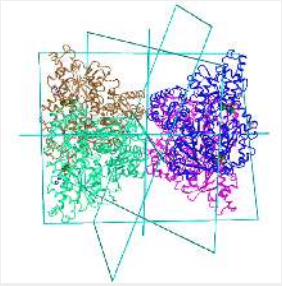
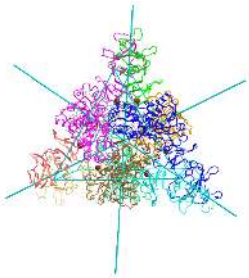
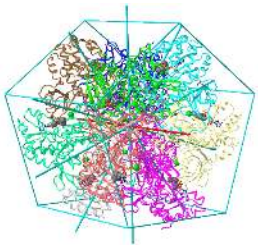


Figure **4.7.3.7**: D_2 symmetry (adapted from Voet and Voet)

A D4 point group has 1 C4 axis and 4 C2 axes, and 2n=8 subunits. An example of a D4 point group is ribulose bisphosphate carboxylase/oxygenase (RuBisCO), which has eight subunits (where a subunit, or more technically, the asymmetric subunit, is a dimer of a small and large molecular weight protein). This point group could arise from the quaternary structure of two C4 tetramers or four C2 dimers.

Here are some homo n-mer protein examples with D_n symmetry.

Symmetry/ homo n-mer **Protein (**click link for new window)	NCBI iCn3D **Symmetry** click for the popup model
D2: homo 4-mer phosphofructokinase (4XYJ)	
D3: hetero 12-mer A_6B_6 asparatate carbamoyltransferase **1Q95**	

D5: homo 10-mer glutamine synthetase (2OJW)	

4.7.3.5.3: Cubic Groups

Now, let's consider some cases in which many rotation axes exist that fit the symmetry of simple geometric shapes called the Platonic solids, in which all faces, edges, and angles are congruent. The five Platonic shapes described by Plato are shown inscribed in spheres in Figure **4.7.3.8**:

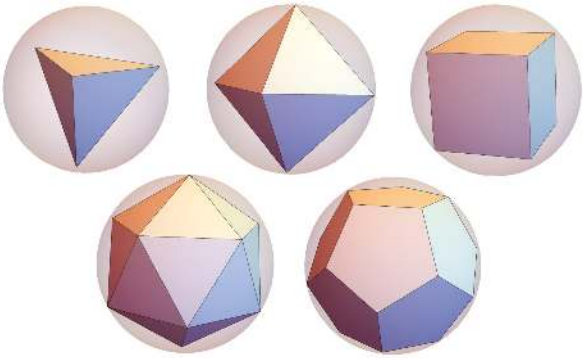


Figure **4.7.3.8**: The five Platonic solids (from left to right): tetrahedron (4 sides), octahedron (8 sides), cube (6 sides), icosahedron (20 sides) and the dodecahedron (12 sides). Quantum 4:293 (2020) DOI: 10.22331/q-2020-07-09-293. Creative Commons License: CC BY 4.0

These can also be inscribed in cubes. A tetrahedron and octahedron (actually the overlap of 2 tetrahedra) are shown inscribed in a cube in Figure 4.7.3.9.

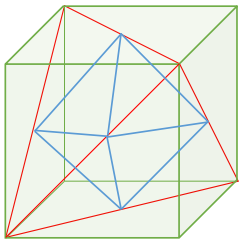


Figure 4.7.3.9: Tetrahedron (red) and Octahedron (blue) inscribed in a Cube (green).

Since all of the Platonic solids can be inscribed in a cube, they all have basic **cubic symmetry**. Cubes have 13 symmetry axes: three C4 axes passing through the centers of opposite faces, four C3 axes passing through opposite vertices (diagonals), and six C2 axes passing through the centers of opposite edges. The other Platonic solids have related C3 axes (diagonals connecting opposite corners for cubes, diagonals from a vertex to the opposite face for tetrahedra, lines connecting two opposite faces for octahedra, for example). Still, they can all be considered part of the cubic point group. The rotation axes for each of the Platonic solids are shown in Table 4.7.3.3below.

# and type rotation axes	# monomers/asymmetric units	Shape
3 C2, 4 C3	12	tetrahedron
6 C2, 4 C3, 3 C4	24	cube/octahedron
15 C2, 10 C3, 6 C5	60	dodecahedron/icosahedron

Table 4.7.3.3: Cn axes and number of monomers in Platonic solids symmetries

We'll consider examples of tetrahedral, octahedral, and icosahedral symmetries, which depend on the overall shape and number of monomers in the functional unit's structure. In some cases, the symmetry of the packed monomers is not perfect. This applies to monomers in clathrin since the monomers can form different structures, and also hemoglobin (in which the four subunits are very similar (alpha and beta) but not identical, as described above.

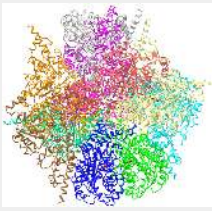
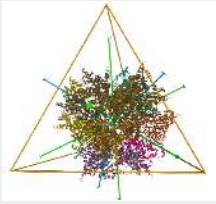
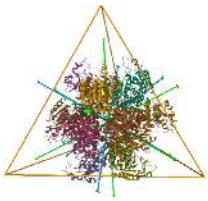
When many monomers form oligomers, homomers seem to be favored evolutionarily over heteromers. One explanation is that, in symmetrical arrangements of monomers, fewer unique subunit-subunit interfaces have evolved to support complementary shape and noncovalent interactions than in heteromers. The same argument applies to the formation of symmetric vs asymmetric arrangements of homo-n-mer.

A final note: The words tetrahedron, cube, octahedron, dodecahedron, and icosahedron refer to the shape of structures. The words tetrahedral, cubic, octahedral, dodecahedral, and icosahedral should be reserved for the type of symmetry. This can be a great source of confusion. Take the case of the E2p, dihydrolipoyl acyltransferase. The homo 24-mer version (from *Azotobacter vinelandii*, 1DPB) has a cube-like shape with octahedral symmetry. In contrast, the homo 60-mer (from *Bacillus stearothermophilus*, 1B5S) has the shape of a dodecahedron with icosahedral symmetry.

It is sometimes difficult to determine the actual biological structure and its symmetry from the crystal structure, given the artificial packing of the protein in the crystal state. Also, aside from icosahedral virus assemblies, there aren't many examples of proteins that exhibit tetrahedral, octahedral, and dodecahedral symmetries. No structure in the Protein Data Bank is listed with a global cubic symmetry. Knowing that they all belong to the cubic point group, we'll describe a few structures with tetrahedral, octahedral, and icosahedral global symmetries.

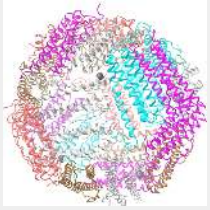
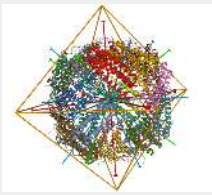
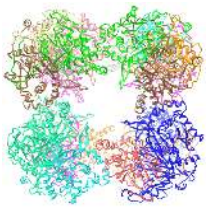
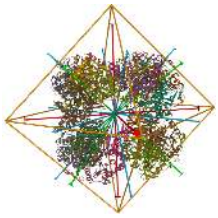
Tetrahedral - 4-sided

A tetrahedron has four C_3 axes - diagonals from four corners/vertices to the opposite faces and three C2 axes, the same as for the cube (since a tetrahedron can be inscribed inside a cube).

Protein (pdb)	**Symmetry/Homo X-mer** (click link)	**Structure - subunits**	**Symmetry**
L-aspartate beta-decarboxylase (2zy2)	tetrahedral/homo 12-mer		
ornithine carbamoyltransferase (1A1S)	tetrahedral, homo 12-mer		

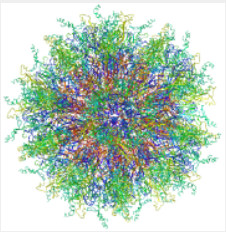
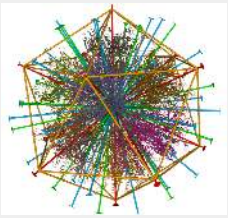
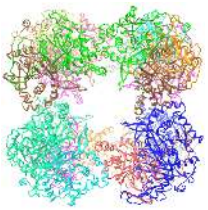
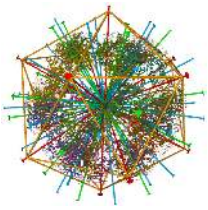
Octahedral - 8-sided

C3 axes - a line connecting two opposite faces of an octahedron. Since an octahedron can be aligned with a cube, it also has the same symmetry axes. Examples include human ferritin, octahedral, and 24 asymmetric subunits. octahedral, homo 24-mer (4V6B) and dihydrolipoyl acyltransferase from *Azotobacter vinelandii*, octahedral, homo 24-mer (1DPB)

Protein (pdb)	**Symmetry/Homo X-mer** (click link)	**Structure - subunits**	**Symmetry**
human ferritin (4V6B)	octahedral/homo 24-mer		
dihydrolipoyl acyltransferase from *Azotobacter vinelandii* (1DPB)	octahedral/homo 24-mer		

Icosohedral - 20-sided, 60 mers

Examples include adenovirus Ad3 virus, Icosahedral, homo 60-mer 4AQQ, and dihydrolipoyl transacetylase from *Bacillus stearothermophilus*, iscosahedral, homo 60-mer (1B5S)

Protein (pdb)	**Symmetry/Homo X-mer** (click link)	**Structure - subunits**	**Symmetry**
Adenovirus Ad3 virus (4AQQ)	Icosahedral/homo 60-mer		
dihydrolipoyl transacetylase from *Bacillus stearothermophilus* (1B5S)	iscosohedral/homo 60-mer		

In summary, our study of symmetry is not in vain since almost half of all proteins appear to form complexes of identical or similar monomers. This probably stabilizes them and adds functional attributes such as kinetic regulation of their activities. Homo n-mers display symmetry, with C_n and D_n being the most common. The cubic symmetries are far less common. What is interesting about cubic symmetries is that structures displaying cubic symmetries have significant empty interior volumes, or "holes," which can be used to encapsulate chemical species. They include the icosahedral viruses, whose interiors contain proteins and viral genomes, and cellular ferritin, which houses up to 4500 Fe^{3+} (oxidized) ions in hydroxide and phosphate complexes. Ferritin delivers Fe^{2+} into cells. Storing iron ions in ferritin prevents their spurious oxidation and precipitation in unregulated environments. The encapsulated available volumes have diameters ranging from about 2-4 nm for tetrahedral, 4-8 nm for octahedral, and 8-18 nm for icosahedral symmetries.

4.7.3.6: Filaments

Proteins, especially those involved in cytoskeletal filaments, can form fibers with helical symmetry, which differs from those described above, since the monomers are at the ends of helical fibers. However, they have the same tertiary structures as those in the middle of the helical fibers and do not contact the same number of monomers as monomers within the oligomer. Hence, they have different microenvironments.

Grueninger et al. address whether oligomerization can be programmed into the genome. Can simple amino acid substitutions lead to oligomerization? Oligomerization can be beneficial (formation of cytoskeleton filaments) or detrimental (formation of fibers in sickle cell anemia and prion disease). Oligomers with long half-lives (for example, cytoskeletal filaments such as actin and tubulin) and short half-lives (for example, proteins associated with transient biological activities) are regulated by oligomer formation.

It has long been noted that if a protein chain forms oligomers, then a single amino acid change in the chain would be found n times in an oligomer of n chains. Mutations could either promote chain contact and oligomer formation or, if in a region involved in subunit association (a contact region), promote dissociation into monomeric or other asymmetric subunit compositions. Experimental work in this field is hampered by the fact that mutants generated by site-specific mutagenesis that favor the monomeric state often fail to fold (due to hydrophobic exposure and aggregation). Studies have shown that most contact areas between monomers or other asymmetric units are hydrophobic, and the contact regions must be complementary in shape. Mutations that replace hydrophobic side chains involved in subunit contact with polar, polar charged, or bulkier hydrophobic side chains would inhibit oligomer formation.

Grueninger et al were able to engineer dimer formation and oligomer formation. First, consider the simplest case of a mutation in a monomer that can produce a dimer with C2 symmetry. This is illustrated below in Figure **4.7.3.10**. The figure illustrates how a mutation that produces a weak interaction in a monomer can also yield a long helical aggregate (which can't be crystallized) lacking symmetry (as described above). A mutation at two could promote either oligomer helix formation or dimerization.

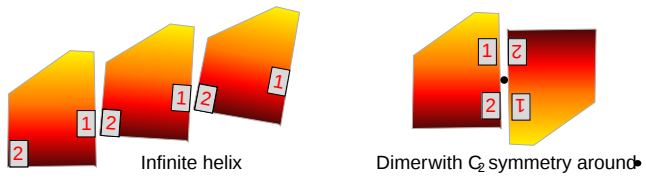


Figure **4.7.3.10**: Mutations causing Dimer with C2 symmetry or Infinite Helix (adapted from Grueninger et al. Science, 319, 206-209 (2008)

It should be noted that mutation could lead to dimer or oligomer formation by producing a more global conformational change in the monomer (not indicated in the example above), which leads to aggregate formation, as we have seen previously in the formation of dimers and aggregates of proteins associated with neurodegenerative diseases (like mad cow disease).

Grueninger produces mutants of two different proteins that showed dimer formation as analyzed by gel filtration chromatography (but did not crystallize, so no 3D structures were determined). In addition, the group modified urocanase, a C2 dimer, at three side chains to form a tetramer with D2 symmetry. Also, they modified L-rhamnulose-1-phosphate aldolase, a C4 tetramer, at a single position to form an octamer with D4 symmetry. The latter two were analyzed through X-ray crystallography. Their work suggests that complex symmetric protein structures arose in nature from simple mutation and evolutionary selection.

4.7.3.7: Summary

This chapter delves into the intricate details of protein tertiary and quaternary structure, emphasizing how a single polypeptide chain achieves its unique three-dimensional conformation and how multiple chains assemble into functional complexes. The discussion begins by exploring the concept of tertiary structure—the overall 3D arrangement of a folded protein. Key to this folding process is the segregation of amino acid side chains: hydrophobic residues tend to be buried within the protein's core, while polar and charged residues are generally exposed to the aqueous environment. Although a simple micelle-like model might suggest perfect segregation, the protein backbone’s inherent polarity and the complex network of noncovalent interactions (including hydrogen bonds, salt bridges, hydrophobic interactions, and aromatic stacking) guide a more nuanced and functionally relevant folding pattern.

Interactive 3D models, such as those of hydroxynitrile lyase and low molecular weight protein tyrosyl phosphatase, illustrate how specific side chains become buried or remain solvent-exposed, reinforcing the role of noncovalent forces in stabilizing the tertiary structure. The chapter also discusses how the local microenvironment within the protein can modulate the pKa values of ionizable side chains, further influencing folding stability and function.

Transitioning to quaternary structure, the chapter explains how individual polypeptide chains (subunits) come together to form multi-subunit complexes. It introduces key symmetry concepts—such as cyclic (Cn), dihedral (Dn), and cubic symmetries—and demonstrates how these symmetrical arrangements not only stabilize protein complexes but also contribute to their functional properties, as seen in hemoglobin, ferritin, and chaperonin complexes. The discussion includes an overview of how complementary hydrophobic interfaces and noncovalent interactions drive protein oligomerization. It highlights how even a single amino acid mutation can dramatically influence the assembly state and symmetry of these complexes.

Overall, this chapter integrates theoretical models with practical structural examples to provide a comprehensive understanding of how proteins fold into their tertiary structures and further assemble into quaternary structures, thereby underpinning the diverse functions of proteins in biological systems.

4.7.3.8: References

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

This page titled 4.7.3: Tertiary, Quaternary, and Symmetrical Structures is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.3: Tertiary, Quaternary, and Symmetrical Structures** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.4: Secondary Structural Motifs and Domains

Learning Goals (ChatGPT o1, 1/26/25)

- **Identify Common Super-Secondary Structural Motifs:**
  - Recognize and describe key motifs such as the alpha-loop-alpha (e.g., helix-turn-helix and EF-hand), beta-hairpin (or beta-turn), Greek key, and beta-alpha-beta motifs.
  - Explain the roles of these motifs in protein function, such as DNA binding in helix-turn-helix and calcium binding in EF-hand domains.
- **Interpret Structural Representations and Topology Maps:**
  - Utilize interactive 3D models (e.g., iCn3D) and 2D wiring/topology diagrams to analyze how secondary structures connect to form super-secondary motifs.
  - Assess the conservation of key residues within motifs (e.g., in the EF-hand of calmodulin) and relate these patterns to functional properties.
- **Relate Motif Structure to Protein Architecture:**
  - Define and describe larger structural motifs (architectures) such as the Rossmann fold, TIM barrel, beta helices, and beta propellers.
  - Explain how these architectures arise from combinations of common secondary structures and discuss their roles in binding cofactors (e.g., NAD^+ in the Rossmann fold) or substrates.
- **Understand the Concept of Domains:**
  - Define a protein domain as an independently folding unit that can also serve as a functional unit.
  - Describe how domains are assembled from recurring motifs and explain the evolutionary mechanisms (duplication, divergence, recombination) that give rise to multidomain proteins.
  - Compare domain classification approaches based on 3D structure (e.g., CATH) and linear sequence alignments (e.g., Pfam).
- **Evaluate Protein Oligomerization and Symmetry:**
  - Discuss how symmetry in protein complexes (cyclic, dihedral, and cubic symmetries) emerges from the assembly of identical or similar domains.
  - Explain the functional implications of symmetric arrangements in oligomers (e.g., kinetic regulation, substrate channeling, and structural stability).
- **Integrate AI and Bioinformatics in Domain Analysis:**
  - Describe the role of modern AI-based tools and databases (e.g., TED) in predicting and classifying protein domains across large datasets such as the AlphaFold database.
  - Evaluate how new domain assignments and novel folds expand our understanding of protein evolution and function.

These learning goals are intended to provide a framework for understanding how recurring structural motifs contribute to the overall protein architecture and function, and how these principles are applied in modern structural biology and bioinformatics.

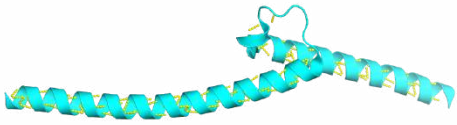
4.7.4.1: Common Structural Motifs

Given the number of possible combinations of 1°, 2°, and 3° structures, one might guess that the 3D structure of each protein is quite distinctive. This is, in general, true. However, similar substructures are found in proteins. For instance, common secondary structures are often grouped into structural motifs called **supersecondary structures**. The same motif is often found in proteins with similar functions (such as proteins that bind DNA, Ca^{2+}, etc). Let's explore some of the common motifs.

4.7.4.1.1: Alpha-loop-Alpha

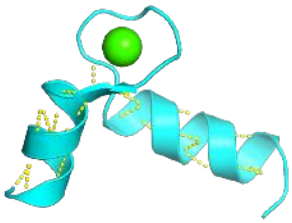
These are found in DNA-binding proteins that regulate transcription and calcium-binding proteins, the motif of which is often called the EF-hand. The loop region in calcium-binding proteins is enriched in Asp, Glu, Ser, and Thr. Why? The EF-hand shown below is from Calmodulin.

Figure **4.7.4.1** shows an interactive iCn3D model of a basic helix-turn-helix from the c-Myc protein (1NKP). The iCn3D model shows helices interacting with the major groove of DNA, depicted in spacefill.



NCBI iCn3D Figure **4.7.4.1**: Basic helix-turn-helix from the c-Myc protein (1NKP). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...kDv9DGzWWWoMZ8

Figure **4.7.4.2** shows an interactive iCn3D model of the "EF-hand" from the calcium-binding protein calmodulin (1cll)



NCBI iCn3D Figure **4.7.4.2**: EF hand from Calmodulin (1cll): Secondary Structure Motif. (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vCYH7EX4sVvtu6

The EF Hand can be envisioned as a hand gripping a ball (calcium ion) with the index finger and thumb representing alpha helices, as shown in Figure **4.7.4.3**.

Figure **4.7.4.3**: "EF hand" gripping a ball

The EF-hand motif of calmodulin is used in various Ca^{2+} binding proteins. Figure **4.7.4.4** shows the alignment of the first 50 residues of human calmodulin with four other human calcium-binding proteins. The EF-hand (F12-L29) of calmodulin consists of the second half of the first helix (F12-L18), an intervening loop (F19-T28), and the second helix (T29-L29). Sometimes, it is annotated to encompass a larger stretch (8-43)

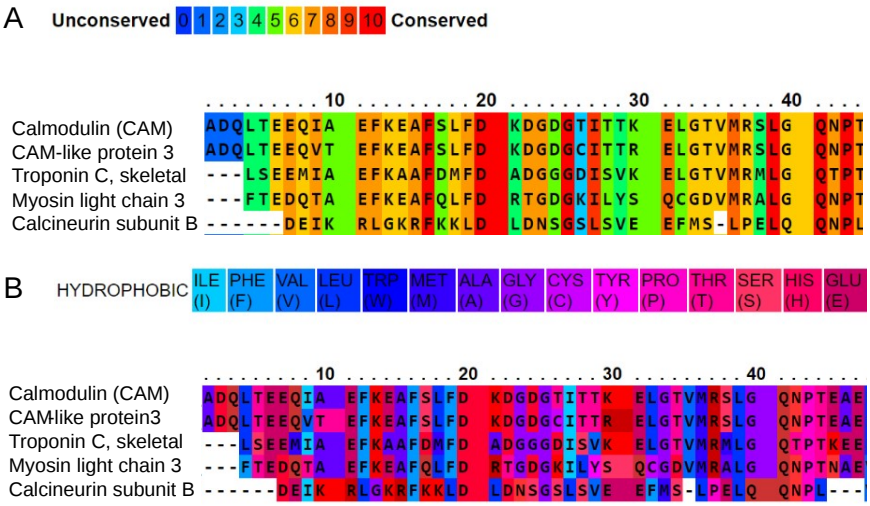
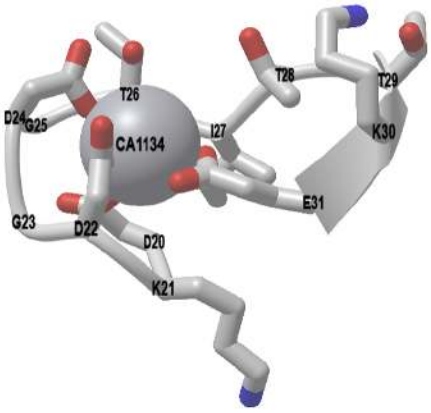


Figure 4.7.4.4: Amino acid sequence homology of Ca^{2+} binding protein

Part A shows the degree of conservation of amino acids in this first Ca^{2+}-binding EF-hand. Part B shows the general conservation of key hydrophobic (F12, F19, I27), as well as those of similar polarity (36 and 39)

Figure 4.7.4.5 shows an interactive iCn3D model of a bound calcium ion and interacting amino acids in human calmodulin, with key amino acids labeled.



NCBI iCn3D Figure 4.7.4.5: Bound calcium ion and interacting amino acids in human calmodulin (1cll) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?bCf4mtNbk4kjkCHw6

Hover over the amino acid side chains that are coordinating the Ca^{2+} ion. Are they what you would expect?

You learned about transition metal complexes and their geometries (tetrahedral with 4 ligand interactions, trigonal bipyramidal with five, and octahedral with six) in introductory chemistry. Ca^{2+} (a nontransition Group II metal ion) has a diameter of 100 pm compared to Fe^{2+} and Fe^{3+} of about 76 and 68 pm, respectively, for octahedral geometry with six ligand interactions. Hence, more ligand groups can crowd around the calcium ion. In the calmodulin EF hand, the coordination number is 7 with a distorted pentagonal bipyramidal geometry, allowing high affinity for the large Ca^{2+} ion and conformational flexibility. Mg^{2+}, a smaller

nontransition Group II metal ion, is typically octahedral, precluding Ca^{2+} binding to Mg^{2+}-binding molecules. Figure **4.7.4.5*b*** shows the distorted pentagonal bipyramidal geometry of Ca^{2+}-EF hand interactions.

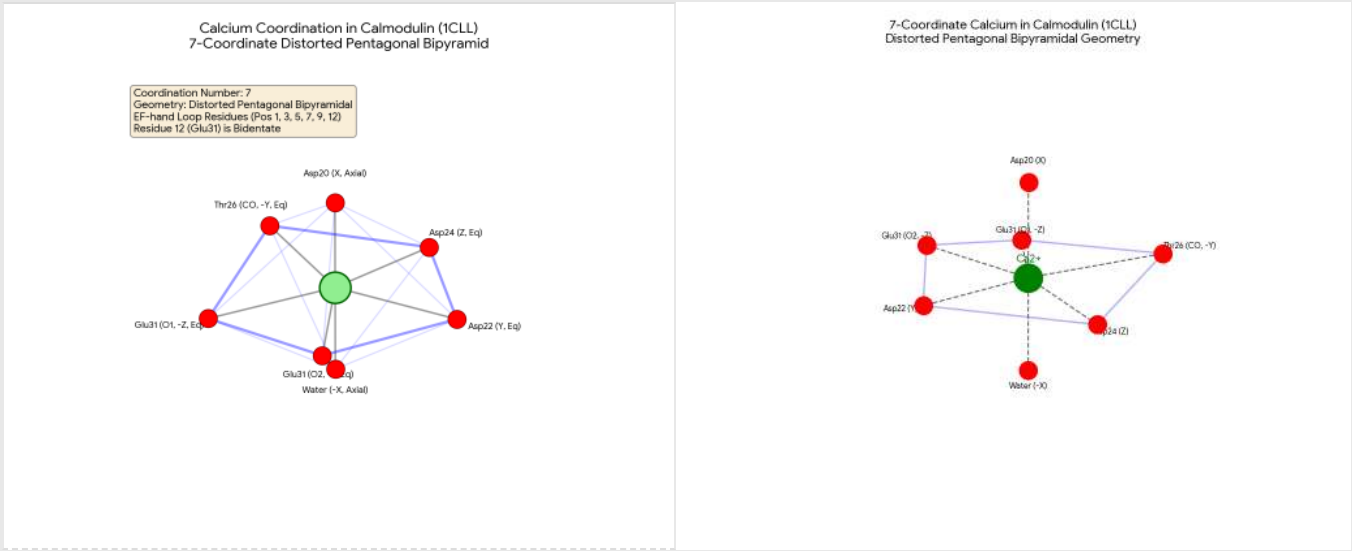


Figure **4.7.4.5*b*** shows the distorted pentagonal bipyramidal geometry of Ca^{2+}-EF hand interactions. Image created by Google Gemini.

The geometry is significantly distorted because Glu31 acts as a bidentate ligand and "pinches" the geometry of the complex.

A linear connectivity "wiring" diagram showing a secondary structure connected by connecting regions is shown in Figure **4.7.4.6**. This wiring diagram shows a 2-residue beta strand, which is too short to be considered a true strand.

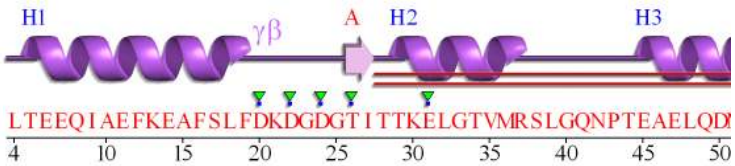


Figure **4.7.4.6**: Wiring Diagram for first 50 amino acids in calmodulin

A more complicated 2D topology map is shown in Figure **4.7.4.7**. In this case, it is linear, given the small section of amino acids depicted. We will see more complicated 2D topology maps with more complicated structures below.

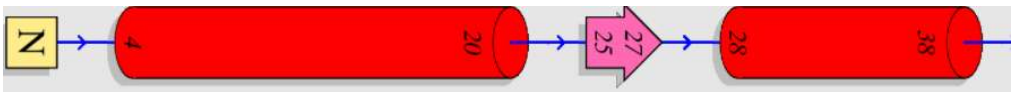


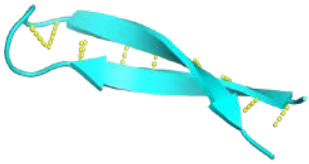
Figure **4.7.4.7**: 2D Topology Map EF Hand Calmodulin

It is presented on its side to save space on this page.

4.7.4.1.2: Beta-hairpin or beta-turn

This motif is present in most antiparallel beta structures, both as an isolated ribbon and as part of beta sheets.

Figure **4.7.4.8** shows an interactive iCn3D model of the beta hairpin from bovine pancreatic trypsin inhibitor (1k6u).



Beta hairpin from bovine pancreatic trypsin inhibitor (1k6u)

Figure 4.7.4.8: Beta hairpin from bovine pancreatic trypsin inhibitor (1k6u) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...eMFdHkGogJHCCA

Figure 4.7.4.9 shows the 2D homology map for the beta-hairpin.

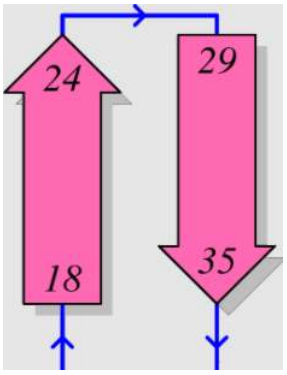


Figure 4.7.4.8: 2D homology map of beta-hairpin from bovine pancreatic trypsin inhibitor (1k6u)

4.7.4.1.3: Greek Key

The "Greek Key" symbol represents infinity and the eternal flow of things and resembles, in part, primitive keys. The Greek Key motif in proteins can be seen in the structure of antiparallel beta sheets, ordering four adjacent antiparallel beta strands, as shown in Figure 4.7.4.9. The figure also shows the repetitive Greek key, which you will see many times if you visit Greece and tour its antiquities.

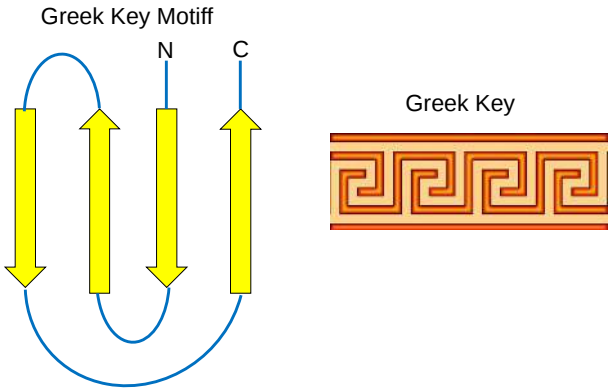


Figure 4.7.4.9: Greek Key Motif

Figure 4.7.4.10s shows a partial 2D topology map of Staphylococcus nuclease (2SNS).

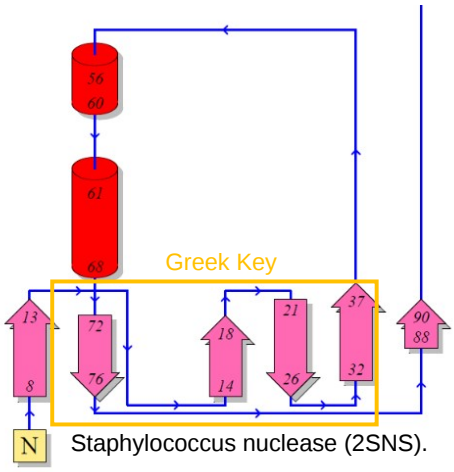


Figure 4.7.4.10: 2D topology map of Staphylococcus nuclease (2SNS)

Figure 4.7.4.11 shows an interactive iCn3D model of the Greek Key motif from Staphylococcus nuclease (2SNS). The involved beta strands are shown in yellow.

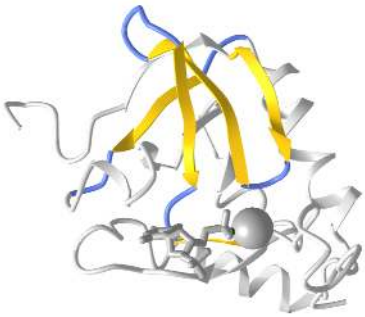


Figure 4.7.4.11: Greek Key motif from Staphylococcus nuclease (2SNS) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...x2ef4xpttXrFb9

4.7.4.1.4: Beta-Alpha-Beta

The motif is a common way to connect two **parallel** beta strands, unlike beta hairpins, which connect antiparallel beta strands.

Figure 4.7.4.12 shows an interactive iCn3D model of the beta-alpha-beta structure from triose phosphate isomerase (1amk).

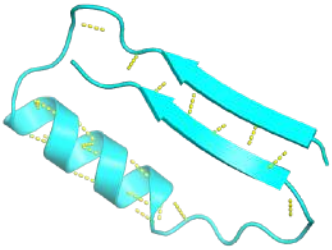


Figure 4.7.4.12: Beta alpha beta motif from triose phosphate isomerase (1amk) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8L3VTKtYjVPy97

Figure 4.7.4.13 shows the 1D wiring diagram for the first beta-alpha-beta motif in triose phosphate isomerase.

Figure 4.7.4.13: 1D wiring diagram for the first beta-alpha-beta motif in triose phosphate isomerase

Figure 4.7.4.14 shows the 2D topology diagrams showing this motif.

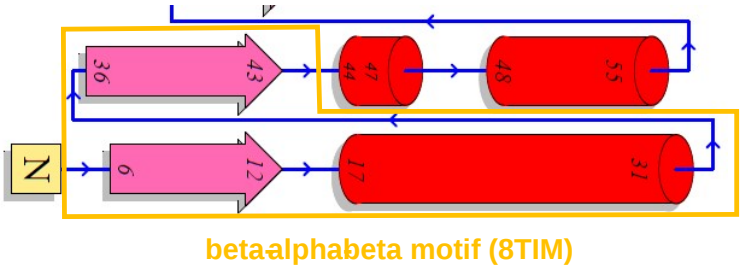


Figure 4.7.4.14: 2D topology diagrams of beta-alpha-beta motif (8tim).

4.7.4.2: Larger Structural Motifs - Protein Architecture

Some proteins combine larger secondary and supersecondary structural components, often repeatedly, to produce more complex structures. We've seen this with larger twisted sheets and beta barrels, such as the TIM barrel. Let's consider three of these as examples of protein architectures, without considering connectivity within the protein.

4.7.4.2.1: The Rossman Fold

Structural motifs can serve particular functions within proteins, such as enabling the binding of substrates or cofactors. For example, the Rossmann fold is responsible for binding to nucleotide cofactors such as nicotinamide adenine dinucleotide (NAD^+), as shown in Figure 4.7.4.15. The Rossmann fold comprises six parallel beta strands forming an extended beta sheet. The first three strands are connected by α-helices, resulting in a beta-alpha-beta-alpha-beta structure. This pattern is duplicated once to produce an inverted tandem repeat with six strands. Overall, the strands are arranged in the order of 321456 (1 = N-terminal, 6 = C-terminal). Five stranded Rossmann-like folds are arranged in the sequential order 32145. The overall tertiary structure of the fold resembles a three-layered sandwich, with the filling composed of an extended beta sheet and the connecting parallel alpha helices forming the two slices of bread.

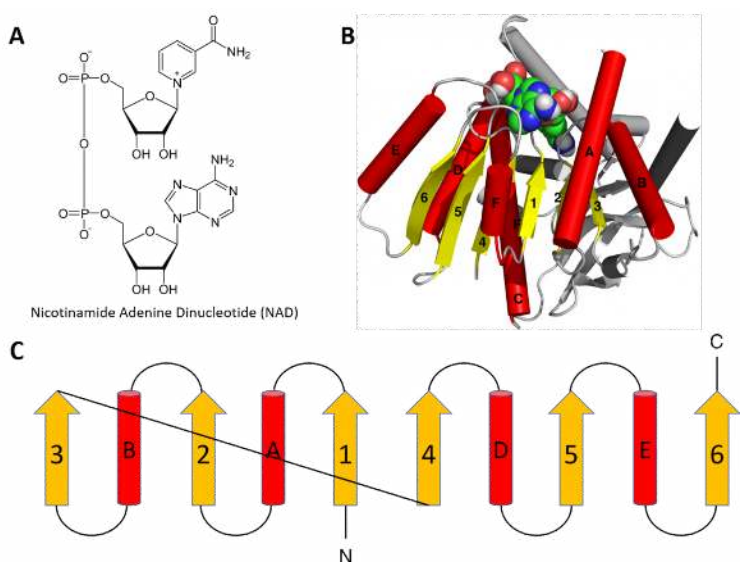


Figure 4.7.4.15 The Rossman Fold. (A) Structure of Nicotinamide Adenine Dinucleotide (NAD^+) (B) Cartoon diagram of the Rossmann Fold (helices A-F red and strands 1-6 yellow) from *E. coli* malate dehydrogenase enzyme. The NAD^+ cofactor is shown binding as the space filling molecule. (C) Schematic diagram of the six stranded Rossmann fold.

Image modified from: Boghog

One feature of the Rossmann fold is its cofactor-binding specificity. The most conserved segment of Rossmann folds is the first beta-alpha-beta segment. Since this segment interacts with the ADP portion of dinucleotides such as FAD, NAD, and NADP, it is also called an "ADP-binding beta-beta fold."

Figure **4.7.4.16** shows an interactive iCn3D model of the Rossman fold of malate dehydrogenase (5KKA) from E. Coli. The beta strands (yellow) connecting the alpha helices (red) and the coil (blue) of the Rossman fold are shown in the context of the rest of the monomeric protein, which is shown in gray.

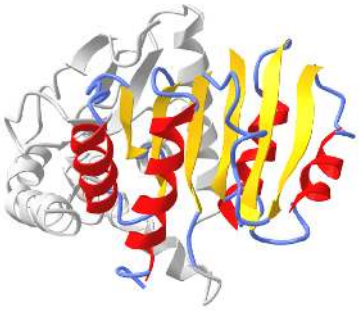


Figure **4.7.4.16**: Rossman fold of malate dehydrogenase (5KKA) from E. Coli (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/icn3d/share.html?igcXKbc8aU96eG896

4.7.4.2.2: The TIM barrel revisited.

Interestingly, similar structural motifs do not always have a common evolutionary ancestor and can arise from convergent evolution. This is the case with the TIM Barrel, a conserved protein fold composed of eight α-helices and eight parallel β-strands, arranged along the peptide backbone. It is illustrated in Figure **4.7.4.17**. The structure is named after triosephosphate isomerase, a conserved metabolic enzyme. TIM barrels are one of the most common protein folds. One of the most intriguing features among members of this class of proteins is that although they all exhibit the same tertiary fold, there is very little sequence similarity between them. At least 15 distinct enzyme families use this framework to generate the appropriate active site geometry, always at the C-terminal end of the eight parallel beta-strands of the barrel.

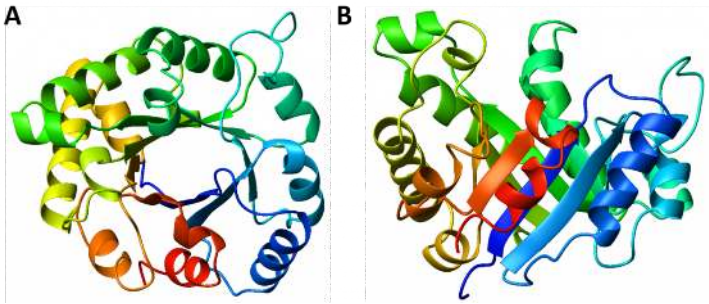


Figure **4.7.4.17** The TIM Barrel. TIM barrels are considered α/β protein folds because they include an alternating pattern of α-helices and β-strands in a single domain. In a TIM barrel, the helices and strands (usually 8 of each) form a solenoid that curves around to close in a doughnut shape, topologically known as a torus. The parallel β-strands form the inner wall of the doughnut (hence, a β-barrel), whereas the α-helices form the outer wall of the doughnut. Each β-strand connects to the next adjacent strand in the barrel through a long right-handed loop that includes one of the helices, so that the ribbon N-to-C coloring in the top view (A) proceeds in rainbow order around the barrel. The TIM barrel can also be thought of as composed of 8 overlapping, right-handed β-α-β supersecondary structures, as shown in the side view (B).

Image modified from: WillowW

Although the ribbon diagram of the TIM Barrel shows a hole in the protein's central core, the amino acid side chains are not shown in this representation (Figure 2.26). The protein's core is tightly packed, mostly with bulky hydrophobic amino acid residues. However, a few glycines are needed to allow wiggle room for the highly constrained center of the eight approximate repeats to fit

together. The packing interactions between the strands and helices are also dominated by hydrophobicity, and the branched aliphatic residues valine, leucine, and isoleucine comprise about 40% of the total residues in the β-strands.

The figure **4.7.4.18** below shows an interactive iCn3D model of the TIM barrel (1WYI) from Chapter 4.2).

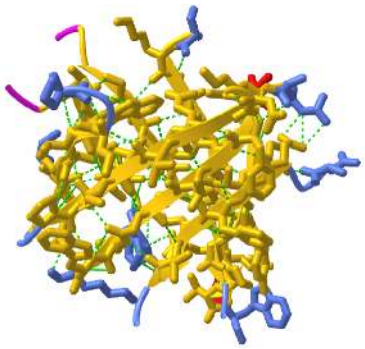


Figure **4.7.4.18**: TIM barrel (1WYI) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7dki3WjRDtiaRA

As our knowledge of the myriad structural motifs found in nature's treasure trove of protein structures continues to increase, we gain insight into how protein structure relates to function and are better enabled to characterize newly acquired protein sequences using *in silico* technologies.

4.7.4.2.3: Beta Helices

These right-handed parallel helical structures consist of a contiguous polypeptide chain with three parallel beta strands separated by three turns, forming a single rung of a larger helical structure, which might contain as many as nine rungs. The intrastrand H-bonds are between parallel beta strands in separate rungs. These are prevalent in proteins from pathogens (bacteria, viruses, and toxins) that facilitate pathogen binding to a host cell.

Figure **4.7.4.19** shows an interactive iCn3D model of the C-terminal fragment of the phage T4 GP5 beta helix (4osd).

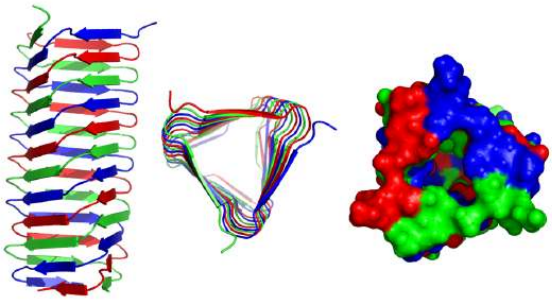


Figure **4.7.4.19**: Beta helix from the C-terminal fragment of the phage T4 GP5 (4osd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...jmRABoThf5dXJ9

Beta helices are found in the following organisms (with the diseases they cause in humans): *Vibrio cholerae* (cholera), *Helicobacter pylori* (ulcers), *Plasmodium falciparum* (malaria), *Chlamydia trachomatis* (VD), *Chlamydophila pneumoniae* (respiratory infection), *Trypanosoma brucei* (sleeping sickness), *Borrelia burgdorferi* (Lyme disease), *Bordetella parapertussis* (whooping cough), *Bacillus anthracis* (anthrax), *Neisseria meningitides* (meningitis), and *Legionella pneumophila* (Legionnaire's disease).

4.7.4.2.4: Beta Propellers

Proteins with this structure have 4-8 blade-shaped beta sheets arranged around a central axis, forming an active site shaped like a funnel.

Figure **4.7.4.20** shows an interactive iCn3D model of the C-terminal domain of Tup1 (1ERJ), a yeast transcription factor, which has a seven-bladed beta propeller. Each blade contains a **WD40 repeat** sequence (around 40 amino acids) that often ends in tryptophan-aspartic acid (W-D). The particular protein has four WD dipeptide sequences, shown as sticks in CPK colors.

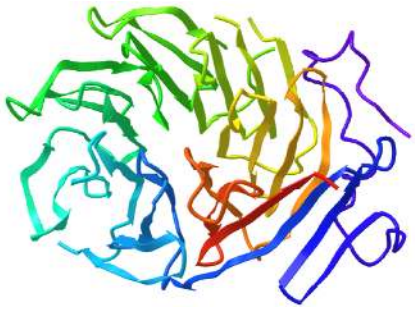


Figure 4.7.4.20: Seven-bladed beta propeller from the C-terminal domain of Tup1 (1ERJ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?ydS1vZedDUPupfN28

The funnel provides binding sites for proteins and other molecules, with the ones with more blades usually acting as enzymes.

4.7.4.3: Domains

Domains are the fundamental unit of tertiary (3°) structure. Domains can be considered a chain or part of a chain that can independently fold into a stable tertiary structure. Domains are units of structure but can also be units of function. Some proteins can be cleaved at a single peptide bond, forming two domains. Often, these can fold independently of each other, and sometimes, each unit retains an activity it had in the uncleaved protein. Sometimes, binding sites on proteins are found at the interface between structural domains. Many proteins appear to share functional and structural domains, suggesting that the DNA encoding each domain may have arisen from the duplication of a primordial gene with a particular structure and function.

Evolution has increased complexity, requiring proteins to have new structures and functions. Increased and diverse protein functionalities have been achieved by adding domains to base proteins. Chothia (2003) has defined a domain in an evolutionary and genetic sense as "an evolutionary unit whose coding sequence can be duplicated and/or undergo recombination." Proteins range from small with a single domain (typically from 100-250 amino acids) to large with many domains. From recent analyses of genomes, new protein functionalities appear to arise from the addition or exchange of other domains, which, according to Chothia, result from

- duplication of sequences that code for one or more domains
- divergence of duplicated sequences by mutations, deletions, and insertions that produce modified structures that may have helpful new properties to be selected
- recombination of genes that form a novel arrangement of domains.

Structural analyses show that about half of all protein-coding sequences in genomes are homologous to other known protein structures. There appear to be about 750 different families of domains (i.e., small proteins derived from a common ancestor) in vertebrates, each with about 50 homologous structures. About 430 of these domain families are found in all solved genomes.

Proteins with multiple domains are less likely to misfold if each domain can fold somewhat autonomously. In addition, they provide a myriad of binding sites, increasing the number of biological functions expressed by a single protein. Multidomain proteins can also express multiple catalytic activities, allowing for a reaction product from one domain to diffuse to another catalytic domain (or interface between domains). This would reduce the dimensionality of the search for a substrate from 3D to more of a 1D or 2D search, enormously speeding up the net reaction. The process is often called **substrate channeling**.

Figure 4.7.4.21 shows an interactive iCn3D model of the three domains of the enzyme pyruvate kinase (1pkn). These include a nucleotide (ADP/ATP) binding domain (blue) made of beta strands, a substrate binding domain (green) in the middle composed of alpha/beta structure, and a regulatory domain (red) composed of alpha/beta structure. These domains were analyzed by a web program called CATH-Gene3D.

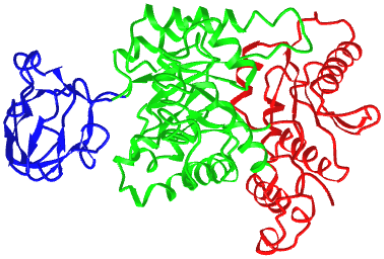
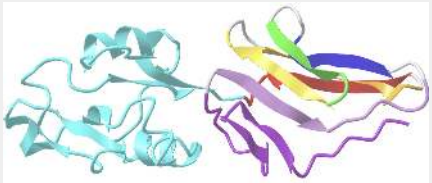
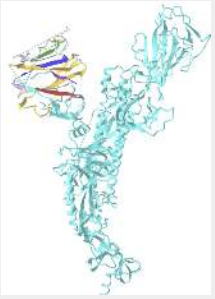
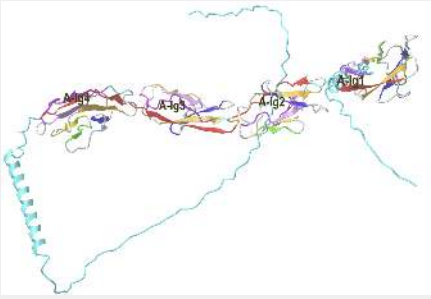


Figure 4.7.4.21: three domains of the enzyme pyruvate kinase (1pkn) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...3eMGhKSto7ceW7

One ubiquitous domain is the **Immunoglobulin Fold (IGF), also called the Immunoglobulin Domain (IgD)**. They are abundantly found in immune proteins, cell surface proteins involved in recognition, and other proteins. They are described in detail in Chapter 5.4: Recognition of Self and Nonself - The Immune System. Here are some images and iCn3D displays of proteins with Ig Domains in bacteria (left), viruses (center), and humans (right). (some have long load times)

Intimin-190 (Int190) from Enteropathogenic E. coli (1E5U)	Ig Domain in SARS-CoV-2 Spike Glycoprotein (6VXX)	AlphaFold ID Carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1) (P13688)
 Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?AcPowvr2Uz37Y1Rh7	 Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...hPGzYJF9wV7gt5	 Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GXmU8dwdNQ2w69

Recent Updates: 11/1/2024

With recent advances in AI-driven prediction of protein structure and function, there is a greater need to refine and develop programs that can determine the domain structures of the over 200 million protein structures in the AlphaFold database. Until recently, there were two different ways to determine domain structures in proteins:

- based on the 3D structure of a protein. The program CATH does this.
- based on 1D (linear) sequences. An example of a program that uses this approach is Pfam.

CATH classifieds protein structure based on the following hierarchy of organization: Class, Architecture, Topology, and Homologous Superfamilies

- **C**lass: the highest level of organization, which consists of four classes - mainly alpha, mainly beta, alpha-beta, and a few secondary structures
- **A**rchitecture (40 types): describes the shape of the domain based on secondary structures, but doesn't describe how they are connected. Ex: beta-barrel, beta-propeller

- **T**opology (or fold group, 1233 types): Members of topology groups share a common fold or topology in the "core" of the domain structure.
- **H**omologous Superfamilies (2386 types): These groups are homologous in sequence or structure and derive from a common precursor gene or protein.

Pfam uses multiple alignments of sequences.

The pyruvate kinase example above shows three structural domains. Pfam finds two major domains: a pyruvate kinase beta-barrel domain and an alpha/beta domain. The domains determined by both programs show about a 75% overlap.

At a simpler level, domains are built from the kinds of motif structures we discussed above. Since proteins are highly compact structures, their organization can be thought of as closely packed motifs, but not all possible combinations are observed. For example, if you have two beta hairpins next to each other to form a 2-unit Greek key, there are 24 possible ways to connect them, but only eight are common. The two below account for more than the sum of the other 22. These are shown in Figure **4.7.4.22**.

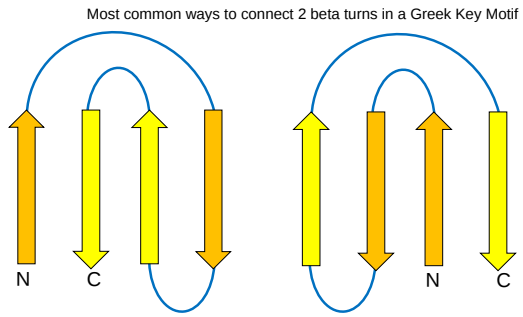
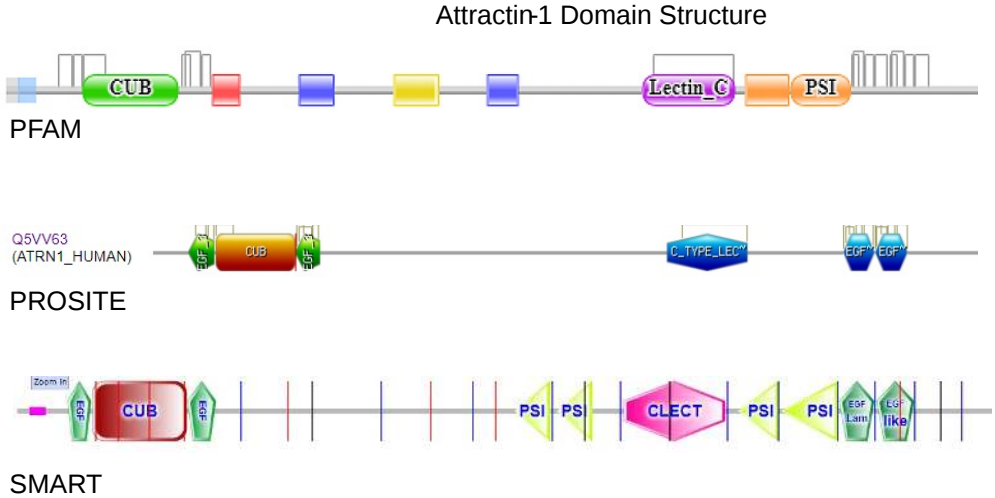


Figure **4.7.4.22**: Most common ways to connect two beta turns in Greek Key Motif

Figure **4.7.4.23** shows an example of the architecture of the multi-domain protein, human Attractin-like protein 1. This protein is an example of a lectin, a carbohydrate-binding protein, which we will explore in a subsequent chapter. It binds Ca^{2+}, so it is considered a C-Lectin. Three different programs were used to analyze the domain structure.



4.7.4.3.1:

Figure **4.7.4.23**: Architecture of the multi-domain protein, human Attractin-like protein 1

A new AI-based "TED - Encyclopedia of Domains" has been developed to identify and classify around 365 million domains in the AlphFold database. Around 1/3 of these were not predicted through 1D sequence alignments and comparisons. Around 3/4 of the "nonredundant" domains were similar to domains predicted by the 3D structure alignments of CATH. TED identified new domain interactions between superfamilies and many new protein folds. New folds across life likely suggest a common function, whereas new folds with a given lineage suggest evolutionary changes.

Some new folds had higher symmetry, including the beta-propeller, which repeats to give high symmetry (see below for lower C3 symmetry). Additional AI-based sequence (1D) information was used to help infer function. For example, putative Zn^{2+} finger-like binding sites containing 2 Cys and 2 His but lacking a traditional Zn-finger motif were found. Most putative heme binding sites contained the CXXCH motif found in heme c. Inspection of the structures predicted to have these features supported their likely functions.

CATH and TED are 3D structure-based, so comparing their domains and domain interactions is warranted. TED found over 27 million examples of interacting domains, with about 14,000 interacting superfamilies pairs compared to around 200,000 with 5000 pairs for CATH

Figure **4.7.4.24** below shows the classification of TED domains using the CATH hierarchy.

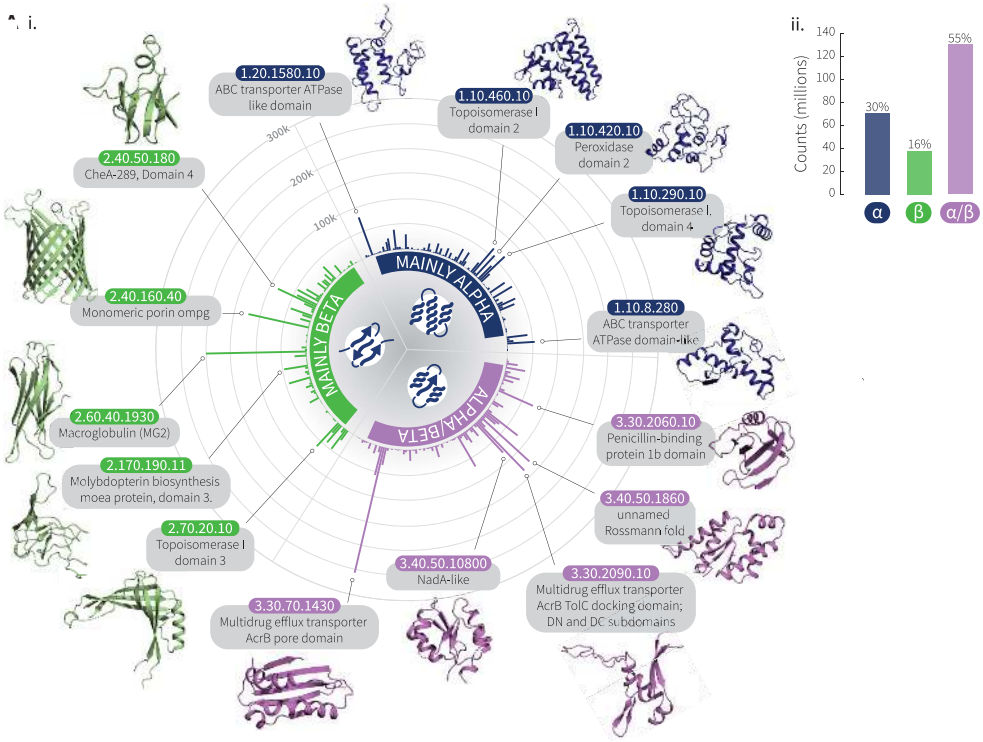
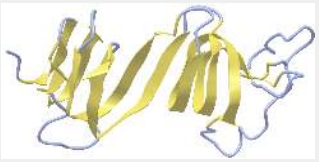
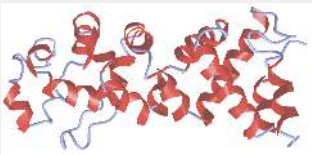
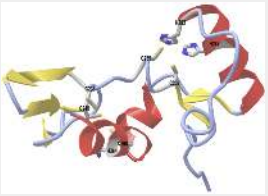
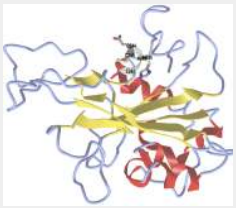


Figure **4.7.4.24**: **Classification of TED domains using the CATH hierarchy.** The top 100 superfamilies in TED-100 for each CATH class, where more matches to CATH superfamilies have been identified through structural hits in TED compared with sequence hits in Gene3D. (ii) Proportion of domains matched to CATH classes (*n* = 238,569,631). Andy M. Lau *et al.* Exploring structural diversity across the protein universe with The Encyclopedia of Domains. *Science* **386**, eadq4946(2024). DOI:10.1126/science.adq4946. Author Accepted Manuscript (AAM) version available under a CC BY public copyright license." Manuscript published in Science, Volume 386, Issue 6721, 1 Nov 2024

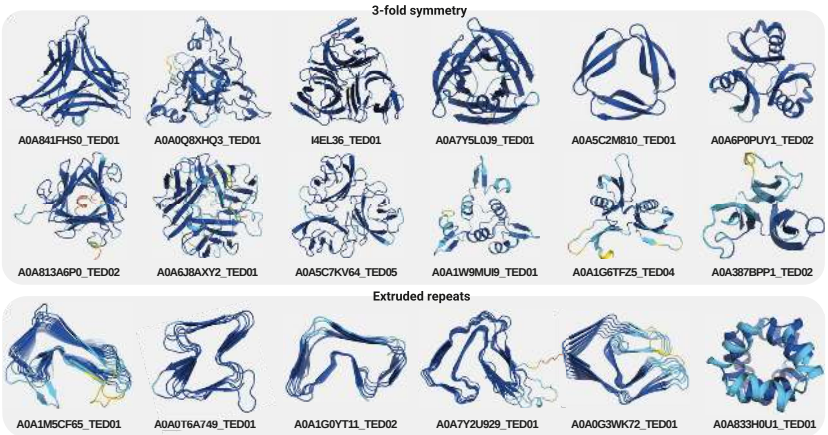
Table **4.7.4.1** below shows examples of novel domain folds and probable functional sites described in the paper by Lau et al.

TED: A0A7M3WA57_TED05 - paired beta-strands in a closed, twisted hairpin with both termini adjacent. (4Ci from Lau et al.)	TED: E1Z635_TED02 - Alpha-helical variant, Eukaryotic. (4Dii from Lau et al).	TED: A0A2J6RQN3 - tentative Zn^{2+}-binding protein	TED: M5FA19 - tentative heme c-binding protein
 Download iCn3D png file	 Download iCn3D png file	 Download iCn3D png file	 Download iCn3D png file

View this in iCn3D as follows:

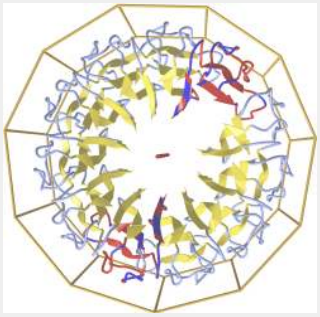
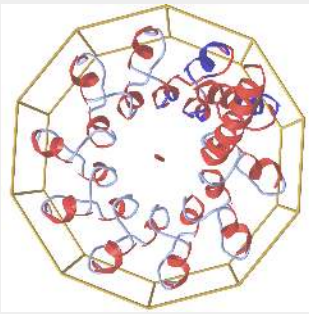
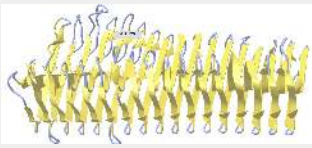
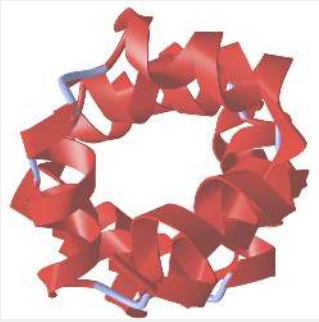
- download the above files to your computer. IMPORTANT: If the file opens as an image in a new browser window, right-click the image and save the file to download it!
- open iCn3D
- **File, Open File, iCn3D appendable,** navigate to the folder with the downloaded png file, and select it.

Figure 4.7.4.25 below shows new examples of symmetry (C3) and extruded repeat domains found using TED.



Examples of high-symmetry domains and extruded repeats. Domains are identified as part of the novel domain identification pipeline, and domains with high internal symmetry are identified through scoring with the SymD program. Extruded repeats are domains with many ordered cyclical repeats projecting along one axis. Coloration follows plDDT confidence bins as per the AFDB. Dark blue indicates very high confidence: plDDT ≥ 90; blue indicates high confidence (90 > plDDT ≥ 70); yellow indicates low confidence (70 > plDDT ≥ 50); and orange indicates very low confidence (plDDT < 50). Andy M. Lau *et al. ibid.*

Table 4.7.4.2 below shows iCn3D examples of higher symmetry and extruded repeat domains described in the paper by Lau et al.

C11 symmetry-A0A1V6M2Y0	C10 symmetry-A0A6C0LIE9	Extruded-A0A1M5CF6	Extruded- A0A833H0U1
 Download iCn3D png file	 Download iCn3D png file	 Download iCn3D png file	 Download iCn3D png file

View this in iCn3D as follows:

- download above files to your computer. IMPORTANT: If the file opens as an image in a new browser window, right-click the image and save the file to download it!
- open iCn3D
- **File, Open File, iCn3D appendable**, navigate to the folder with the downloaded png file, and select it.

Individual protein annotations can also be browsed from the TED website (https://ted.cathdb.info).

TED structural domain assignments for AlphaFold Database v4 and associated codes are available for download at Zenodo. The deposition contains domain assignments for TED, PDB files for novel folds, and individual domain assignments from Chainsaw, Merizo, and UniDoc to facilitate further benchmarking efforts. Specifically:

- **novel_folds_set_models.tar.gz** contains PDB files for all novel-fold representatives identified in TED100.

- **high_symmetry_folds_set_models.tar.gz** contains PDB files of all highly symmetrical fold representatives identified in TED100.

4.7.4.4: Summary

This chapter explores the recurring structural motifs and higher-level organizational units that underpin protein architecture. It begins by examining common super-secondary structures—distinct arrangements of secondary structure elements that recur across different proteins despite vast sequence diversity. Examples include the alpha-loop-alpha motif, exemplified by the helix-turn-helix found in DNA-binding proteins and the EF-hand motif in calcium-binding proteins such as calmodulin, where two alpha helices flank a loop enriched in acidic and polar residues. The chapter also reviews beta-hairpins (or beta-turns), which connect antiparallel beta strands, as well as the Greek key motif, which organizes four adjacent beta strands into a characteristic pattern reminiscent of ancient keys. Another common motif, the beta-alpha-beta structure, links two beta strands via an intervening alpha helix, as seen in enzymes such as triose phosphate isomerase.

The discussion then transitions to larger structural motifs that underpin protein architecture. The Rossmann fold, for instance, is highlighted for its role in binding nucleotide cofactors such as NAD^+, comprising a repeated beta-alpha-beta arrangement that forms a three-layered sandwich structure. Similarly, the TIM barrel, a ubiquitous fold found in many enzyme families, is an alternating arrangement of eight beta strands and eight alpha helices that forms a barrel with an internal active site. Other architectures, such as beta helices and beta propellers—characterized by repeating beta-sheet elements arranged in helical or radial fashions—are also presented, underscoring their prevalence in proteins from pathogens and regulatory factors.

The concept of protein domains is then introduced as the fundamental units of tertiary structure. Domains are defined as independently folding units that often carry distinct functions. Their assembly into multidomain proteins, achieved through evolutionary processes such as gene duplication and recombination, enhances protein function by providing multiple binding sites

and catalytic activities. The chapter also discusses the role of symmetry in protein oligomers, emphasizing how common symmetry operations, particularly proper rotations, govern the assembly of homomeric complexes.

Finally, the chapter highlights advances in bioinformatics and AI-driven tools, such as TED, CATH, and Pfam, that are revolutionizing our ability to predict, classify, and compare protein domains at a proteome-wide scale. Together, these insights into structural motifs and domain architecture provide a framework for understanding how proteins achieve their diverse functions despite limited primary sequence diversity, setting the stage for further exploration of protein structure-function relationships in advanced studies.

This page titled 4.7.4: Secondary Structural Motifs and Domains is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.4: Secondary Structural Motifs and Domains** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure

Learning Goals (ChatGPT o1, 1/26/55)

- **Classify Protein Structural Classes:**
  - Distinguish between the three major classes of proteins based on secondary structure: α proteins, β proteins, and α/β proteins.
  - Identify examples of each class using representative models (e.g., orthogonal and updown bundles for α proteins, roll and sandwich for β proteins, and two- or three-layer sandwiches for α/β proteins).
- **Analyze Specific Structural Motifs:**
  - Describe the characteristics and functional roles of common motifs, such as:
    - The alpha-loop-alpha motif (e.g., helix-turn-helix and EF-hand) in DNA-binding and calcium-binding proteins.
    - The beta-hairpin (or beta-turn) motif found in antiparallel beta structures.
    - The Greek key motif as a recurring pattern in antiparallel beta sheets.
    - The beta-alpha-beta motif as a connector between parallel beta strands.
  - Use interactive 3D models and topology maps to visualize how these motifs are arranged and how they contribute to the overall protein fold.
- **Understand Structural Motifs in Protein Architecture:**
  - Explain how super-secondary structures (motifs) are integrated into larger architectural frameworks (domains) that determine the protein's tertiary structure.
  - Illustrate how the arrangement of motifs such as the Rossmann fold and TIM barrel relates to cofactor binding and enzymatic activity.
  - Discuss how proteins with little secondary structure, like HIV-1 TAT or metallothionein, may undergo conformational changes upon binding their targets.
- **Connect Structure to Function:**
  - Analyze how the distribution of secondary structures within a protein (e.g., extensive α-helices in globular proteins versus β-sheet-dominated proteins) influences its overall shape, stability, and biological function.
  - Evaluate the functional significance of amphipathic helices and beta propellers in processes such as membrane integration and molecular recognition.
- **Apply Bioinformatics Tools to Structural Classification:**
  - Explore databases and tools (e.g., CATH) to understand how proteins are classified into structural classes based on secondary and tertiary structure.
  - Compare structural motifs across proteins with similar functions to infer evolutionary relationships and predict functional sites.

These learning goals are designed to provide a framework for understanding how common structural motifs contribute to protein architecture and ultimately to the diverse functions of proteins in biological systems.

Proteins can also be classified according to the type and extent of secondary structure they contain. A detailed description of protein classes can be found at CATH. Here, we will describe the basic types with a few examples.

4.7.5.1: Alpha proteins

The core of these proteins is composed of alpha helices. The CATH classification shows two major types

a. *Orthogonal bundles*. Example: the Z[beta] Domain of the RNA-editing Enzyme ADAR1 (1xmk), shown in the interactive iCn3D model in Figure 4.7.5.1.

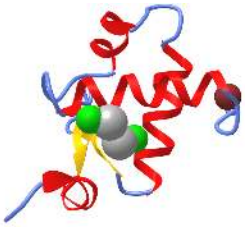


Figure 4.7.5.1: Alpha protein, Orthogonal bundles: Z[beta] Domain of the RNA-editing Enzyme ADAR1 (1xmk) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fxHFM9ErUptjY7

b. *Updown bundles*. Example: Phospholipase A2 from Agkistrodon acutus venom (1mc2), shown in the interactive iCn3D model in Figure 4.7.5.2.

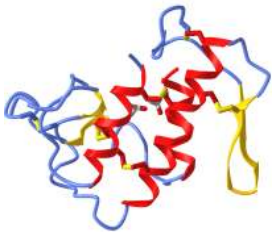


Figure 4.7.5.2: Alpha protein, Updown bundles: Phospholipase A2 from Agkistrodon acutus venom (1mc2) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...NGSeK9JAMe3Qh9

4.7.5.2: Beta proteins

In these proteins, the core is typically an antiparallel beta sheet. There are many types of these, including single sheets, rolls, beta barrels, clams, sandwiches, propellers, etc (some of which we have already discussed). Here are a few interesting examples.

a. *Roll*. Example: The second SH3 domain from ponsin (2O9S), shown in the interactive iCn3D model in Figure 4.7.5.3.

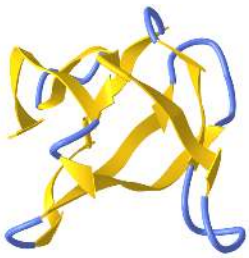


Figure 4.7.5.3: Beta protein Roll: Second SH3 domain from ponsin (2O9S), (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xYZBXHWpt3njf7

b. *Sandwich*. Example: Mcg immunoglobulin light chain variable domain (4unu), shown in the interactive iCn3D model in Figure 4.7.5.4.

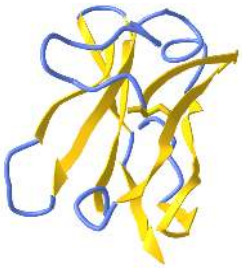


Figure 4.7.5.4: Beta protein Sandwich - Mcg immunoglobulin light chain variable domain (4unu) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...4vcY5svrBveJp7

4.7.5.3: Alpha/Beta proteins

These are the most common class and contain many beta-alpha-beta motifs with mostly parallel beta strands surrounded by alpha helices. Again, these proteins have many variants, including the alpha-beta barrel. We will show two other common ones.

a. *Two-layer sandwich*. Example: HIV-1 Nef-SF2 Core Domain (4U5W), shown in the interactive iCn3D model in Figure 4.7.5.5.

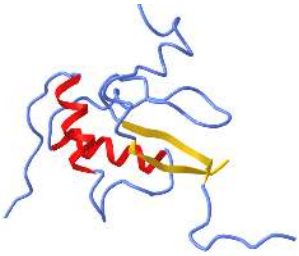


Figure 4.7.5.5: Alpha-Beta *Two layer sandwich* - HIV-1 Nef-SF2 Core Domain (4U5W) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6Fx6J3ecPJbF3A

b. Three-layer *sandwich (aba)*. Example: Human biliverdin IX beta reductase (1hdo), shown in the interactive iCn3D model in Figure 4.7.5.6.

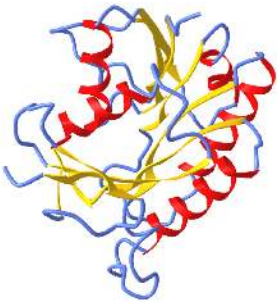


Figure 4.7.5.6: Alpha-Beta Three layer sandwich (aba) - Human biliverdin IX beta reductase (1hdo) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...s8GrdasiecmZC8

4.7.5.4: Little Secondary Structures

These proteins most likely can adopt different conformations on binding to target molecules. Here are two examples

a. HIV-1 TAT (Transactivating) Protein (1JFW), shown in the interactive iCn3D model in Figure 4.7.5.7.

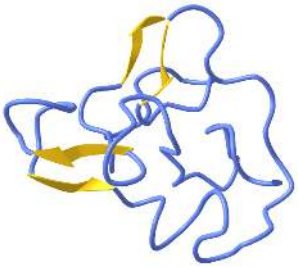


Figure 4.7.5.7: Few Secondary Structures - HIV-1 TAT (Transactivating) Protein (1JFW) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zVYofj3HFasWBA

b. Rat rat metallothionein-2 (4MT2), shown in the interactive iCn3D model in Figure 4.7.5.8.

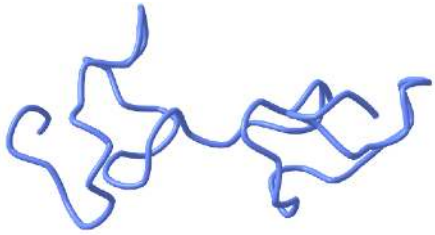


Figure 4.7.5.8: Few Secondary Structures - Rat rat metallothionein-2 (4MT2) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fYY4xF2RWd7pD7

4.7.5.5: Summary

This chapter provides a comprehensive overview of the recurring structural motifs and architectures that form the building blocks of protein structure. Although each protein's 3D structure is unique, many share common substructures that appear repeatedly across different protein families and are often linked to specific functions.

The discussion begins with super-secondary structures, which are groups of secondary structure elements organized in characteristic patterns. Key motifs include:

- **Alpha-Loop-Alpha Motifs:**
  These motifs, exemplified by the helix-turn-helix in DNA-binding proteins (e.g., the basic helix-turn-helix of c-Myc) and the EF-hand found in calcium-binding proteins like calmodulin, involve two alpha helices connected by a loop. In the EF-hand, the loop is enriched in acidic and polar residues that coordinate calcium ions, creating a "hand" that grips the ion.
- **Beta-Hairpins and Greek Key Motifs:**
  The beta-hairpin, a simple turn connecting two antiparallel beta strands, and the more elaborate Greek key motif, which organizes four adjacent beta strands in a distinctive pattern, are common in beta-rich proteins. These motifs provide the structural foundation for larger beta-sheet arrangements, such as rolls, sandwiches, and propellers.
- **Beta-Alpha-Beta Motifs:**
  This motif connects two beta strands with an intervening alpha helix, forming a common structural framework in many enzymes. It plays a crucial role in defining the active sites and the overall fold of proteins such as triose phosphate isomerase.

The chapter then shifts to the higher-order organization of protein structure, discussing how these recurring motifs are integrated into larger structural architectures or domains. Examples include:

- **The Rossmann Fold:**
  A ubiquitous motif that binds nucleotide cofactors (such as NAD^+) through a distinctive beta-alpha-beta arrangement, forming a three-layered sandwich.
- **The TIM Barrel:**
  A common and versatile fold characterized by eight alternating alpha helices and parallel beta strands, which serves as the scaffold for a variety of enzyme families despite low sequence similarity among them.
- **Beta Helices and Beta Propellers:**
  Beta helices form right-handed, parallel helical structures that are often found in proteins from pathogens. In contrast, beta propellers consist of multiple blade-shaped beta sheets arranged radially around a central axis, forming a funnel-shaped binding site.

Finally, the concept of domains is introduced as the fundamental units of tertiary structure that fold independently and often serve as protein functional modules. The chapter discusses how domains arise through evolutionary processes such as duplication, divergence, and recombination, and how modern computational tools (like TED, CATH, and Pfam) are used to classify and predict domain structures across the proteome.

In summary, this chapter illustrates that, despite the complexity of protein structures, a relatively limited set of structural motifs and architectures recur across proteins, providing insight into their functions, evolution, and the principles governing protein folding and assembly.

This page titled 4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.6: Intrinsically Disordered Proteins

Learning Goals (ChatGPT o1, 1/26/25)

- **Define and Characterize Intrinsic Disorder:**
  - Explain what intrinsically disordered proteins (IDPs) and intrinsically disordered regions (IDRs) are, and describe how their amino acid composition (enrichment in polar and charged residues with low hydrophobicity) distinguishes them from ordered proteins.
  - Interpret experimental data (e.g., from X-ray crystallography, NMR, CD spectroscopy, and computational predictions such as AlphaFold confidence scores) that indicate the presence of disorder in proteins.
- **Understand the Functional Role of Protein Disorder:**
  - Discuss how the flexible, dynamic nature of IDPs allows them to bind multiple partners and participate in cell signaling, regulation, and transcription.
  - Describe the concept of coupled folding and binding, and explain how disorder enables a protein to adopt more ordered structures upon ligand binding.
  - Evaluate the idea of the "Protein Trinity," which posits that proteins exist in multiple native states (ordered, collapsed, and extended) that contribute to their biological functions.
- **Examine the Biological and Regulatory Implications of Disorder:**
  - Analyze how intrinsic disorder affects protein turnover, aggregation propensity, and post-translational modifications.
  - Discuss the significance of IDPs in disease contexts, including their roles in neurodegenerative disorders (e.g., tauopathies and alpha-synuclein aggregation) and in the formation of dynamic complexes.
- **Explore Metamorphic Proteins:**
  - Define metamorphic proteins and distinguish them from both fully ordered proteins and prion proteins by discussing their reversible, alternative conformations.
  - Describe how metamorphic proteins, such as lymphotactin (XCL1) and Mad2, can switch conformations in response to subtle changes (e.g., in pH, redox state, or binding interactions) without complete unfolding.
  - Evaluate the evolutionary and functional significance of proteins that can adopt multiple, interconvertible conformations.
- **Apply Computational and Experimental Approaches:**
  - Illustrate how molecular dynamics simulations and AI-based prediction tools (e.g., ALBATROSS, IDRLab) are used to study the conformational ensembles of disordered proteins.
  - Interpret interactive structural models (e.g., using iCn3D) to visualize examples of IDPs (such as SIC1, tau, and alpha-synuclein) and metamorphic proteins, and relate these structures to their functional outcomes.
- **Connect Sequence, Structure, and Function in Disorder:**
  - Explain how synonymous mutations can influence protein folding and structure, particularly for membrane proteins, by affecting the kinetics of translation and cotranslational folding.
  - Discuss how the lack of a stable hydrophobic core in disordered regions underlies their dynamic behavior and functional versatility in the cell.

These learning goals provide a roadmap for understanding the diverse roles of protein disorder and metamorphism in cellular function and regulation, bridging concepts from structural biochemistry, molecular dynamics, and protein evolution.

4.7.6.1: Intrinsically Disordered Proteins (IDPs) and Metamorphic Proteins

Many examples of proteins that are partially or completely disordered yet retain biological function have been identified. At first glance, this might appear unexpected since how could such a protein bind its natural ligand with specificity and selectivity to express its function? Of course, one could postulate that ligand binding would induce conformational changes necessary for

function (such as catalysis), in an extreme example of an induced fit rather than a "lock-and-key" fit. Decades ago, Linus Pauling predicted that antibodies, proteins that recognize foreign molecules (antigens), would bind loosely to the antigen, then undergo a conformational change to form a more complementary, tighter fit. This was the easiest way to allow a finite number of protein antibodies to bind to seemingly endless numbers of foreign molecules. This is indeed one way antibodies can recognize foreign antigens. Antibodies that bind to antigens with high affinity and, hence, high specificity are more likely to bind through a lock-and-key fit. (Pauling, however, didn't know that the genes that encode the protein chains in antibodies are differentially spliced and subjected to enhanced mutational rates, allowing the generation of incredible antibody diversity from a limited set of genes.)

4.7.6.1.1: Intrinsically Disordered Proteins (IDPs)

It's been estimated that over half of all native proteins have regions (greater than 30 amino acids) that are disordered, and upwards of 20% of proteins are completely disordered. Regions of disorder are enriched in polar and charged side chains, since these might be expected to assume many available conformations in aqueous solutions compared to sequences enriched in hydrophobic side chains, which would probably collapse into a compact core stabilized by the hydrophobic effect. Mutations in the disordered regions tend to preserve the disordered region, suggesting that the disordered region is advantageous for "future" function. In addition, mutations that convert a noncoding sequence into a coding one invariably yield disordered protein sequences. Disordered proteins tend to have regulatory properties and bind multiple ligands, whereas ordered proteins are involved in highly specific ligand binding necessary for catalysis and transport. The intracellular concentration of disordered proteins has also been shown to be lower than that of ordered proteins, possibly to prevent inappropriate binding interactions mediated by hydrophobic interactions, for example. Processes to accomplish this include more rapid mRNA and protein degradation and slower translation of mRNA for disordered proteins. For a similar reason, misfolded proteins are also targeted for degradation. Figure **4.7.6.1** shows characteristics of intrinsically disordered proteins.

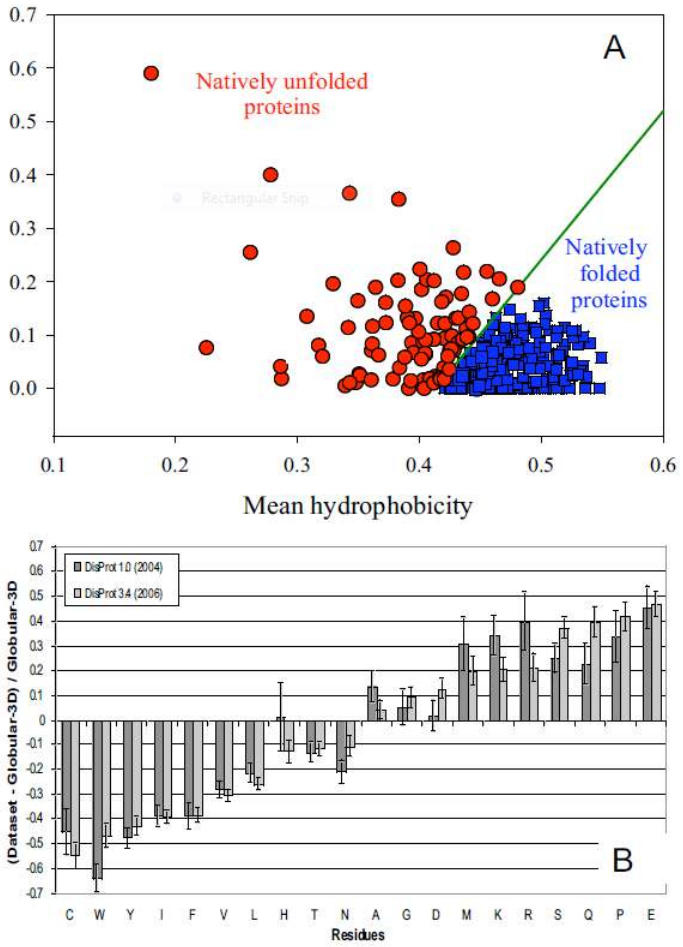


Figure 4.7.6.1: Characteristics of Intrinsically Disordered Proteins. From an open-access journal: Dunker, A. et al. BMC Genomics 2008, 9(Suppl 2):S1 doi:10.1186/1471-2164-9-S2-S1. Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0)

Panel A shows the mean net charge vs the mean hydrophobicity for 275 folded and 91 natively unfolded proteins. Panel B shows the relative amino acid composition of globular (ordered) proteins compared to regions of disorder greater than 10 amino acids in disordered proteins. The two grey bars were obtained using two software versions to analyze the proteins. Again, the graph shows the enrichment of hydrophilic amino acids in disordered proteins.

Many experimental methods can detect disordered regions in proteins. Such regions are not well resolved in X-ray crystal structures (they have high B factors). NMR solution structures would show multiple, distinct conformations. CD spectroscopy likewise would show an ill-defined secondary structure. In addition, solution measurements of size (light scattering, centrifugation) would show more significant size distributions for a given protein.

What types of proteins contain disorder? The above experimental and new computational methods have been developed to classify proteins according to their degree of disorder. There appear to be more IDPs in eukaryotes than in archaea and prokaryotes. Many IDPs are involved in cell signaling processes (when external molecules signal cells to respond by proliferating, differentiating, dying, etc). Most appear to reside in the nucleus. The largest percentage of known IDPs bind to other proteins and DNA. These results suggest that IDPs are essential to protein function and probably confer significant advantages to eukaryotic cells. Multiple functions can be elicited from the interaction of a single IEP (derived from a single gene) with different protein-binding partners. This would greatly extend the effective genome size in humans from around 20,000 protein-encoding genes with specified functions to many more. This doesn't even consider the increase in functionalities derived from post-translational chemical modifications.

Protein structure is fluid and complex, and our simple notions and words to denote proteins as either native or denatured are misguided and constrain our ideas about how protein structure elicits biological function. For example, what does the word "native"

mean if proteins exist in multiple states simultaneously, in vivo and in vitro? Dunker et al. (2001) coined the concept of the "Protein Trinity" to move beyond the notion that a single protein folds to a single state, which elicits a single function. Instead, each of the states in the "trinity", the ordered, collapsed (or molten globule) and extended (random coil), coexist in the cell, as shown in Figure **4.7.6.2**. Characteristics of Intrinsically Disordered Proteins. Hence, all can be considered "native" and all contribute to the cell's function. A single IDP could bind to multiple protein partners, yielding distinct final structures and functions. IDPs would also be more accessible and hence more susceptible to proteolysis, providing a simple mechanism to control their concentrations — an important way to regulate their biological activity. Their propensity to undergo post-translational chemical modification would lead to new biological regulation types.

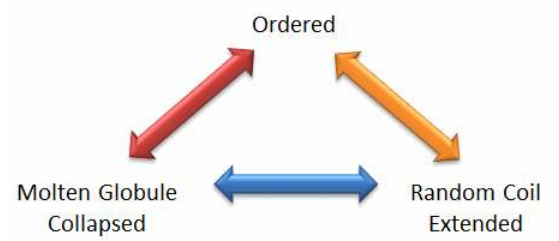


Figure **4.7.6.2**: The Protein Trinity: Ordered, Collapsed and Extended States

These ideas have profound ramifications for our understanding of cellular phenotype expression. In addition, a whole new world of drug targets is available, as drugs that modulate transitions between ordered, collapsed, and extended protein states are being discovered. Likewise, side effects of drugs might be understood by investigating their effects on these transitions in IDPs that were not initially targeted for analysis. Several web databases, including PONDR - Predictor of Naturally Occurring Disorder and Database of Protein Disorder, are available.

IDPs span a spectrum from fully unstructured to partially structured, including random coils, (pre-)molten globules, and large multi-domain proteins connected by flexible linkers. They constitute one of the main types of protein (alongside globular, fibrous, and membrane proteins). Figure **4.7.6.3** shows the conformational flexibility in SUMO-1 protein (PDB:1a5r), which is a composite of 10 NMR structures. The central part shows a relatively ordered structure. Conversely, the N- and C-terminal regions (left and right, respectively) show ‘intrinsic disorder’.

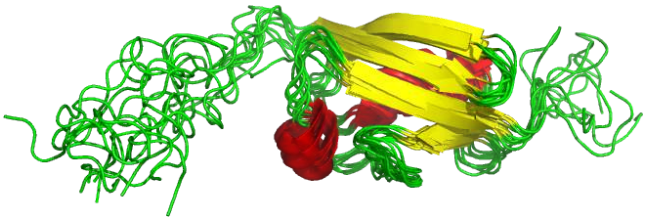


Figure **4.7.6.3**: Conformational flexibility in SUMO-1 protein (1a5r) showing intrinsically disordered regions

History of IDPs

It's interesting to explore the history of our understanding of IDPs. In the 1930s -1950s, the first protein structures were solved by protein crystallography. These early structures suggested that a fixed three-dimensional structure might be generally required to mediate the biological functions of proteins. When stating that proteins have just one uniquely defined configuration, Mirsky and Pauling did not recognize that Fisher's work would have supported their thesis with his 'Lock and Key' model (1894). These publications solidified the central dogma of molecular biology: the sequence determines the structure, which, in turn, determines the function of proteins. In 1950, Karush wrote about 'Configurational Adaptability,' contradicting all the assumptions and research of the 19th century. He was convinced that proteins have more than one configuration at the same energy level and can choose one when binding to other substrates. In the 1960s, Levinthal's paradox suggested that the systematic conformational search of a long polypeptide is unlikely to yield a single folded protein structure on biologically relevant timescales (i.e., seconds to minutes). Curiously, for many (small) proteins or protein domains, in vitro refolding can be rapid *and efficient*. As stated in Anfinsen's Dogma from 1973, the fixed 3D structure of these proteins is

uniquely encoded in their primary structure (the amino acid sequence), is kinetically accessible and stable under a range of (near) physiological conditions, and can, therefore, be considered as the native state of such "ordered" proteins.

During the subsequent decades, however, many large protein regions could not be assigned in X-ray datasets, indicating that they occupy multiple positions, which average out in electron density maps. The lack of fixed, unique positions relative to the crystal lattice suggested that these regions were "disordered". Nuclear magnetic resonance spectroscopy of proteins also demonstrated the presence of large flexible linkers and termini in many solved structural ensembles. It is generally accepted that proteins exist as ensembles of similar structures, with some regions more constrained than others.

Some people differentiate a particular type of IDP called **Intrinsically Unstructured Proteins (IUPs),** which occupy the end of this flexibility spectrum. In contrast, IDPs also include proteins with a strong local structural tendency or flexible multidomain assemblies. These highly dynamic disordered regions of proteins have subsequently been linked to functionally important phenomena such as allosteric regulation and enzyme catalysis.

Many disordered proteins have their binding affinity for receptors regulated by post-translational modifications. Hence, it has been proposed that the flexibility of disordered proteins facilitates the attainment of the conformational requirements for binding their modifying enzymes and receptors. Intrinsic disorder is particularly common in proteins involved in cell signaling, transcription, and chromatin remodeling. Here are some types or characteristics of IDPs.

Flexible linkers

Disordered regions are often found as flexible linkers or loops connecting domains. Linker sequences vary greatly in length but are typically rich in polar uncharged amino acids. Flexible linkers allow the connecting domains to twist and rotate freely, recruiting their binding partners through protein domain dynamics. They also allow their binding partners to induce larger-scale conformational changes by long-range allostery.

Linear motifs

Linear motifs are short, disordered segments of proteins that mediate functional interactions with other proteins or biomolecules (RNA, DNA, sugars, etc.). Many roles of linear motifs are associated with cell regulation, for instance, in controlling cell shape, subcellular localization of individual proteins, and regulated protein turnover. Post-translational modifications, such as phosphorylation, often tune the affinity (not infrequently by several orders of magnitude) of individual linear motifs for specific interactions. Unlike globular proteins, IDPs lack preformed active sites. Nevertheless, in 80% of IDPs (~3 dozen) subjected to detailed structural characterization by NMR, linear motifs termed PreSMos (pre-structured motifs) —transient secondary structural elements primed for target recognition —are identified. In several cases, it has been demonstrated that these transient structures become full and stable secondary structures, e.g., helices, upon target binding. Hence, PreSMos are the putative active sites in IDPs.

Coupled folding and binding

Many unstructured proteins transition to more ordered states upon binding to their targets. Coupled folding and binding may be local, involving only a few interacting residues, or global, involving an entire protein domain. It was recently shown that coupled folding and binding allow the burial of a large surface area, which would be possible only for fully structured proteins if they were much larger. Moreover, certain disordered regions might serve as "molecular switches" that regulate biological functions by adopting ordered conformations upon binding small molecules, nucleic acids, or ions.

Disorder in the bound state (fuzzy complexes)

Intrinsically disordered proteins can retain their conformational freedom even when they bind specifically to other proteins. The structural disorder in the bound state can be static or dynamic. Structural multiplicity is required for function in fuzzy complexes, and manipulating the bound-disordered region alters activity. The conformational ensemble of the complex is modulated via post-translational modifications or protein interactions. The specificity of DNA-binding proteins often depends on the length of fuzzy regions, which vary through alternative splicing. Intrinsically disordered proteins adopt many different structures *in vivo* in response to cellular conditions, forming a structural or conformational ensemble.

Therefore, their structures are strongly function-related. However, only a few proteins are fully disordered in their native state. The disorder is mainly found in ***intrinsically disordered regions (IDRs)*** within an otherwise well-structured protein. Therefore, the term

intrinsically disordered protein (IDP) includes proteins containing IDRs and fully disordered proteins.

The existence and kind of protein disorder are encoded in its amino acid sequence. As described above, IDPs are characterized by a low content of bulky hydrophobic amino acids and a high proportion of polar and charged amino acids, resulting in low hydrophobicity. This property leads to good interactions with water. Furthermore, high net charges promote disorder due to electrostatic repulsion between similarly charged residues. Thus, disordered sequences cannot sufficiently bury a hydrophobic core to fold into stable globular proteins. In some cases, hydrophobic clusters in disordered sequences provide clues to identify regions that undergo coupled folding and binding (see biological roles).

Many disordered proteins lack regular secondary structure. These regions can be considered more flexible than structured loops. While the latter are rigid and contain only one set of Ramachandran angles, IDPs involve multiple sets of angles. The term flexibility is also used for well-structured proteins, but describes a different phenomenon in the context of disordered proteins. Flexibility in structured proteins occurs around an equilibrium state, whereas it doesn't in IDPs. Many disordered proteins also contain low-complexity sequences, i.e., sequences over-representing a few residues. While low-complexity sequences are a strong indication of disorder, the reverse is not necessarily true. That is, not all disordered proteins have low-complexity sequences. Disordered proteins have low predicted secondary structure content.

Recent Updates: October - December, 2024

One way to determine the extent of disorder of a given protein is to examine the certainty in the predicted AlphaFold structure of the protein. Two examples of proteins that are largely disordered in AlphaFold models are shown below. The legend below shows the color corresponding to the confidence level of the predicted structure along the sequence.

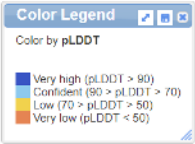
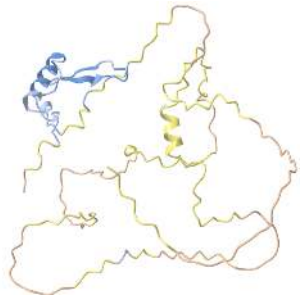
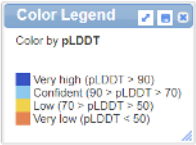


Figure 4.7.6.4 shows interactive iCn3D model of AlphaFold predicted structure of the yeast SIC1 protein (P38634)

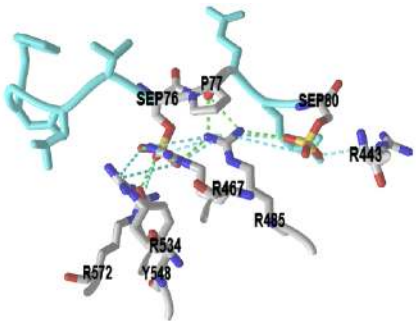




NCBI iCn3D Figure 4.7.6.4: AlphaFold predicted the structure of the yeast SIC1 protein (P38634). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...rJsptsoszZGm29

The SIC1 is a cyclin-dependent kinase (CDK) inhibitor. It interacts with Cdc4, a subunit of ubiquitin ligase, which degrades SIC1. This regulates the G1-to-S cell cycle transition in budding yeast. SIC1 has nine sites for post-translational phosphorylation. Phosphorylation of individual sites enables this IDP to bind the active site of Cdc4. Especially important are phosphorylation of threonines 5, 33, and 45 and serines 69, 76, and 80.

Figure 4.7.6.5 shows interactive iCn3D model of the interactions of a phosphorylated pSIC1 peptide with key residues in Cdc4 from a ScSkp1-ScCdc4-pSic1 peptide complex (3V7D)

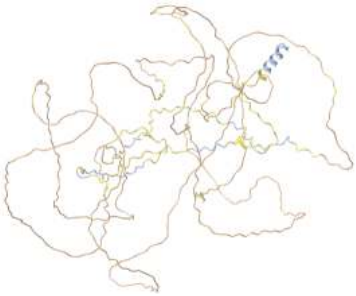


NCBI iCn3D Figure 4.7.6.5: Interactions of a phosphorylated pSIC1 peptide with key residues in Cdc4 from a ScSkp1-ScCdc4-pSic1 peptide complex (3V7D). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PZmqrqwmQyyJJ6

The pSIC1 peptide is shown with a cyan backbone. Ser 76 and Ser 80 are phosphorylated. Note the salt bridges (ion-ion interactions) with arginine side chains in the active site of Cdc4. Figure 4 in this paper shows a model of the interactions between phosphorylated SIC1 and Cdc4. Phosphorylation appears to cause only local, not global, ordering of the SIC1 protein. This local ordering provides sufficient interactions between SIC1 and Cdc4.

The second protein we'll discuss is the human microtubule-associated protein tau.

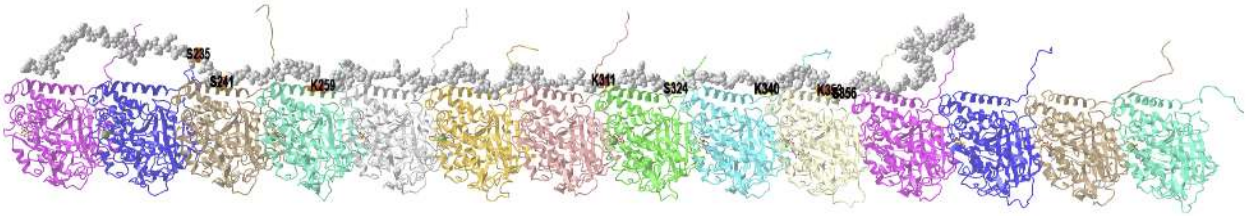
Figure 4.7.6.6 shows interactive iCn3D model of the AlphaFold predicted structure of the human microtubule-associated protein tau (P10636).



NCBI iCn3D Figure 4.7.6.6: AlphaFold predicted structure of the human microtubule-associated protein tau (P10636). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?6gs41N9MEbziWuJe8

This protein is almost entirely disordered. It promotes tubulin association into microtubules (as described in Chapter 5.5: Protein Interactions Modulated by Chemical Energy- Actin, Myosin, and Molecular Motors). It's especially important in neurons for axonal transport.

Figure 4.7.6.7 shows interactive iCn3D model of the cryoEM structure of human tau-tubulin (7PQC).



NCBI iCn3D Figure 4.7.6.7: CryoEM structure of human tau-tubulin (7PQC). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?nFiZfGoWZuSzDdT6A

This is quite an amazing structure! A microtubule consists of multiple copies of a dimer of two monomeric tubulins, alpha and beta. The tubulin monomers are shown in colored cartoons in the model above. Tau is shown in gray as the elongated spacefill

structure. It consists of the tau-F isoform's microtubule-binding region (amino acids 202-395), which has 441 residues. Tau is largely still disordered even in the bound state! There are regions of closer noncovalent interactions with the microtubule. Key interacting residues in Tau are labeled and shown in red spacefill.

Tau has four highly homologous repeats R1–R4. Figure **4.7.6.8** shows electron microscopy (EM) models of Tau/MTs interaction showing two of the repeats, R1 and R2. Many of the interactions are salt bridges.

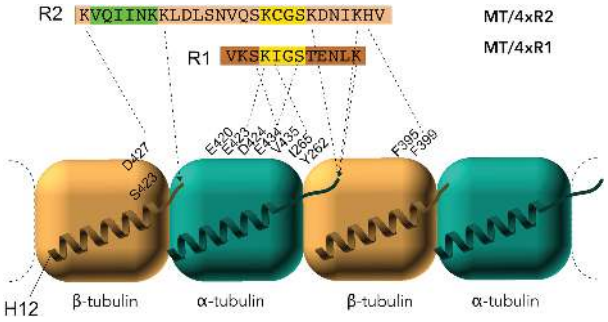


Figure **4.7.6.8**: Electron microscopy (EM) models of Tau/MTs interaction. Barbier P, Zejneli O, Martinho M, Lasorsa A, Belle V, Smet-Nocca C, Tsvetkov PO, Devred F, and Landrieu I (2019). Role of Tau as a Microtubule-Associated Protein: Structural and Functional Aspects. Front. Aging Neurosci. 11:204. doi: 10.3389/fnagi.2019.00204 **Creative Commons Attribution License (CC BY)**.

The α-tubulin subunit is represented as a blue cube, and the β-tubulin subunit as a brown cube. H12 C-terminal helix is schematized; the C-terminal tail prolongs the H12 helix. Models based on cryo-EM coupled to Rosetta modeling: contacts of the R1 repeat with the α-tubulin subunit at the inter-dimer interface: S258 and S262 of R1 make hydrogen bonds with E434; K259 of R1 interacts with an acidic patch formed by E420, E423, and D424; I260 of R1 is in a hydrophobic pocket formed by residues I265, V435 and Y262; K267 of R1 is in contact with the acidic C-terminal tail. Additional contacts for the R2 repeat with the β-tubulin subunit at the intra-dimer interface: K274 of R2 interacts with an acidic patch formed by D427 and S423; K281 of R2 is in contact with the acidic C-terminal tail of the β-tubulin subunit. The PHF6* peptide (highlighted green) is close to this tail and localizes at the intra-dimer interface. Additional contacts for the R2 repeat with the α-tubulin subunit: K294 and K298 are in contact with the acidic C-terminal tail of the α-tubulin subunit. Finally, H299 of R2 is buried in a cleft formed by residues F395 and F399 of the β-tubulin subunit.

Tau also interacts with RNA and DNA in the nucleus and at cell membranes. It self-associates to form aggregates associated with Alzheimer's and other neurodegenerative diseases (collectively called tauopathies). The latter is described in Chapter 4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules. Many post-translational modifications, including phosphorylation, ubiquitination, glycosylation, nitration, acetylation, and limited proteolysis, regulate tau's structure and activity. As above, phosphorylation plays a key role in regulating tau activity. Phosphorylated and acetylated tau have a lower affinity for the microtubules, which dysregulates their binding. The protein in aggregates is hyperphosphorylated.

Repeats 1-4 have both weak and strong interaction motifs. One strong site is the SK(I/C)GS motif. Phosphorylation within this motif is associated with Alzheimer's Disease. Experimental evidence implicates S235, S241, S262, S324, S356, K259, K311, K340 and K353 as key interaction residues. These are shown in Figure 7 above as labeled red spacefill amino acids in tau (gray) and are in close contact with the microtubule. The lysine interactions are modified and weakened by acetylation.

AI and machine learning have been deployed to analyze sequence data to identify intrinsically disordered regions (IDRs) in protein sequences. One such program is ALBATROSS, **A** deep-**l**earning **A**pproach for predic**t**ing the p**r**operties **O**f di**s**ordered protein**s.** It can predict conformational "ensembles" of disordered proteins, as disordered regions would be expected to have multiple conformations. It is available through Google Colaboratory. Another is IDRLab, which allows molecular dynamics simulations of intrinsically disordered proteins through Google Colab.

Figure **4.7.6.9** below shows a simplified molecular dynamics simulation of the intrinsically disordered protein human alpha-synuclein (Uniprot ID P37840). (Molecular dynamics were discussed in Chapter 3.4: Analyses of Protein Structure.) α-Synuclein (140 amino acids, MW 14,460) is expressed in the brain and presynaptic terminals in the central nervous system, but also in more

distal neurons, and is involved in the regulation of neurotransmitter release and in the synaptic vesicles that hold them. Its aggregation is a cause or consequence of Parkinson's Disease and Lewy Body Dementia (see more details in Chapter 4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules).

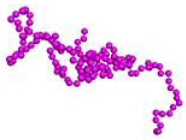
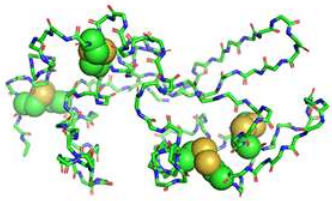


Figure 4.7.6.9: Simplified molecular dynamics simulation of the intrinsically disordered protein human alpha-synuclein (Uniprot ID P37840). G. Tesei, A. I. Trolle, N. Jonsson, J. Betz, F. Pesce, K. E. Johansson, K. Lindorff-Larsen. **Conformational ensembles of the human intrinsically disordered proteome: Bridging chain compaction with function and sequence conservation** *bioRxiv* 2023 2023.05.08.539815 DOI: https://doi.org/10.1101/2023.05.08.539815

This Colab notebook enables running molecular dynamics (MD) simulations of intrinsically disordered proteins (IDPs) and protein regions (IDRs) and to study their conformational ensembles. MD simulations employ the coarse-grained force field CALVADOS2, in which each residue is mapped to a single bead and modeled with a stickiness parameter and electrostatics.

Note the extreme range of conformations of α-Synuclein, allowing this protein to interact with many possible protein targets and also itself when it forms aggregates. Compare it to a one-nanosecond molecular dynamics simulation of two proteins of similar molecular weight shown below in Figure 4.7.6.10. On the left, the hen egg white lysosome (2VB1) has four intramolecular disulfide bonds. The eight cysteines involved in the disulfide bonds are shown in spacefill. On the right is sperm whale myoglobin with the heme removed before the simulation (1A6M). This protein does not have disulfide bonds. Note that the lysozyme is dynamically more constrained by the 4 S-S bonds, making the whole structure slightly more rigid.



2VB1: Hen Egg White Lysozyme (4 S-S bonds)
iCn3D Render

1A6M: Myoglobin (apo, no S-S bonds)

Figure 4.7.6.10: 1-nanosecond simulations were performed on Gromacs using Neuroapp. The input PDB files for the simulations were first processed using Charmm-GUI. The output simulation files (input_protein and md_center.xtc from the output folder) from the Gromacs simulation were opened in Pymol and processed to make the animations shown above (File, Export Movie, as png. then animated gif in APS). Abraham, J. M. et al., GROMACS: High-performance molecular simulations through multi-level parallelism from laptops to supercomputers, https://www.sciencedirect.com/, September 2015, https://doi.org/10.1016/j.softx.2015.06.001. Neurosnap Inc. - Computational Biology Platform for Research. Wilmington, DE, 2022. https://neurosnap.ai/.

? Another interesting IDP!

Click the link below for more information.

The tick saliva glycine-rich protein!

As described above, IDPs are prone to aggregate, with some aggregates being quite amorphous. One example of such an aggregate is the "cement cone" formed by a tick that allows it to stick to its host. About 20% of the aggregate consists of the tick saliva protein glycine-rich protein (GRP). Noncovalent interactions hold together the aggregate/cement. These processes are illustrated in Figure **4.7.6.x**, along with the amino acid characteristics of the protein.

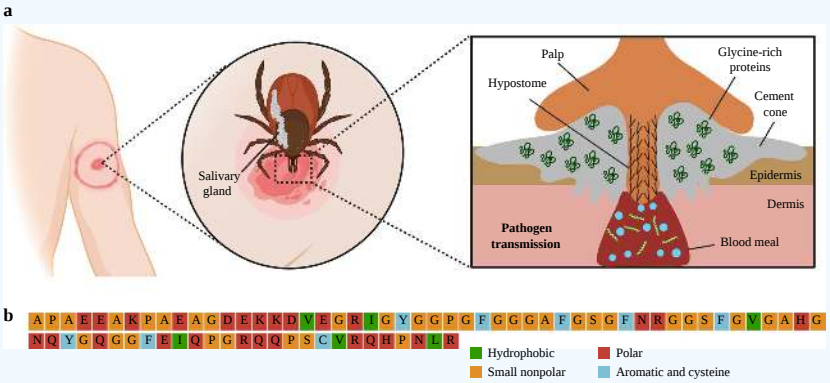
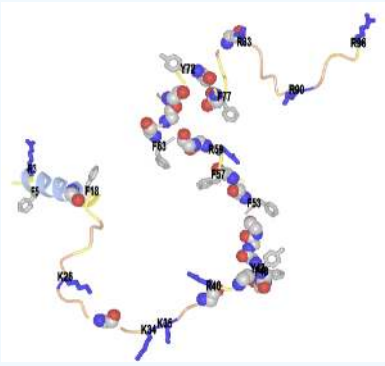


Figure **4.7.6.x**: Tick Glycine Rich Protein Secretion. Ganar, K.A., Nandy, M., Turbina, P. *et al*. Phase separation and aging of glycine-rich protein from tick adhesive. *Nat. Chem.* (2024). https://doi.org/10.1038/s41557-024-01686-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel a: Schematic overview showing the consequences of a tick bite. The tick inserts its hypostome into the host epidermis and secretes a protein-rich saliva abundant in GRPs. The saliva undergoes a liquid-to-solid transition, forming a hard cement cone that allows the tick to feed on the host for several days and facilitates pathogen transmission (the shown 'bull's-eye' rash is typical in the case of *Borrelia* infection, causing Lyme borreliosis).

Panel b: Amino-acid composition of tick-GRP77 shows a high proportion of non-polar amino acids (44%, of which 26% are glycine). Predicts the N-terminal signal peptide of GRP as an α-helix, while the rest of the sequence (tick-GRP77) remains unstructured, indicating a disordered region.

Figure **4.7.6.x** shows interactive iCn3D model of the AlphaFold structure of Tick saliva Glycine Rich Protein (Q4PME3). Glycines are shown in spacefill. Basic side chains (Arg, Lys) are shown in blue sticks, and the aromatic (Tyr and Phe) are shown in gray sticks.



NCBI iCn3D Figure **4.7.6.x**: AlphaFold structure of Tick saliva Glycine Rich Protein (Q4PME3). (Copyright; author via source). Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...656J4hCuHDDup9.

The first 19 amino acids, which form a predicted alpha-helix, form the signal peptide cleaved from the pro-protein to form the mature protein. The remaining portion of the protein (77 amino acids) is called tick-GRP77. The predicted AlphaFold model for GRP77 is completely disordered.

Tick saliva, on evaporation, becomes more concentrated and undergoes a liquid-liquid phase separation to form condensates. Results from lab studies of GRP77 are shown in Figure **4.7.6.x** below. Panel d shows the condensates. Panels a and b show key amino acid positions in the N-terminus (of the mature protein, the C-terminus, and mutants of the C-terminus in which the amino acids phenylalanine and tyrosine (ΔFY) or arginine (ΔR) have changed to alanine to probe the importance of F, Y, and R for aggregation.

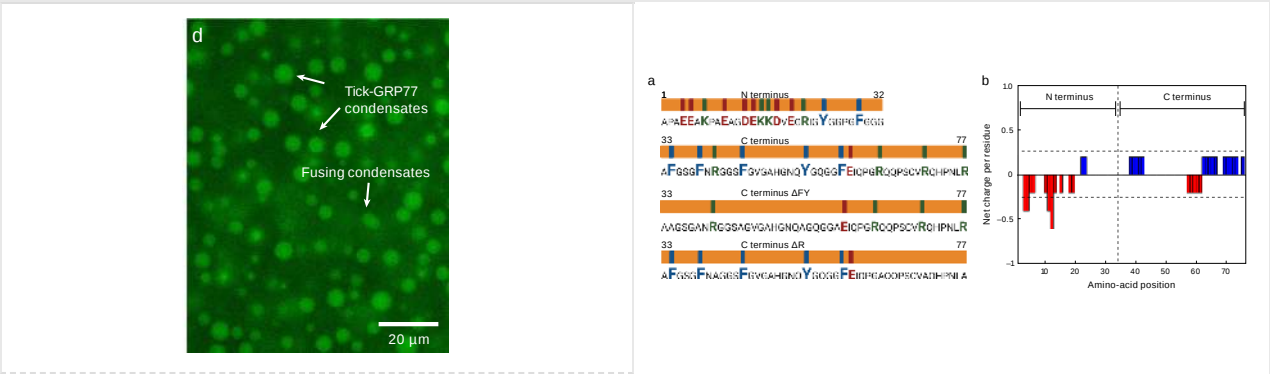


Figure **4.7.6.x**: Characteristics of peptides used in condensation studies. Ganar, K.A. et al. ibid

Panel d: Evaporation of tick-GRP77 on a surface-passivated glass slide showing numerous spherical tick-GRP77 condensates.

Panel a: The four fractions of tick-GRP77: N terminus (1–32), C terminus (33–77), C terminus mutant without aromatic amino acids (ΔFY mutant), and C terminus mutant without arginine residues (ΔR mutant). They all have a similar glycine content but differ in the number of basic (green), acidic (red), and aromatic (blue) amino acid residues. **Panel b**: Net charge per residue as a function of amino-acid position for the two termini, obtained using CIDER. The N terminus is strongly negatively charged, and the C terminus is moderately positively charged.

In vitro condensation experiments were conducted using droplet evaporation assays for the N-terminal and C-terminal peptides and respective mutants. It was found that the N-terminus formed underwent phase separation and condensation on a similar time scale as GRP77. The C-terminal separated quickly (9 min) when placed on specially treated surfaces. Since they differ in charge, it's unlikely that simple salt bridges mediate condensate formation. More likely, cation:pi (between Arg/Lys and the aromatic electron clouds of Tyr and Phe) and pi:pi interactions are key in aggregate formation and stabilization. Indeed, mutations that replaced aromatic side chains with the nonpolar amino acid alanine led to weak condensate formation. Likewise, replacing Arg with Ala reduced condensate formation.

4.7.6.2: Silent Single-nucleotide polymorphisms (SNPs)

For some amino acids, multiple triplet nucleotide sequences (codons) in the gene coding regions for a protein lead to the incorporation of the same amino acid in the protein sequence. Hence, two proteins identical in amino acid sequence might have slightly different nucleotide sequences in the encoded gene. Such single-nucleotide polymorphisms (SNPs) in coding regions were thought not to affect the tertiary structure and biological function of a protein if the single nucleotide variation did not lead to the insertion of a different amino acid into the growing peptide chain (i.e, the codons were synonymous and the mutations presumably silent with no effect). Recently, single-nucleotide polymorphisms (SNPs) in the MDR1 (multidrug resistance 1) gene, which encodes P-glycoprotein, were shown to result in a protein with altered substrate specificity and inhibitor interactions, and hence a different 3D structure. One possible explanation for this observation is a difference in the translation rate of the mRNA encoding this membrane protein. Different rates might yield different intra- and intermolecular associations, resulting in distinct final 3D

structures as the protein cotranslationally folds and inserts into the membrane. This would be especially true if two possible structures were close enough in free energy but separated by a significant activation energy barrier, precluding simple conformational rearrangement between them.

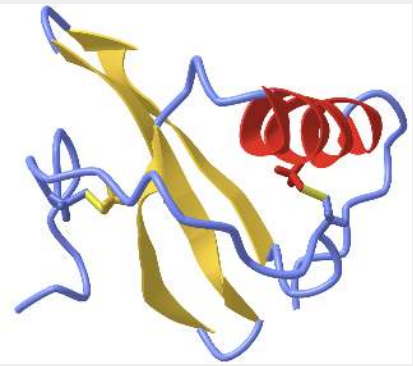
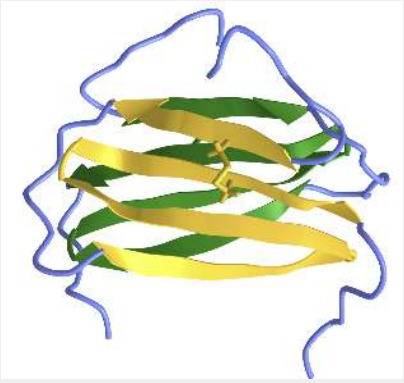
It has been shown in yeast that synonymous mutations (those that don't change the amino acid on the mutation of the DNA encoding the particular amino acid) generally have the same effect on the "health" of yeast as do non-synonymous mutations (those that change the amino acid). This rather startling result upends much dogma. Some possible effects of synonymous mutation include alterations in gene expression of the mutated gene and effects on the stability of the transcribed RNA. Both types of mutations reduce mRNA levels and the fitness of the yeast, as defined by growth rate.

4.7.6.3: Metamorphic Proteins

In addition to prion proteins, many proteins can adopt multiple conformations under the same conditions. In contrast to prion proteins, for which the formation of the beta-structure variant is irreversible since the conformational change is associated with aggregation, many proteins can change conformations reversibly. Often, these changes are not associated solely with binding interactions that trigger them. Murzin has described proteins that change conformation upon changes in pH (viral glycoproteins), redox state (chloride channel), disulfide isomerization (lysozyme), and bound ligand (RNA polymerase as it initiates and then elongates the growing RNA polymer). He cites two proteins that appear to change state without external signals. These include Mad2, in which the two conformers share an extensive similarity, and Ltn10 (lymphotactin), in which they don't. One form of lymphotactin (Ltn 10) binds to similar lymphokine receptors, while the other (Ltn 40) binds to heparin. Folding kinetics may play a part in these examples as well, as proteins capable of folding to two conformers independently and quickly might prevent misfolding and aggregation that might occur if they had to unfold completely first before a conformational transition. Both Mad2 and Ltn10 alter conformation through transient dimer formation, facilitating conformational changes without widespread unfolding. Mutations in Ltn10 can cause the protein to adopt the Ltn40 conformation. Hence, primordial "metamorphic" proteins could, by simple mutation, produce new protein functionalities.

Metamorphic proteins, which undergo large structural changes, often involving shifts in hydrogen bonding and hence secondary structure, differ from simpler allosteric proteins with smaller conformational changes. Few metamorphic proteins have been found, but some speculate they could account for as much as 5% of proteins. An excellent example of such a protein is the human chemokine XCL1 (lymphotactin), an immune regulatory protein. It undergoes a major transition from a typical chemokine fold to a dimer with extensive beta structure. Lymphotactin lacks one of the two disulfide bonds found in other chemokines, allowing greater conformational flexibility.

Figure **4.7.6.11** shows interactive iCn3D models of the solution structures of monomeric XCL1 (lymphotactin), PDB 2HDM (left), and a dimeric form, PDB 2JP1 (right).

Monomeric XCL1 (lymphotactin PDB **2HDM**)	Dimeric XCL1 (lymphotactin PDB 2JP1)
 NCBI iCn3D Monomeric XCL1 (lymphotactin) PDB 2HDM. (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...b8BwBXkuTAJvx6	 NCBI iCn3D Dimeric XCL1 (lymphotactin) PDB 2JP1. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1fuU5jg76RAbVA

4.7.6.4: Summary

This chapter examines the dynamic world of intrinsically disordered proteins (IDPs) and metamorphic proteins, challenging the traditional view that proteins must adopt a single, well-defined structure to be functional. It begins by defining IDPs and intrinsically disordered regions (IDRs) as protein segments that, due to their enrichment in polar and charged residues and low hydrophobicity, fail to form a stable hydrophobic core. Despite this apparent lack of structure, many IDPs perform crucial regulatory roles, particularly in cell signaling and transcription, often through mechanisms such as coupled folding and binding. In this process, an otherwise unstructured protein adopts a more ordered conformation upon binding its partner, a concept that supports the idea of a dynamic "Protein Trinity," in which ordered, collapsed, and extended states coexist and contribute to biological function.

The chapter further discusses the biological implications of disorder, highlighting how IDPs are predisposed to post-translational modifications and are subject to rapid degradation, which together help regulate their cellular concentrations and interactions. Detailed examples, including proteins like SIC1 and tau, illustrate how intrinsic disorder facilitates multiple binding interactions and functional versatility, while also predisposing some proteins to form pathological aggregates in diseases such as Alzheimer's and Parkinson's.

In addition to IDPs, the chapter introduces metamorphic proteins—those capable of adopting two or more distinct conformations under the same environmental conditions. Unlike prion proteins, the conformational transitions in metamorphic proteins are reversible and are not solely triggered by external ligands. Proteins such as lymphotactin and Mad2 serve as prime examples, demonstrating how alternative conformations can yield distinct functional outcomes.

Finally, the text highlights advances in experimental and computational techniques, including molecular dynamics simulations and AI-based prediction tools, which have expanded our understanding of the conformational ensembles of disordered proteins and the factors governing metamorphic transitions. This emerging perspective underscores the fluidity of protein structure. It challenges the conventional "one sequence – one structure – one function" paradigm, opening new avenues for drug targeting and functional regulation in complex biological systems.

4.7.6.5: References

f17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

This page titled 4.7.6: Intrinsically Disordered Proteins is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.6: Intrinsically Disordered Proteins** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.7: Fibrillar Proteins

Learning Goals (ChaptGPT o1, 1/26/25)

After working through this section, **students should be able to:**

1. **Define fibrillar (fibrous) proteins** and distinguish them from globular and membrane proteins based on structure, solubility, and biological function.
2. **Describe structural motifs common in fibrillar proteins** (e.g., α-keratin coiled coils, collagen triple helices) and explain how repetitive sequences give rise to elongated architectures.
3. **Explain the molecular basis of β-sheet formation in fibrillar proteins**, including hydrogen bonding patterns and side-chain packing, and relate these to mechanical properties like strength and elasticity.
4. **Compare and contrast classic fibrillar proteins** (e.g., keratin, collagen) with **spider silk proteins**, emphasizing the role of sequence composition and hierarchical assembly.
5. **Describe how spider silk proteins (spidroins) transition from soluble precursors to solid fibers**, including the roles of:
   - intrinsically disordered protein (IDP) domains,
   - multivalent side-chain interactions (e.g., Arg–Tyr cation–π contacts),
   - liquid–liquid phase separation (LLPS),
   - mechanical alignment and dehydration in the spinning duct.
6. **Summarize the distinct roles of phase separation vs. β-sheet crystallization** in silk fiber formation, and explain why silk represents a unique case among fibrillar proteins.
7. **Compare fibril formation mechanisms** in functional materials (spider silk) with pathological β-sheet assemblies (e.g., amyloid fibrils), identifying fundamental differences and shared motifs.

4.7.7.1: Introduction

Most proteins have a roughly spherical or "globular" tertiary structure. However, many proteins form elongated fibrils with properties such as elasticity: the protein deforms when a force is applied and returns to its original state. Elastic molecules must store energy (go to a higher energy state) when the elongating force is applied, and the energy must be released on return to the equilibrium resting structure. Structures that can store and release energy when subjected to a force have resiliency. Proteins that stretch with an applied force include elastin (in blood vessels, lungs, and skin where elasticity is required), resilin in insects (which stretches on wing beating), silk (found in spider web and whose partial structure we showed in 4.2), and fibrillin (found in most connective tissues and cartilage). Some proteins have high resiliency (90% in elastin and resilin), while others are only partially resilient (35% in silk, which has a tensile strength approaching that of stainless steel).

In contrast to rubber, which has an amorphous structure that imparts elasticity, these proteins, although they have dissimilar amino acid sequences, appear to share a common structure inferred from their DNA sequences. In some (such as fibrillin), the protein has a folded beta-sheet domain that unfolds like a stretched accordion. Others, like elastin and spider silk, have a beta-sheet domain and other secondary structures (alpha-helices and beta turns) along with Pro and Ala repetitions. Scientists are studying these structures to help synthesize new elastic and resilient products.

Elongated structures characterize fibrous proteins. These proteins often aggregate into filaments or bundles, forming structural scaffolds in biological systems. Within animals, the two most abundant fibrous protein families are collagen and α-keratin. Let's start our exploration of fibrillar proteins with these.

4.7.7.2: Collagen

Collagen is the most abundant protein in mammals, making 25% to 35% of the whole-body protein content. It is predominantly found in the extracellular space of various connective tissues. Collagen has a unique quaternary structure: three protein strands

wound together to form a triple helix. It is mostly found in fibrous tissues such as tendons, ligaments, and skin.

Depending upon the degree of mineralization, collagen tissues may be rigid (bone), compliant (tendon), or have a gradient from rigid to compliant (cartilage). It is also abundant in corneas, blood vessels, the gut, intervertebral discs, and dentin. In muscle tissue, it serves as a major component of the endomysium. Collagen constitutes one to two percent of muscle tissue and accounts for 6% of the weight of strong, tendinous muscles. The fibroblast is the most common cell that creates collagen. Gelatin, used in food and industry, is collagen that has been irreversibly hydrolyzed. In addition, partially and fully hydrolyzed collagen powders are used as dietary supplements. Collagen also participates in many binding interactions with target proteins, in addition to its structural role.

The name *collagen* comes from the Greek (*kólla*), meaning "glue," and the suffix *-gen*, denoting "producing." This refers to the compound's early use in boiling the skin and tendons of horses and other animals to make glue.

Over 90% of the collagen in the human body is type I. However, as of 2011, 28 types of collagen had been identified and described, divided into several groups based on the structures they form. The five most common types are:

- Type I: skin, tendon, vasculature, organs, bone (the main component of the organic part of bone)
- Type II: cartilage (the main collagenous component of cartilage)
- Type III: reticulate (the main component of reticular fibers), commonly found alongside type I
- Type IV: forms basal lamina, the epithelium-secreted layer of the basement membrane
- Type V: cell surfaces, hair, and placenta

Let's focus on Type I collagen, which has an unusual amino acid composition and sequence:

- Glycine is found in almost every third residue.
- Proline makes up about 17% of collagen.
- Collagen contains many hydroxyproline and hydroxylysine, which are formed by post-translational modifications catalyzed by different enzymes, both of which require vitamin C as a cofactor.
- Some hydroxylysines are glycosylated, mostly with disaccharides.

Figure **4.7.7.1** shows the post-translational hydroxylations of lysine and proline.

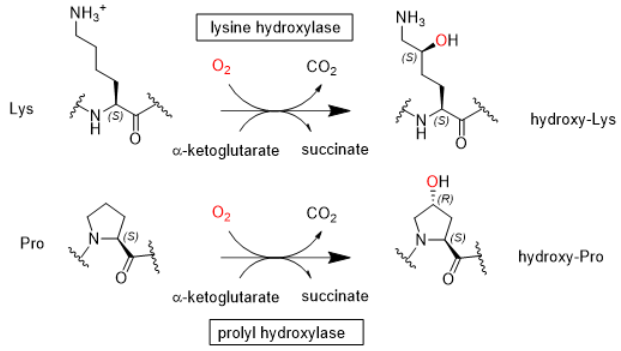


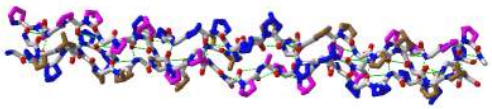
Figure **4.7.7.1**: Post-translational hydroxylations of lysine and proline

Vitamin C deficiency causes scurvy, a serious and painful disease in which defective collagen prevents the formation of strong connective tissue. Gums deteriorate and bleed, with loss of teeth; skin discolors, and wounds do not heal. Before the 18th century, this condition was notorious among long-duration military expeditions, particularly naval ones. Participants were deprived of foods containing vitamin C. Many bacteria and viruses secrete virulence factors, such as the enzyme collagenase, which destroys collagen or interferes with its production.

Collagen has many $(GXY)_n$ repeats, where G is glycine (Gly), and X and Y are frequently proline (Pro) and hydroxyproline (Hyp). Three strands of collagen self-associate to form a triple-stranded helix with 10 GXY triplets in 3 complete turns of the helix. Others suggest seven triplet units in 2 turns of the stands. Note that the helix of each strand in the triple helix is **not** an alpha helix and has different phi/psi angles. Each strand is "frameshifted" by one amino acid, resulting in a staggered arrangement of the individual strands and helices. The glycines are buried along the central axis, so no essential hydrophobic core exists. The X and Y amino

acids are solvent-exposed. All the other side chains, both hydrophobic and hydrophilic, are likewise exposed to solvent. Hydrogen bonding occurs between the amide hydrogen of the peptide bond of Gly and the carbonyl O of an X amino acid in another chain.

Figure 4.7.7.2 shows an interactive iCn3D model of a **triple helical collagen-like peptide** (4Z1R). The main chain atoms, shown in CPK colors, form hydrogen bonds with neighboring chains. The side chains are colored based on the three chains (blue, brown, and magenta). Two sets of Pro-HPro-Gly repeats are labeled.



NCBI iCn3D Figure 4.7.7.2: Triple helical collagen-like peptide (4Z1R). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?TCrji1wPhekypJtc6

Recent Updates: September 16, 2024

Given its widespread nature and its role in the structural stability of organisms, collagen is a very stable biomolecule. It is so stable that traces of it have even been found in the fossil remains of dinosaurs like T. rex from 80 million years ago and a Lufengosaurus from 195 million years ago. The peptide bond is thermodynamically unstable in an aqueous solution, so collagen, like other proteins, would be expected to hydrolyze over a period of hundreds of years. Its unique repetitive sequence, aggregation into a triple helix, and biological environment (such as in bone) help protect it from degradation, but not for millions of years. There must be another previously overlooked structural feature that inhibits its hydrolysis at the electrophilic carbonyl of its peptide bond.

To investigate this extreme stability, the resistance of small nitrophenyl esters of N-acylated prolines in both cis and trans conformations to hydrolysis was studied in an aqueous solution using NMR. Figure 4.7.7.3 below shows the structure of these cis and trans isomers.

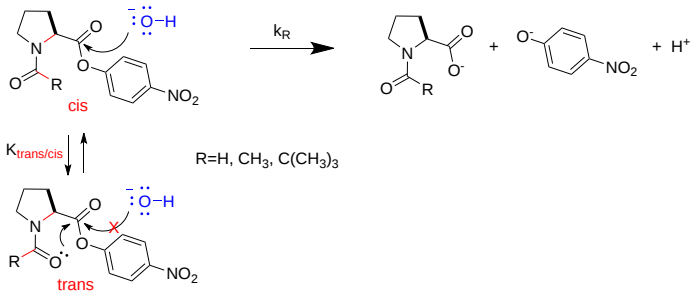


Figure 4.7.7.3: Hydrolysis of N-acylated-Pro-p-nitrophenol esters

The equilibrium constant ($K_{cis/trans}$) for the cis ↔ trans isomers increases with R group size. Remember, for X-Y peptide bonds in dipeptides, the Kcis/trans is about 1000 (trans favored), but for X-Pro, it's around 4. However, in collagen, all X-Pro bonds are trans to enable triple helix formation.

The kinetics of spontaneous hydrolysis in solution (Tris-HCl buffer, pH 8, 0.1 M NaCl, and 1% DMSO to increase solubility) were studied by measuring $A_{400\ nm}$ on the release of the nitrophenolate anion. There is a remarkable resistance to hydrolysis in the trans isomer that can best be accounted for by an interaction between a nonbonded (n) electron pair on the acyl group (corresponding to the X C=O in an XY dipeptide sequence) and the C=O of the ester. This likely results from some stabilizing electron donation from the n orbital of the O to the pi* antibonding C=O orbital. In hydrolysis, water or OH^- would donate a lone pair. This becomes disallowed in the sterically restricted trans form since only two electrons can occupy an orbital (in this case, the pi* antibonding orbital). The authors describe this as an example of the Pauli Exclusion principle (only 2 electrons per orbital), which you encountered in introductory chemistry classes.

Figure 4.7.7.4 below shows representations of the interactions and an energy diagram from Chapter 2.4 (Solubility in an aqueous world - noncovalent interactions in depth) to illustrate these points.

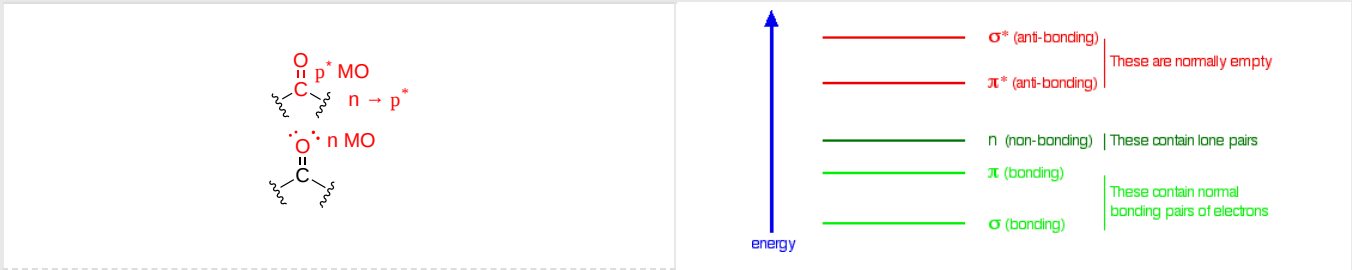


Figure 4.7.7.4: Left, Interactions between a lone pair of electrons (n) on an adjacent group and the pi* MO of an adjacent C=O. Right, Relative energy levels of molecular orbitals (not to scale). https://chem.libretexts.org/Bookshel...c_Spectroscopy

In the triple collagen helix, the n→ pi* interactions are on the same strand, while stabilizing hydrogen bonds are between donors and acceptors on adjacent strands, as illustrated in Figure 4.7.7.5 below.

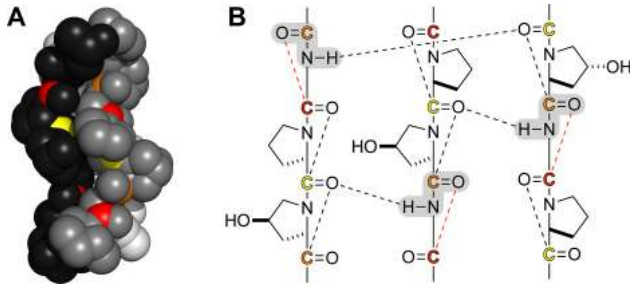


Figure 4.7.7.5: Interactions in a collagen triple helix.

(A) Crystal structure of a (Pro-Hyp-Gly)n triple helix. The three strands are white, gray, or black. In each strand, peptide-bond carbons are yellow (Pro-Hyp), orange (Hyp-Gly), or red (Gly-Pro). The image was generated with PyMOL using PDB entry 1v7h.59.

(B) n→π* Interactions and hydrogen bonds within the main chain of a collagen triple helix. Carbons are colored as in panel A. Interstrand hydrogen bonds polarize the shaded peptide bonds. Jinyi Yang, Volga Kojasoy, Gerard J. Porter, and Ronald T. Raines. *ACS Central Science* **Article ASAP.** DOI: 10.1021/acscentsci.4c00971. Sept 2024. DOI: 10.1021/acscentsci.4c00971. CC-BY 4.0

4.7.7.3: α-Keratin

α-keratin is the key structural element of hair, nails, horns, claws, hooves, and the outer layer of skin. Due to its tightly wound structure, it can function as one of the strongest biological materials and has various uses in mammals, from predatory claws to hair for warmth.

Hanukoglu and Fuchs determined the first α-keratin sequences. These sequences revealed two distinct but homologous keratin families: Type I and Type II keratins. There are 54 keratin genes in humans, 28 of which code for type I and 26 for type II. Type I proteins are enriched in Asp and Glu amino acids, while type II proteins contain more basic amino acids, such as lysine. This differentiation is especially important in α-keratins because, in forming a keratin dimer, the ***coiled-coil requires that one protein coil be type I and the other*** type II. Some pairs of types I and II are particularly complementary within each organism. For example, in human skin, K5, a type II α-keratin, pairs primarily with K14, a type I α-keratin, to form the α-keratin complex of the epidermis layer of cells in the skin.

Coiled-coil dimers then assemble into a tetramer of two staggered coiled-coil dimers. Two tetramers can then pack together to form an elongated protofilament, a very stable, left-handed superhelical structure, as shown in the figure below. The keratin filaments stay associated through hydrophobic interactions between apolar residues along the keratin's helical segments. This is illustrated in Figure **4.7.7.6**.

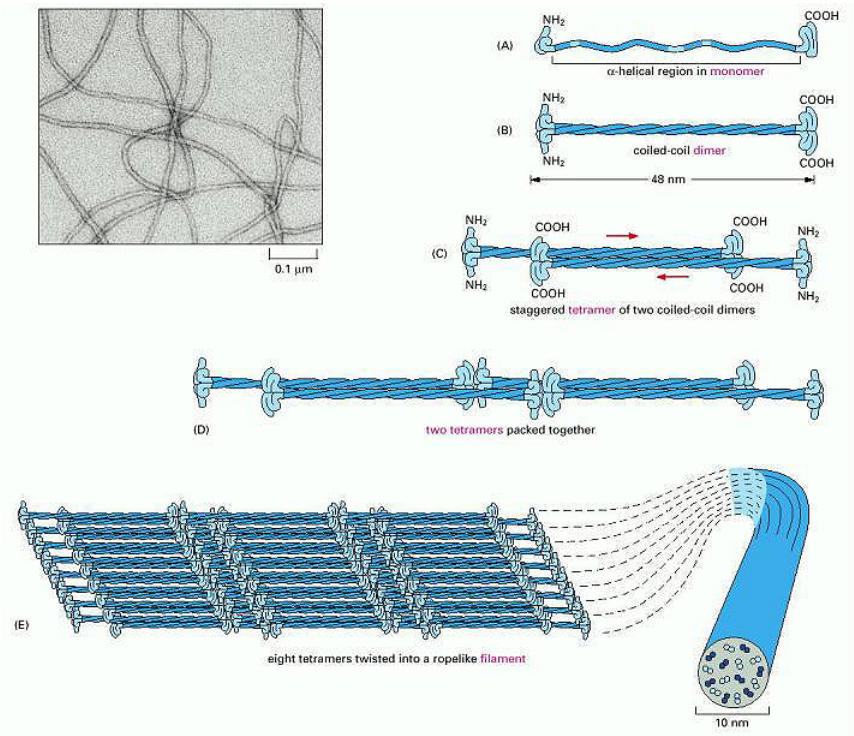
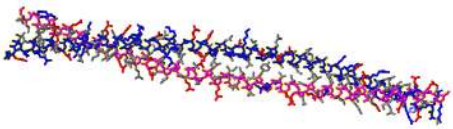


Figure **4.7.7.6**: Assembly of Keratin Fibers. Wiki lectures. https://www.wikilectures.eu/w/Indivi...e_and_function

Initially, two keratin monomers (A) form a coiled-coil dimer structure (B). Two coiled-coil dimers join to form a staggered tetramer (C); next, the tetramers start to join together (D), ultimately forming a sheet of eight tetramers (E). The sheet of eight tetramers is then twisted into a left-handed helix, forming the final intermediate filament (E). An electron micrograph of the intermediate filament is shown in the upper left-hand corner.

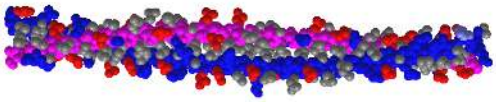
Figure **4.7.7.7** shows an interactive iCn3D model of a dimer of Type I alpha-keratin (magenta backbone) and Type II (blue backbone) (6JFV).



NCBI iCn3D Figure **4.7.7.7**: Dimer of a Type I and Type II alpha-keratin backbones (6JFV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...g3L3sQh9dj5pw8

Acidic (red) side chains (Asp and Glu) and basic (blue) side chains (Lys) can be seen projecting away from the dimer. Both A and B chains have negative and positive side chains. This dimer's A (more acidic) chain has 5 Lys, 7 Arg, 1 Asp, and 16 Glu side chains, for a net charge of +12 -17 = -5. The dimer's B (more basic) chain has 8 Lys, 9 Arg, 5 Asp, and 12 Glu side chains for a net charge of +17 -17 = 0. It is more basic, with 17 lysine and arginine side chains, than the A chain, with 12. Depending on their 3D orientation, they could present a positive face to the more negative monomer in the dimer.

Figure **4.7.7.8** shows an interactive iCn3D model of a spacefill model of the dimer. Note that a significant fraction of the nonpolar side chains are pointed inward between the two monomers and are much less exposed to solvent.



NCBI iCn3D Figure 4.7.7.8: Dimer of a Type I and Type II alpha-keratin backbones in spacefill (6JFV). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...uyJrw28mZ6x5h7

4.7.7.4: Elastin

As its name implies, the protein confers elasticity in target structures such as connective tissue and blood vessels. It has low-complexity hydrophobic domains, and the protein is cross-linked to form larger structures. It contains repeating hydrophobic amino acid sequences, mostly valine, proline, glycine, and alanine, and mimetics of the repeating structure (LGGVG)6 have been studied. This protein also displays significant disorder.

4.7.7.5: Resilin

The following description of resilin is taken directly from an article under Creative Commons Attribution 4.0 International License available at http://creativecommons.org/licenses/by/4.0/. Balu, R., Dutta, N.K., Dutta, A.K. *et al*. Resilin-mimetics as a smart biomaterial platform for biomedical applications. *Nat Commun* **12,** 149 (2021). https://doi.org/10.1038/s41467-020-20375-x

Resilin-Mimetics

Native elastomeric proteins are biomaterials that natural selection has perfected over billions of years to act as molecular springs in a wide range of biological systems, driving unique functions. Among native proteins, resilin is considered one of the most efficient elastic proteins. It is a structural protein mainly found in insect exoskeletons and exhibits outstanding resilience and fatigue life. The first description of resilin was made in the 1960s as a rubber-like protein observed in the locust wing hinge and the dragonfly tendon. Early studies on the composition and structure of resilin revealed that the protein contains about 66% hydrophobic residues (much lower than in elastin), with about 45% proline and glycine combined. In their native state, resilins exist as di- and trityrosine crosslinked hydrogels and exhibit highly amorphous structures as observed by X-ray diffraction and electron microscopy. During biosynthesis, pro-resilins (uncrosslinked) are secreted from the apical surface of the epidermal cells into the subcuticular space, where an enzyme-mediated process crosslinks them to form hydrogels. Over the next three decades, resilin was also identified in many other insects and arthropods, including copepods, reduviidae, and moths. In arthropods, resilin is largely involved in several different functions, including the flexibility and deformability of membranous cuticle and joint systems, the storage of elastic energy in locomotion (jumping, flying, etc.) and catapulting systems, the adaptability to surface topography by multiple contact attachment, and prey catching systems and the reduction of fatigue and damage in feeding and traumatic reproductive system.

The amino acid sequence of resilin was first identified in the early 2000s from the CG15920 gene segment of the fruit fly *Drosophila melanogaster. This* opened up new opportunities for the synthesis and development of biomimetic resilins. The CG15920 gene comprises N-terminal (exon-1), C-terminal (exon-3), and the middle chitin-binding (exon-2) domains, where exon-1 and exon-3 consist of 18 and 11 copies of consensus amino acid sequences: GGRPSDSYGAPGGGN and GYSGGRPGGQDLG, respectively. The first recombinant pro-resilin or resilin-like polypeptide (RLP), namely Rec1-resilin (encoded from the exon-1 of CG15920 gene) was synthesized in mid-2000s as a water soluble polypeptide expressed in the bacteria *Escherichia coli*. The synthesized pro-resilin was photo-crosslinked (dityrosine) using a ruthenium-persulfate crosslinking system to form hydrogels, which exhibited 97% resilience, outperforming native resilin dissected from dragonfly tendon (92%), natural elastin (90%), and synthetic polybutadiene rubber (80%)

The synthesized RLPs have several advantages over other elastomeric polypeptides, such as elastin-like polypeptides (ELPs), silk-like polypeptides (SLPs), and collagen-like polypeptides (CLPs). These include:

- Unique sequences rich in uncharged, polar amino acids and devoid of canonical hydrophobic residues, and contain high proportions of glycine- and proline-rich segments.
- Average negative hydropathy index.

- Intrinsically disordered protein structure with rapidly interchangeable conformational ensemble in physiological conditions.
- Multi-stimuli (pH, temperature, ions, mechanical stress, other molecules, etc.) responsiveness, including dual-phase transition behavior (existence of both upper critical solution temperature, UCST, and lower critical solution temperature, LCST).
- Low stiffness, high extensibility, outstanding resilience, and excellent fatigue life.
- No inflammatory response.

Figure **4.7.7.9** shows possible transitions in the resilin protein that demonstrate such resiliency.

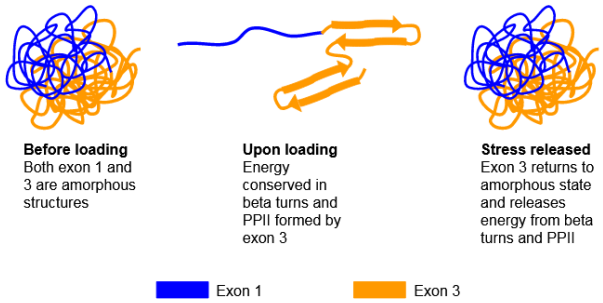


Figure **4.7.7.9**: Transitions in the resilin protein By MTLE4470 grp10 ss - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/inde...curid=58499811

Given its disordered nature, no PDB structures of resilin are available.

4.7.7.6: Fibrinogen

This very large molecule is a hexamer of three monomers (Aα, Bβ, and γ), each present in two copies ($\alpha_2\beta_2\gamma_2$). Disulfide bonds connect one structural unit (αβγ) with another. When two fibrinopeptides (FpA and FpB) are cleaved from the amino ends of the α and β chains by the clotting enzyme thrombin, small structural "knobs" form that bind to "holes" on another fibrinogen, causing the formation of large fibrils of fibrin clots. Figure **4.7.7.10** shows an interactive iCn3D model of human fibrinogen (3ghg). The "floppy" parts of the alpha chain (*α*C region) and FpA and FpB peptides are not shown as they were not resolved (due to their disorder) in the crystal structure.

NCBI iCn3D Figure **4.7.7.10**: Human fibrinogen (3ghg). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...S3dXzzsrkujKZ8

It is a very long, flexible molecule. The alpha chains are magenta and light green, the beta chains are dark blue and gray, and the gamma chains are brown and orange. Note that the helical chains are *alpha*-helical for this molecule.

Figure **4.7.7.11** shows the domain structure and hints at the flexibility of this long molecule, which is required to form a fibrous mesh clot as well as to be accessible to an enzyme, plasmin, which cleaves fibrin clots, facilitating their removal. The A*α*, B*β*, and *γ* are shown in blue, red, and green, respectively. Carbohydrates are shown in orange in the space-filling model (b).

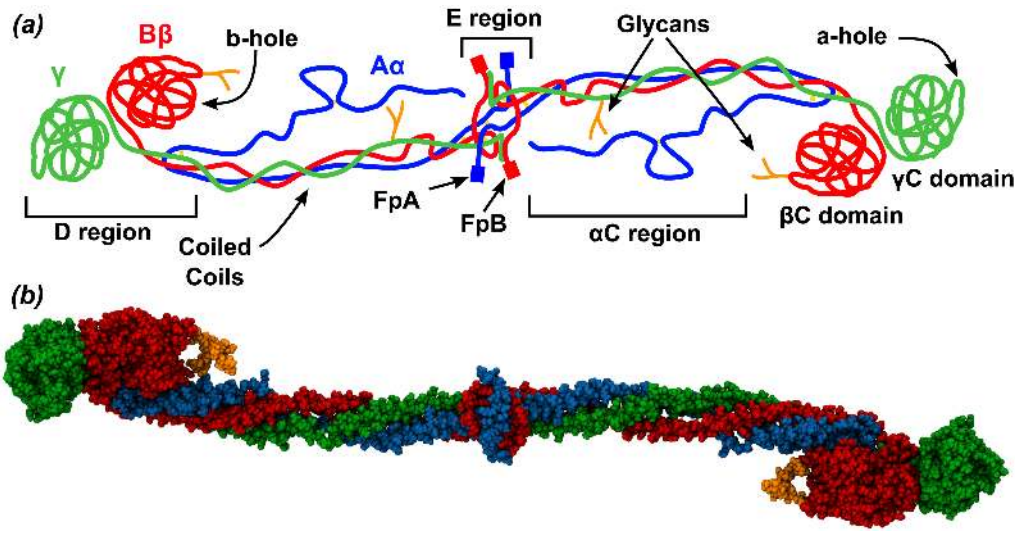


Figure 4.7.7.11: Domain Structure of FibrinogenKöhler S, Schmid F, Settanni G (2015) The Internal Dynamics of Fibrinogen and Its Implications for Coagulation and Adsorption. PLoS Comput Biol 11(9): e1004346. https://doi.org/10.1371/journal.pcbi.1004346. Creative Commons Attribution License.

The coiled coils are mostly alpha-helical. The central E region is where the N-terminal ends of all of the fibrinogen chains are located and where the fibrinopeptides A (FpA, 16 residues) and B (FpB, 14 residues) are located, where they will be cleaved by thrombin in clot formation. The C-termini of the chains are located at the distal D region, which houses two C domains. After cleavage of the fibrinopeptides, conformational changes occur in the central E region to produce the "knobs" A (with starting sequence Gly-Pro-Arg) and B (Gly-His-Arg). These knobs bind through noncovalent interactions corresponding to "holes" a and b at the distal D regions of another fibrinogen to form a dimer and, subsequently, a fibrin clot.

4.7.7.7: Myosin Heavy Chain

Myosin is a chief component of muscles and, in complex with actin and other proteins, allows muscle contraction. It has two distinct domains and an elongated, rod-like tail with 28-residue repeats of 4 heptapeptides, characteristic of alpha-helical proteins that form coiled-coil quaternary structures. Figure 4.7.7.12 shows the domain structure of human myosin heavy chain 1. The orange represents the motor domain of the protein, which binds and hydrolyzes ATP, providing the free energy that powers muscle contraction.

Figure 4.7.7.12: Domain structure of human myosin

Figure 4.7.7.13 shows a cartoon of myosin heavy chains (blue) associated with myosin light chains and how they interact with actin in the actinomyosin complex in striated muscle cells. The motor domain also binds actin filaments. Simplistically, the thick myosin filaments can slide back and forth in muscle contraction and extension.

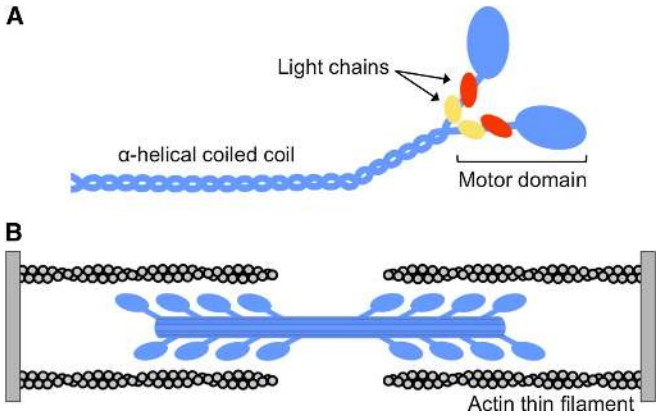
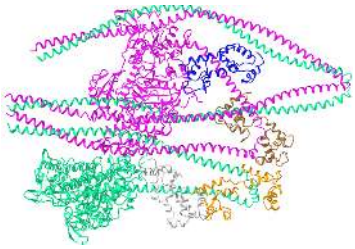


Figure 4.7.7.13: Myosin-Actin Complex. Lee, L.A., Karabina, A., Broadwell, L.J. *et al.* The ancient sarcomeric myosins found in specialized muscles. *Skeletal Muscle* **9,** 7 (2019). https://doi.org/10.1186/s13395-019-0192-3Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/

Myosin can exist in two major conformations. One is the "6S" (extended tail) form that assembles into myosin filaments, which interact with actin as shown in Figure 4.7.7.10 to transduce the chemical energy from ATP hydrolysis into mechanical forces and filament sliding. The other is the "10S" conformation, which is folded on itself. The heads interact with each other and the tail. All necessary steps (ATP cleavage, filament assembly, actin activation) required for actin/myosin-mediated contraction are inhibited in this compact form.

Figure 4.7.7.14 shows an interactive iCn3D model of an *inactive* (nonextended) conformation of myosin heavy chain (6xe9) from smooth muscle. The long cyan and green chains are the myosin (II) heavy chains from smooth muscle.



NCBI iCn3D Figure 4.7.7.14: Inactive (nonextended) conformation of myosin heavy chain (6xe9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...thwKbyQQ5ezB96

Recent Updates: February 2026

4.7.7.8: Spider Silk - an unusual fibrillar protein

Silk from both spiders and silk worms is a fibrillar protein. In its final state, silk has long, aligned protein chains that form extensive β-sheets, which in turn form fibrils, which in turn form fibers that are insoluble and extremely strong. However, before the spider spins its web, the protein exists in a soluble form called a **spidroin** (250–400 kD), which forms a family. They aggregate together into a liquid condensate in which the spidroins are intrinsically disordered (a topic we discussed in Chapter 4.6: Intrinsically Disordered Proteins) and are stored in the silk gland. The process of liquid condensate formation is another example of **liquid-liquid phase separation (LLPS)** or liquid-liquid demixing, a topic we have covered in previous sections (1.1: Cellular Foundations and 4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules). Potassium phosphate promotes LLPS. LLPS, along with changes in pH and ion exchanges in the LLPS, lead to the formation of the β-sheets fibrils, which self-associate into macroscopic fibers. This drastic conformational change during spinning is an amazing and unusual example of a rapid conversion of an IDP into a fibril-forming protein with extensive organization.

The structure of a typical spider web is shown below in Figure 4.7.7.15. The dragline is the strongest.

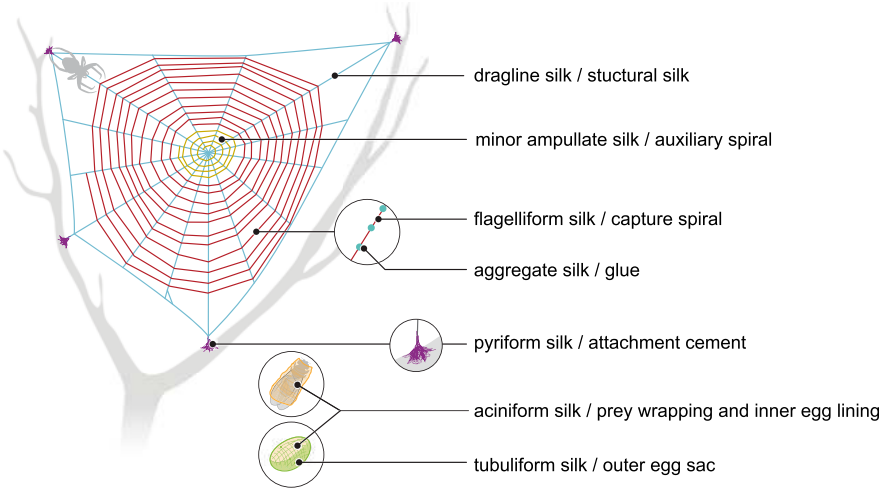


Figure 4.7.7.15: Schematic representation of different types of spider silk and their role in spider webs (compiled according to **Vollrath and Knight 2001**). Ramezaniaghdam Maryam , Nahdi Nadia D. , Reski Ralf. Recombinant Spider Silk: Promises and Bottlenecks. Frontiers in Bioengineering and Biotechnology, 10, 2022. https://www.frontiersin.org/journals...oe.2022.835637. **Creative Commons Attribution License (CC BY).**

The mechanical properties of different types of spider silk and some commonly used fibers are shown in Table 4.7.7.1 below.

Materials	Strength [MPa]	Strain [%]	Toughness [MJ m^{-3}]	Young's modulus [GPa]
MA silk (*L. venusta*)	1469	23.3	151	10.6
B. mori silk	500–600	15.0–18.0	70–80	7
Elastin	2	150	1.6	0.0011
Tendon collagen	150	12	7.5	1.5
Kevlar 49	3600	2.7	50	130
Nylon 66	750–950	–	80	2–3.6
Carbon fiber	4000	1.3	25	300
High-tensile steel	1500	0.8	6	200

Table 4.7.7.1: The mechanical properties of different types of spider silk and some commonly used fibers. Li J, Li S, Huang J, Khan AQ, An B, Zhou X, Liu Z, Zhu M. Spider Silk-Inspired Artificial Fibers. Adv Sci (Weinh). 2022 Feb;9(5):e2103965. doi: 10.1002/advs.202103965. Epub 2021 Dec 19. PMID: 34927397; PMCID: PMC8844500. **CC BY 4.0**

Here is the meaning of the terms:

- Strength: How much pulling force the fiber can withstand before it breaks
- strain: How much the fiber stretches before breaking
- toughness: How much energy the fiber can absorb before breaking
- Young's modulus: How stiff the fiber is (higher value stiffer, lower value more flexible)

We'll explore dragline silk from the black widow spider. The dragline (the strongest type of silk fiber) is predominantly composed of two types of **spidroins** called **major ampullate silk protein (MaSP).** The two types are MaSp1 (most abundant) and MaSp2. These proteins have stretches of repeating poly-alanine residues that form crystalline beta-sheet fibers. They have many glycine-rich repeats of the sequence GGX, where X is Ala, Gln, and to a lesser degree Tyr, Arg, and Ser. These are between the poly-Ala

regions. Tyr is found in all the glycine-rich regions. The motif RGG is found in every other Gly-rich region in MaSP1, while RQQ is found in the MaSP2 spidroin repeats.

Detailed structural characterization of the protein before and after spinning is complicated. Structural information from NMR spectroscopy of the MaSP1 before spinning indicates it is intrinsically disordered, with little secondary structure. Molecular structure predictions show that a MaSP forms a micelle-like structure with a buried hydrophobic core enriched in poly-Ala, with the more hydrophilic side chains (such as Arg and the hydroxyl group of Tyr) on the surface and solvent-exposed.

In the spun fiber, the poly-Ala tract forms a β-sheet crystalline structure. The Gly-rich are generally disordered. Given their high Gly content, they are likely in reverse turns (for GPGXX turns in MaSp2) and tight 3_{10}-helical (for GGX).

In general, the spidroin proteins can exist in a variety of states and undergo many transitions from the soluble monomer to the insoluble β-sheet crystalline fiber. These are described in Table 4.7.7.2 below.

Protein State	aggregation state	changes in transition to next state
1. Soluble spidroin "dope" in silk gland reservoir	monomer, "micelle" like in that poly-Ala generally buried; Gly, Arg, Tyr mostly exposed; mainly disordered; neutral pH, NaCl elevated, high protein concentration	increase phosphate ($H_2PO_4^-$); ion exchange with NaCl; phosphate interacts with Arg^+, displacing H_2O; increase Arg–Tyr cation–π interactions
2. Liquid-liquid phase separation (LLPS)	Proteins are very condensed but still hydrated with enhanced Arg-Tyr noncovalent interactions; little β-sheet, mostly disordered. Many weak side-chain interactions between monomers (multivalency)	LLPS moves through the spinning duct, leading to pH drop, shear forces, removal of water, chain alignment, β-sheet nucleation, and β-sheet formation
3. Fiber	Poly(Ala) hydrogen bonding drives β-sheet nanocrystals forming rigid domains; glycine-rich more flexible and elastic; Arg, Try are at β-sheet interfaces	

Table 4.7.7.2: Structures and transition in the spidron protein on fiber formation.

Recent structural work using NMR spectroscopy and structural predictions using Google ColabFold (v1.5.5 - AlphaFold2 using MMseqs -2) for the prespun silk and AlphaFold3 for the spun silk has been performed on shorter fragments and multimers (hexamers and trimers) of the fragments of *Latrodectus hesperus* (Western Blackwidow spider) MaSp1. The structures described in the Table above are illustrated below.

Figure 4.7.7.14 below shows the sequence of the 116-residue MaSp1 fragment from *Latrodectus hesperus (Western Blackwidow spider)* used in structural predictions and molecular dynamic simulations, with poly(Ala) (red), Arg (purple), and Tyr (blue) highlighted. It represents a typical repeat within the protein sequence.

1AAAAAAAAGGAGQGGQGGYGQGGYGQGGAGQGGAAAAAAAAGGAGQGGYGRGGAGQGG58

59AAAAAGAGQGGYGGQGAGQGGAGAAAAAAAAGGAGQGGQGGYGRGGYGQGGAGQGGAG116

Figure 4.7.7.14: sequence of the 116-residue MaSp1 fragment from *Latrodectus hesperus (Western Blackwidow spider)*

First, let's look at the simulated/predicted structure of the soluble (before spinning) disordered MaSp1 116-mer fragment in NaCl. It's shown in Figure **4.7.7.14** below (left panel). It looks similarly globular in KH_2PO_4.

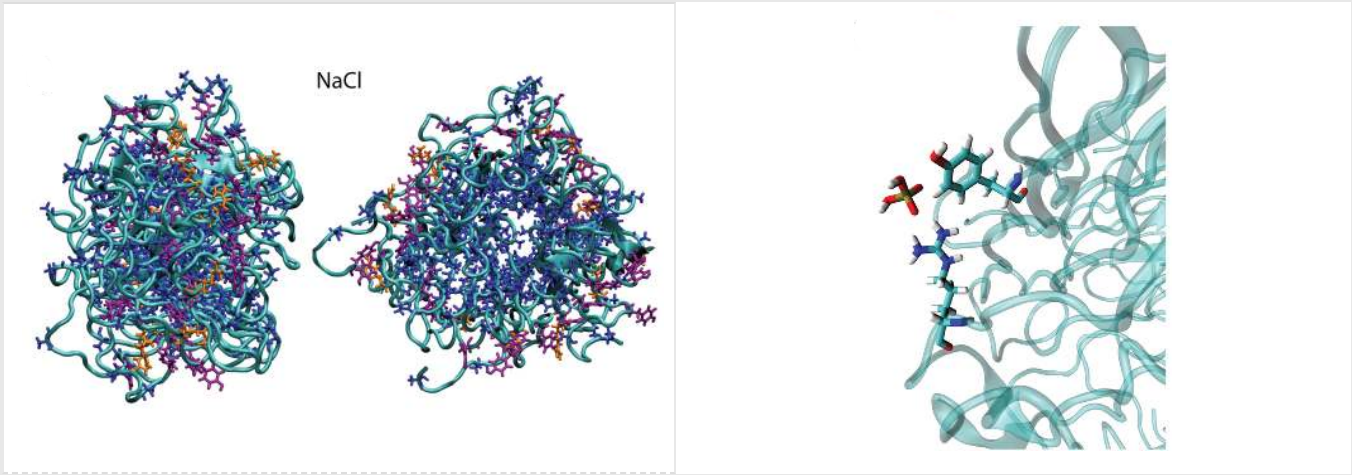


Figure **4.7.7.14**: Left: Representative structures from simulations run with NaCl. Right: Example of a phosphate-bridged Arg–Tyr interaction observed in KH_2PO_4. H.R. Johnson, K. Chalek, N. Elathram, A.T. Chau, A.R. Domingo, J.E. Aldana, H. Nguyen, A. de Loera, B.A. Duarte, L. Shapakidze, D. Onofrei, G.T. Debelouchina, C.D. Lorenz, & G.P. Holland, Arg–Tyr cation–π interactions drive phase separation and β-sheet assembly in native spider dragline silk, Proc. Natl. Acad. Sci. U.S.A. 122 (52) e2523198122, https://doi.org/10.1073/pnas.2523198122 (2025). Creative Commons Attribution License 4.0 (CC BY).

Google CoLabFold seed structure: Figure **4.7.7.14**

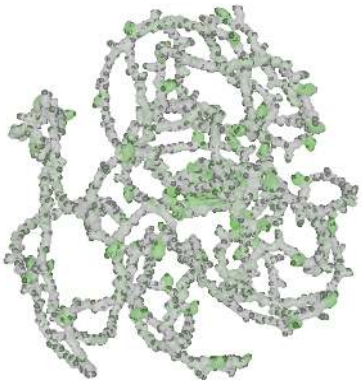


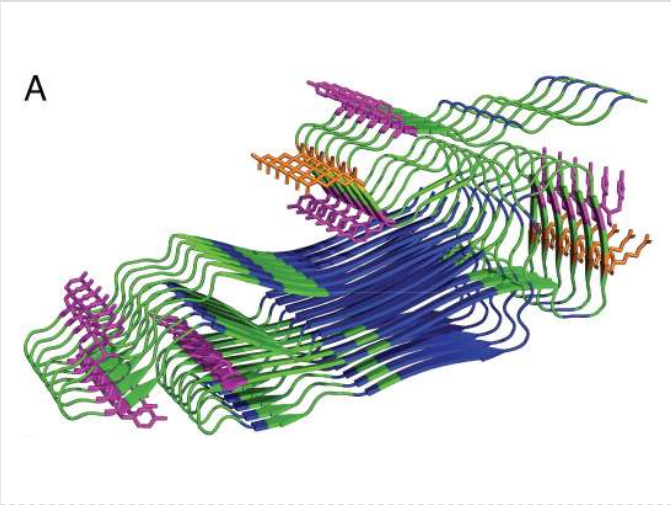
Figure **4.7.7.14**: Google CoLabFold seed structure

To view the AlphaFold3 model below in iCn3D:

1. Download this PNG appendable file. This might download as an image. If so, then resave the png file again using right-click on PC. 2. Open iCn3D. 3. **File, Open File, iCn3D png appendable,** and select the downloaded file on your computer (long load time)

The right panel shows a key interaction between the phosphate anion and proximal Arg and Tyr side chains. Molecular dynamics simulations reveal that phosphate promotes Arg–Tyr interactions and disrupts Arg–Ala contacts, freeing Ala for eventual β-sheet fiber formation.

Figure 4.7.7.15 shows that AlphaFold3 models with β-sheet fiber architecture of a hexamer of the 116mer with Arg–Tyr cation–π interactions at structured–unstructured interfaces.



To view the AlphaFold3 model below in iCn3D:
1. Download this PNG appendable file. This might download as an image. If so, then resave the png file again using right-click on PC. 2. Open iCn3D. 3. **File, Open File, iCn3D png appendable,** and select the downloaded file on your computer (long load time)

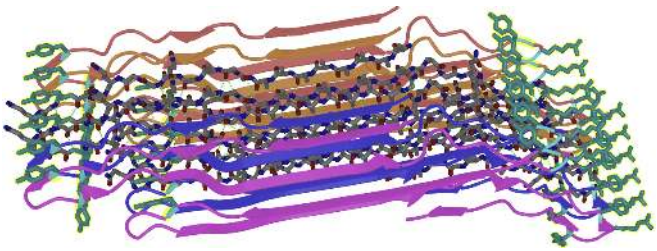
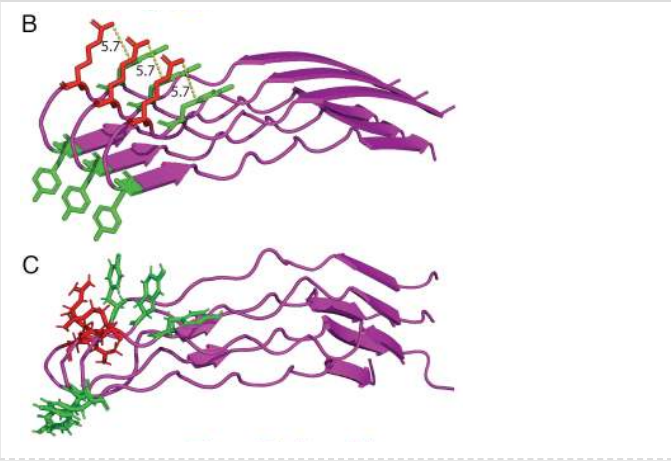


Figure 4.7.7.15: Left - (A) Hexamer AlphaFold3 model of a 116-residue MaSp1 hexamer repeat predicts highly ordered β-sheet architecture, with poly(Ala) regions forming the β-sheet core (blue), and Arg (orange) and Tyr (purple) side chains positioned near β-sheet interfaces. H.R. Johnson et al., ibid.

Right: iCn3D AlphaFold3 model of a hexamer of 116-residue MaSP1. Two monomers are shown in sticks to illustrate the interchain hydrogen bonds (dotted lines). The Tyr and Arg side chains are both shown in cyan.

Figure 4.7.7.16 below shows AlphaFold3 models of the β-sheet fiber architecture of a trimer of the 43mer with Arg–Tyr cation–π interactions at structured–unstructured interfaces.



To view the AlphaFold3 model below in iCn3D:
1. Download this PNG appendable file. This might download as an image. If so, then resave the png file again using right-click on PC. 2. Open iCn3D. 3. **File, Open File, iCn3D png appendable,** and select the downloaded file on your computer (long load time)

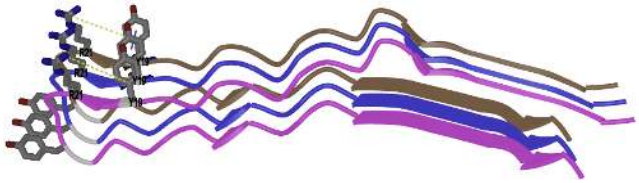


Figure 4.7.7.16: H.R. Johnson, et al., ibid.

Panel B: Trimer model of a 43-residue sequence containing Arg–Tyr motifs exhibits short β-sheets in poly(Ala) regions (purple). Arg (red) and Tyr (green) residues are located at β-sheet interfaces, with Arg side chains positioned near Tyr aromatic rings. The average distance between the Tyr Y19 ring (center-of-mass) and Arg R21 Nη atoms is ~5.7 Å.

Panel C: All-atom CHARMM36m MD simulation (200 ns) preserves poly(Ala) β-sheet structure in the trimer assembly, but reveals noteworthy sidechain disorder in interfacial Arg and Tyr residues.

Right Panel: iCn3D AlphaFold3 Model of 43-mer model. Tyrosine and arginine are shown in sticks and labeled. Two gold dotted lines show the ion-pi interactions between the adjacent Arg and Tyr on the same chain. The red dotted line between adjacent aromatic rings of the tyrosines indicated pi-pi interactions. If you open the iCn3D model and measure the distance between the Arg

and Try side chains, it is too long (around 8 Å) for an actual ion-pi interaction. The AlphaFold3 program used to make the program returned the structure shown. The manuscript shows that the distance is closer to 5.7 Å, and in MD simulations, sometimes closer.

We will discuss another type of protein, the β-amyloid proteins (Chapter 4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules) that form insoluble aggregates enriched in beta-sheets. They differ from silk, for example, in many ways. The amyloid protein monomers are soluble but are typically enriched in alpha-structure. On misfolding (which can be caused by mutations), as the initial form is metastable. The misfolded or monomers induced to misfold aggregate to form fibrils with an interchain beta-sheet structure.

4.7.7.9: Summary

This chapter explores the diverse structures and functions of elastic and fibrous proteins, which play critical roles in maintaining the mechanical integrity and specialized functions of tissues in multicellular organisms. Although many proteins adopt a roughly spherical (globular) structure, a significant subset forms elongated, fibrous structures that provide elasticity, resilience, and strength. These proteins not only support structural integrity but also enable dynamic responses to mechanical forces.

Elastic Proteins: Storage and Release of Energy

- **Elasticity and Resiliency:**
  Elastic proteins such as elastin, resilin, silk, and fibrillin store energy when stretched and release it when the force is removed. These proteins differ in their degree of resiliency: elastin and resilin exhibit up to 90% resiliency, whereas silk shows partial resilience despite its high tensile strength.
- **Molecular Features:**
  Many elastic proteins share recurring sequences and structural motifs. For example, elastin and spider silk incorporate beta-sheet domains, along with alpha-helices and beta-turns, often featuring repetitive sequences rich in proline and alanine. In resilin, high proportions of glycine and proline contribute to its intrinsically disordered structure, allowing for rapid conformational changes and responsiveness to environmental stimuli.

Fibrous Proteins: Structure and Assembly

- **Collagen:**
  Collagen is the most abundant protein in mammals, forming the extracellular scaffolds of connective tissues. Its unique triple helix, created by three polypeptide chains with a repeating (Gly-X-Y) sequence (where X is frequently proline and Y often hydroxyproline), is stabilized by interchain hydrogen bonds and $n \rightarrow \pi^*$ interactions. Post-translational modifications, such as hydroxylation (requiring vitamin C), are essential for its stability and proper function. Remarkably, collagen's stability has enabled its preservation in the fossil record for millions of years.
- **α-Keratin:**
  As a key component of hair, nails, and skin, α-keratin forms robust intermediate filaments through the assembly of coiled-coil dimers and tetramers. The formation of heterodimers between type I (acidic) and type II (basic) keratins is critical for their proper assembly into strong, insoluble fibers that rely heavily on hydrophobic interactions.
- **Fibrinogen:**
  Fibrinogen is a large, flexible glycoprotein involved in blood clotting. Comprised of six polypeptide chains ($\alpha 2\beta 2\gamma 2$) linked by disulfide bonds, fibrinogen transforms into fibrin upon cleavage by thrombin. This conversion exposes "knobs" that interact with complementary "holes" in other fibrinogen molecules, leading to the formation of a fibrous clot network essential for hemostasis.
- **Myosin Heavy Chain:**
  Myosin heavy chain is a central component of muscle contraction, containing a motor domain that hydrolyzes ATP and an elongated, alpha-helical tail that forms coiled-coil structures. Myosin exists in different conformations – an extended 6S form that assembles into filaments and a compact 10S form that is inactive – illustrating how structural dynamics are critical for its function in the actin-myosin complex.
- **Spider dragline silk:**

This is a distinctive example of a fibrillar protein. The silk proteins (spidroins) contain alternating poly(Ala) regions and glycine-rich sequences. The poly(Ala) segments form β-sheet nanocrystals in the final fiber, while the glycine-rich regions contribute flexibility and extensibility.

Unlike many classical fibrillar proteins, spidroins are initially stored in the silk gland as highly concentrated, largely disordered proteins. During fiber formation, environmental changes such as ion exchange, pH shifts, mechanical shear, and water removal trigger hierarchical assembly. Multivalent side-chain interactions, including arginine–tyrosine cation–π contacts, promote liquid–liquid phase separation, concentrating the proteins prior to fiber formation. Subsequent alignment and dehydration allow poly(Ala) segments to form hydrogen-bonded β-sheet domains that give the silk its strength. The final fiber consists of β-sheet crystalline regions embedded in a more disordered, elastic matrix.

Thus, fibrillar proteins illustrate how repetitive primary structure leads to organized higher-order assemblies. Spider silk expands this paradigm by showing how intrinsically disordered proteins can undergo controlled phase separation and then transition to β-sheet–rich fibrils, linking weak multivalent interactions with mechanically robust materials.

Integrating Structure with Function

Throughout the chapter, the interplay between protein structure and mechanical function is emphasized. Elastic proteins exemplify how specific sequence motifs and folding patterns confer the ability to absorb and release energy efficiently. In contrast, fibrous proteins such as collagen and keratin are highlighted for their role in forming robust, load-bearing structures. Moreover, the chapter delves into the molecular basis of protein stability—such as the role of post-translational modifications and noncovalent interactions—that ensures these proteins perform reliably under varying physiological conditions.

This integrated overview not only underscores the diversity of protein structures but also connects molecular details to their broader biological roles, providing a comprehensive foundation for further studies in biochemistry and materials science.

4.7.7.10: References

OpenStax, Proteins. OpenStax CNX. Sep 30, 2016 http://cnx.org/contents/bf17f4df-605c-4388-88c2-25b0f000b0ed@2.

File:Chirality with hands.jpg. (2017, September 16). *Wikimedia Commons, the free media repository*. Retrieved 17:34, July 10, 2019 from commons.wikimedia.org/w/index.php?title=File:Chirality_with_hands.jpg&oldid=258750003.

Wikipedia contributors. (2019, July 6). Zwitterion. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:48, July 10, 2019, from en.Wikipedia.org/w/index.php?title=Zwitterion&oldid=905089721

Wikipedia contributors. (2019, July 8). Absolute configuration. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:28, July 14, 2019, from en.Wikipedia.org/w/index.php?title=Absolute_configuration&oldid=905412423

Structural Biochemistry/Enzyme/Active Site. (2019, July 1). *Wikibooks, The Free Textbook Project*. Retrieved 16:55, July 16, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Enzyme/Active_Site&oldid=3555410.

Structural Biochemistry/Proteins. (2019, March 24). *Wikibooks, The Free Textbook Project*. Retrieved 19:16, July 18, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins&oldid=3529061.

Fujiwara, K., Toda, H., and Ikeguchi, M. (2012) Dependence of a α-helical and β-sheet amino acid propensities on teh overall protein fold type. BMC Structural Biology 12:18. Available at: https://bmcstructbiol.biomedcentral.com/track/pdf/10.1186/1472-6807-12-18

Wikipedia contributors. (2019, July 16). Keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:50, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Keratin&oldid=906578340

Wikipedia contributors. (2019, July 13). Alpha-keratin. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:17, July 19, 2019, from en.Wikipedia.org/w/index.php?title=Alpha-keratin&oldid=906117410

Open Learning Initiative. (2019) Integumentary Levels of Organization. Carnegie Mellon University. In Anatomy & Physiology. Available at: https://oli.cmu.edu/jcourse/webui/syllabus/module.do?context=4348901580020ca6010f804da8baf7ba.

Wikipedia contributors. (2019, July 16). Collagen. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Collagen&oldid=906509954

Wikipedia contributors. (2019, July 2). Rossmann fold. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:01, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Rossmann_fold&oldid=904468788

Wikipedia contributors. (2019, May 30). TIM barrel. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:46, July 20, 2019, from en.Wikipedia.org/w/index.php?title=TIM_barrel&oldid=899459569

Wikipedia contributors. (2019, July 16). Protein folding. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:30, July 20, 2019, from https://en.Wikipedia.org/w/index.php?title=Protein_folding&oldid=906604145

Wikipedia contributors. (2019, June 11). Globular protein. In *Wikipedia, The Free Encyclopedia*. Retrieved 18:49, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Globular_protein&oldid=901360467

Wikipedia contributors. (2019, July 11). Intrinsically disordered proteins. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:52, July 20, 2019, from en.Wikipedia.org/w/index.php?title=Intrinsically_disordered_proteins&oldid=905782287

Comprehensive Database for Protein Analysis - Biozon

SCOP: Structural Characterization of Proteins - Database showing folds, superfamiles, families, and domains

---

This page titled 4.7.7: Fibrillar Proteins is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.7: Fibrillar Proteins** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics

Learning Goals (ChatGPT o1, 1/26/25)

1. **Free Energy Landscapes and Protein Folding Dynamics**
   - Explain the concept of free energy landscapes in protein folding, including the "funnel" model versus rugged landscapes with local minima and a global minimum.
   - Analyze how thermodynamic principles and activation energy barriers govern the transitions between unfolded, intermediate, and native protein states.
2. **Molecular Dynamics and the Levinthal Paradox**
   - Describe how molecular dynamics (MD) simulations provide insight into the 3D, time-dependent process of protein folding.
   - Discuss the Levinthal paradox and explain why proteins can fold rapidly despite the astronomical number of possible conformations.
3. **Experimental Approaches to Studying Protein Folding**
   - Outline the key experimental techniques (e.g., fluorescence, circular dichroism, NMR D/H exchange) used to monitor folding transitions and measure protein stability.
   - Interpret how kinetic methods (like stop-flow kinetics) and thermodynamic measurements are used to detect folding intermediates.
4. **Anfinsen's Experiment and the Role of Disulfide Bonds**
   - Summarize Anfinsen's classic experiments on RNase folding, including how disulfide bonds are reformed during refolding and the impact of catalytic amounts of reducing agents.
   - Relate the "socks and disulfides" exercise to the probability of achieving the correct native disulfide pairing during protein folding.
5. **In Vitro vs. In Vivo Protein Folding and Aggregation**
   - Compare the conditions and challenges of protein folding in vitro versus in vivo, including the effects of concentration, aggregation, and cellular crowding.
   - Explain the formation of inclusion bodies in recombinant protein expression and the role of the cellular environment in promoting correct folding.
6. **Molecular Chaperones and Proteostasis**
   - Describe the function and mechanisms of molecular chaperones (e.g., Hsp70/Hsp40, chaperonins like GroEL/GroES and CCT/TRiC) in assisting protein folding and preventing aggregation.
   - Analyze the ATP-dependent catalytic cycles of chaperones, including how nucleotide binding and hydrolysis regulate substrate affinity and folding kinetics.
7. **Unfolded Protein Response (UPR) and ER Homeostasis**
   - Explain the unfolded protein response (UPR) and its role in managing ER stress by modulating chaperone levels, lipid synthesis, and protein degradation pathways.
   - Discuss how the UPR helps maintain proteostasis in the context of protein folding and secretion.
8. **Redox Chemistry in Protein Folding**
   - Understand the role of redox reactions in forming disulfide bonds, including the importance of oxidizing environments (e.g., ER, mitochondrial IMS) and reducing agents (e.g., glutathione).
   - Compare the enzymatic systems (e.g., PDI/Ero1 in the ER versus Mia40/Erv1 in mitochondria) that catalyze disulfide bond formation and how these processes impact protein stability.
9. **Protein Translocation Across Membranes**

- Describe the mechanisms by which proteins are targeted and translocated across cellular membranes, including the role of signal peptides, the signal recognition particle (SRP), and the Sec translocon.
- Analyze how hydrophobicity and amino acid partitioning (using scales such as the Wimley-White hydrophobicity scale) determine whether a peptide inserts into a membrane or translocates across it.

10. **Integration of Folding, Translocation, and Quality Control**

- Synthesize how protein folding is coupled with membrane translocation and how cellular quality control systems (e.g., molecular chaperones, proteasomes) prevent aggregation and misfolding in vivo.
- Evaluate how both intrinsic protein properties and cellular factors contribute to the efficiency and fidelity of protein folding under physiological conditions.

These learning goals will help you integrate concepts from thermodynamics, kinetics, and cell biology to understand the complex, dynamic process of protein folding and its regulation in living systems.

4.7.8.1: Introduction

We've seen many static and rotating models of lipid aggregates (the micelle) and proteins. We have learned some rules about the disposition of amino acid side chains in folded proteins. However, we must think dynamically and thermodynamically when considering protein folding. Figure 4.7.8.1 shows a fun but clearly unrealistic animation of how a protein might fold from an unfolded state with exposed hydrophobic side chains (orange) to a folded state when they are mostly, but not fully, buried.

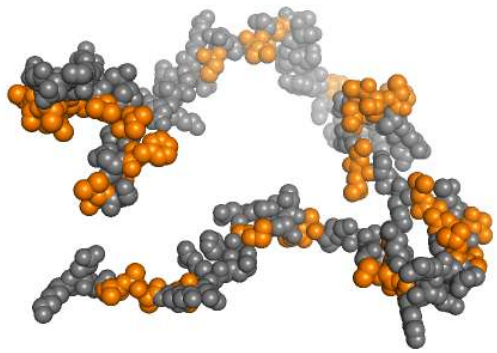


Figure 4.7.8.1: A morph showing an unfolded protein collapsing to the folded state. (CC BY-NC; Henry Jakubowski via LibreTexts)

Luckily, we have the tools of molecular dynamics (MD) at our fingertips, which help us imagine how these processes take place and, concomitantly, how to probe protein folding experimentally.

As protein folding occurs in 3D, let's explore a free energy (G) landscape for folding from an extraordinarily large number of unfolded states of higher free energy to a single low-energy folded state. Two landscapes, created by Ken Dill's group, are shown in Figure 4.7.8.2.

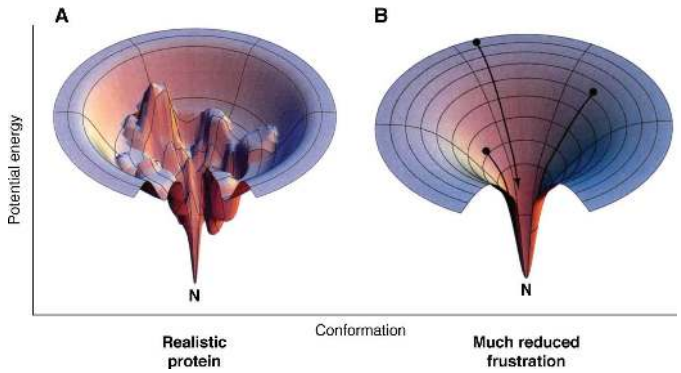
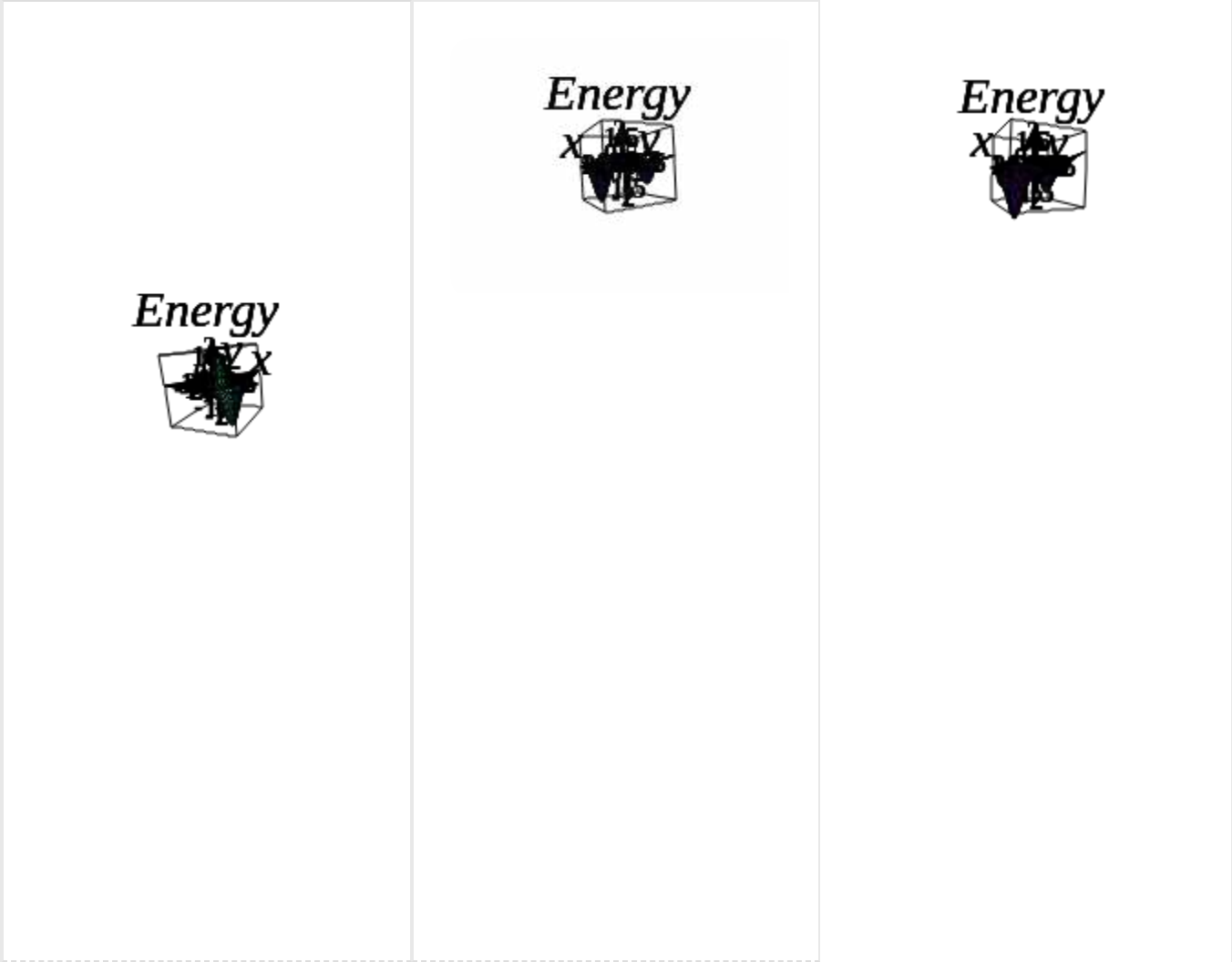


Figure 4.7.8.2: Free energy protein folding landscapes. Gershenson et al. JBC Reviews 295, 15-33 (2020). Open Access. https://doi.org/10.1074/jbc.REV119.006794. Creative Commons BY 4.0 license

The right image B shows a simplified ("much-reduced frustration") funnel view with a one-step path for any unfolded protein to reach a single global free-energy minimum state. This process occurs without any intermediates. A more realistic view is shown in A, which depicts a series of local minima and a single global minimum. Like in any activation energy curve you have encountered in chemistry courses, traversing from a local minimum to other local minimums or the global minimum is possible if enough energy is provided to overcome the activation energy for that step. Panel A implies that many intermediates lie along the path to the final global free energy minimum native state. Local minima could be populated and either stable or metastable depending on their activation energies. Intermediates might be "trapped" in these local minima. The view also conforms to our view that proteins are conformationally flexible and can adopt a variety of conformations. This is especially true for intrinsically disordered proteins.

Here are several 3D topology graphs that mimic (to some extent) the actual folding topology diagrams of proteins. Rotate them to view their shapes. Identify local and global minima, and the activation energies between them, for various paths. Use the mouse scroll wheel to resize the graphs and display the axis labels.



Now let's add some additional complexity. A protein on the path to a folded state has more hydrophobic exposure than the native state, so you would expect that it could aggregate with other self-proteins and form intermediates and end products off the normal folding pathway. These are illustrated in Figure **4.7.8.3**, which shows the "normal" protein folding (blue) and a misfolding (orange) landscape. The misfolded landscape is populated with amorphous aggregates, oligomers, protofibrils, and fibrils.

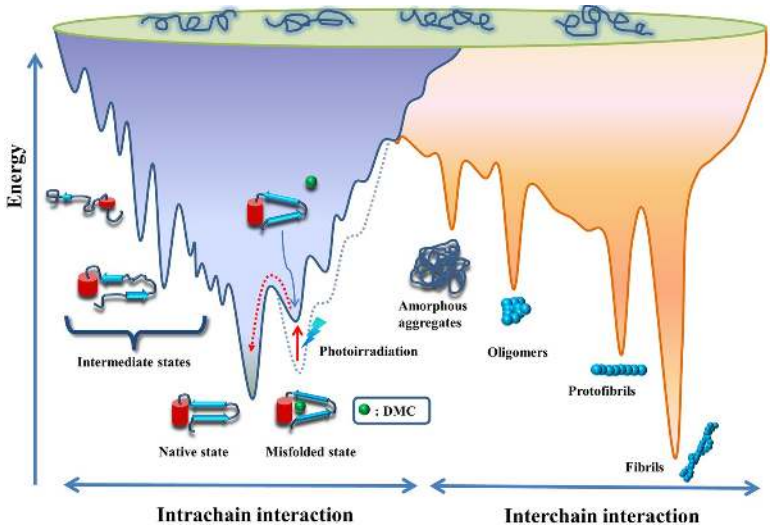


Figure 4.7.8.3: Protein Folding and Misfolding Landscape Eric H.-L. Chen et al. Scientific Report 7: 8691 DOI:10.1038/s41598-017-08385-0 . Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

We will discuss protein folding in vitro (in the lab) and in vivo (in the cell). In vitro folding is performed under very defined conditions, typically using low concentrations of small proteins to minimize the likelihood of misfolding. Folding in vivo occurs when a protein is made on a ribosome. It also occurs when a fully folded protein misfolds (such as during fevers in disease states) and is prevented from folding by association with other molecules. Folding in vivo is often assisted by other proteins called folding chaperones.

In either case, given the number of possibly nonnative states, it is amazing that proteins fold to the native state at all, let alone in a reasonable time frame. Consider this greatly simplified view of protein folding for a protein containing 100 amino acids. If each amino acid can adopt only three possible conformations, the total number of conformations could be $3^{100} = 5 \times 10^{47}$. Assuming it would take 10^{-13} s to change each conformation, the time required to "test" all conformations would be 5×10^{34} s or 10^{27} years, longer than the universe's age (13.8×10^{9} yr). Yet the protein can fold within seconds. This paradox is called the **Levinthal paradox**, after Cyrus Levinthal.

Lubert Stryer, (in his classic Biochemistry text), shows a way out of this dilemma by using an analogy of a monkey sitting at a typewriter, and typing this line out of Hamlet: "Me thinks it is like a weasel." Random typing would produce that line after 10^{40} keystrokes on average, but if the correct letters were maintained, the number of keystrokes would be a few thousand. Proteins could fold more quickly if they retain native-like intermediates along the way. Also, remember that much of conformational space is already restricted by allowed phi/psi angles (remember the empty areas in the Ramachandran plots).

We will explore the classic study of RNase folding by Anfinsen, for which he won the Nobel Prize. RNase is a small protein with four disulfide bonds and can refold from denaturing and reducing conditions in vitro. If he had chosen a larger protein, he might not have succeeded, as they are more prone to aggregation. We'll discuss mostly small proteins in this section.

Before we start, let's anticipate what might happen to fully reduced and denatured RNase with eight free Cys side chains. If the denatured protein is suddenly placed in a refolding solution lacking a denaturant and in an oxidizing environment (such as oxidized glutathione), reduced cysteine side chains could begin forming disulfide pairs. Still, only one such pair would be native. What would be the probability that a random process of forming disulfides would produce four correct ones and a fully native state? The following exercise will lead you to an answer.

? Exercise 4.7.8.1 Socks and Disulfides

You have four pairs of socks, each with a different color, as shown in Figure 4.7.8.4. You threw them into a drawer unpaired after washing. Now, without looking, take one sock out, then a second, and tie them together to form a pair. Continue doing this without looking until all four are paired. What is the probability that all the socks will be correctly matched when you are done? Once you have the answer, you're ready to understand Anfinsen's classic experiment.

PROBABILITY AND SOCKS : AN ANALOGY FOR PROTEIN FOLDING

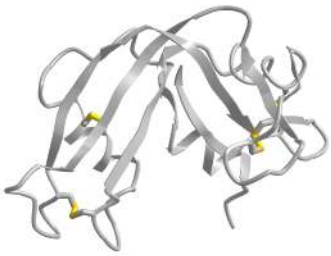
Figure 4.7.8.4: Socks and protein folding. (CC BY-NC; Ümit Kaya via LibreTexts)

Answer

We'll calculate the probability of getting four perfect matches. You pick one sock initially. You have a 1/7 chance of getting the match from the remaining seven unpaired socks. You have six socks left. Pick one again. From the remaining 5, you have a 1/5 chance of getting a match. Now, pick another. From the remaining 3, you have a 1/3 chance of getting a match. Now, pick another. You have a 1/1 chance of getting the match. Each is independent, so the total probability of getting four matched pairs of socks is given by the product of the probabilities at each step, or

$$(1/7) \times (1/5) \times (1/3) \times (1/1) = 1/105 \text{ or about } 1\%$$

The classic experiment of Anfinsen showed that, at least for some proteins, all the necessary and sufficient information required to direct the folding of a protein into the native state is present in the primary sequence of a protein. Anfinsen studied the in vitro folding of a single-chain protein, RNase, which has four intrachain disulfide bonds, as shown in the interactive iCn3D model in Figure 4.7.8.5.



NCBI iCn3D Figure 4.7.8.5: RNase with four intrachain disulfide bonds (yellow sticks) (1KF5). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Risnvb1SfNrww6

We previously discussed how chemical agents (such as beta-mercaptoethanol, a disulfide-reducing agent) can covalently modify specific protein functional groups. Some can bind via complementary intermolecular forces to the active site or other surface cavities. Other reagents, such as urea, can alter protein folding through generalized solvent effects or nonspecific interactions with the protein. Anfinsen combined two reagents, 8 M urea and beta-mercaptoethanol (βME), to unfold RNase into a nonnative state. He then removed the βME using dialysis, allowing the disulfides to reform. Next, he removed the denaturing reagent, urea. To monitor whether the protein was correctly refolded or renatured, he measured its activity relative to that of the native protein. He found that the "refolded" protein retained only 1% of its initial activity (compare this to the matching socks activity). If, however, he added catalytic amounts of βME, the protein soon retained 100% of its initial activity. He was awarded the Nobel Prize in Chemistry in 1972 for his work. A general outline of his experiment is shown in Figure 4.7.8.6.

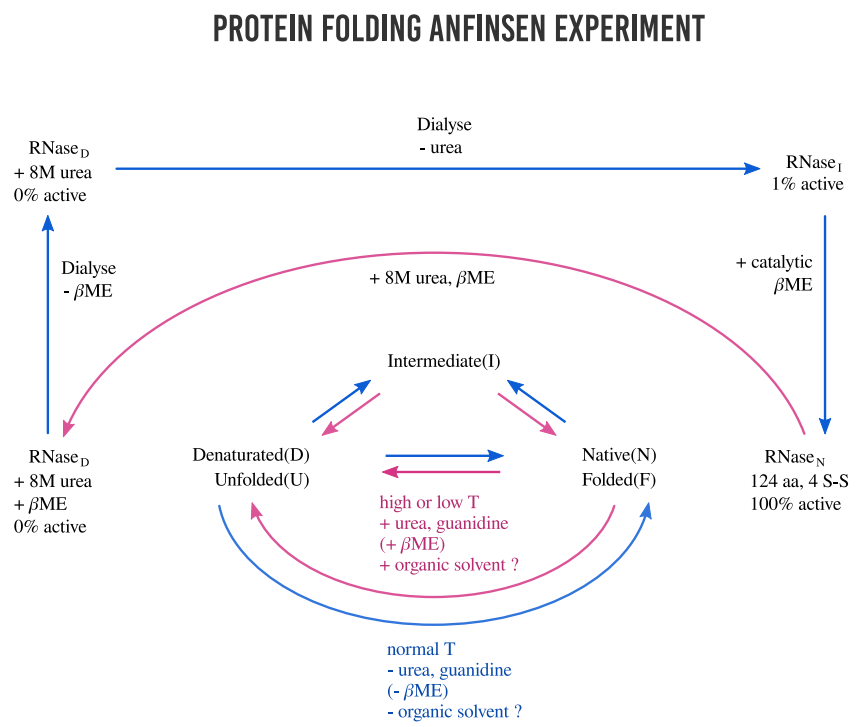


Figure 4.7.8.6: Anfinsen Experiment: Folding of RNase. (CC BY-NC; Ümit Kaya via LibreTexts)

Figure 4.7.8.7 presents a chemical mechanism to show how catalytic (non-stoichiometric) amounts of beta-mercaptoethanol can lead to full recovery of the most stable set of a protein with two disulfide bonds (right-hand side).

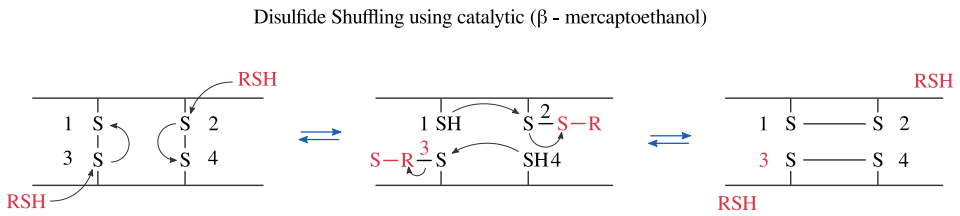


Figure 4.7.8.7: Catalytic Shuffling of Disulfides With Beta-mercaptoethanol. (CC BY-NC; Ümit Kaya via LibreTexts)

Scientists have investigated protein folding both in vitro and in vivo. In vitro experiments use urea, guanidine hydrochloride, or heat to denature the protein. Refolding is started by removing the perturbing agent (denaturing agent), using spectral techniques to follow the process. In vivo experiments involve studying intracellular proteins that assist in protein folding. An understanding of protein folding can not be separated from an understanding of protein stability and an understanding of the nature of the native and denatured state, as illustrated in the protein folding landscapes shown in Figure 4.7.8.1 and Figure 4.7.8.2.

In studying protein folding and stability/structure of the native and denatured states, equilibrium (thermodynamic) and time-dependent (kinetic) measurements are made. Folding occurs over the millisecond-to-second time scale, limiting the ability to detect intermediates. Some clever methods have been developed to study intermediates in protein folding by trapping specific intermediate structures and investigating their structure and stability in a "leisurely" fashion. Alternatively, intermediates can be studied as they occur using **stop-flow kinetics**. In this technique, a protein under denaturing conditions is rapidly mixed with a solution containing no denaturant or protein by injecting both solutions into a mixer/cuvette using syringes. The denaturant in the protein solution has been diluted to allow renaturation. Spectral measurements can begin at once.

Figure 4.7.8.8 shows a diagram summarizing these methods. Study it in conjunction with the text that follows.

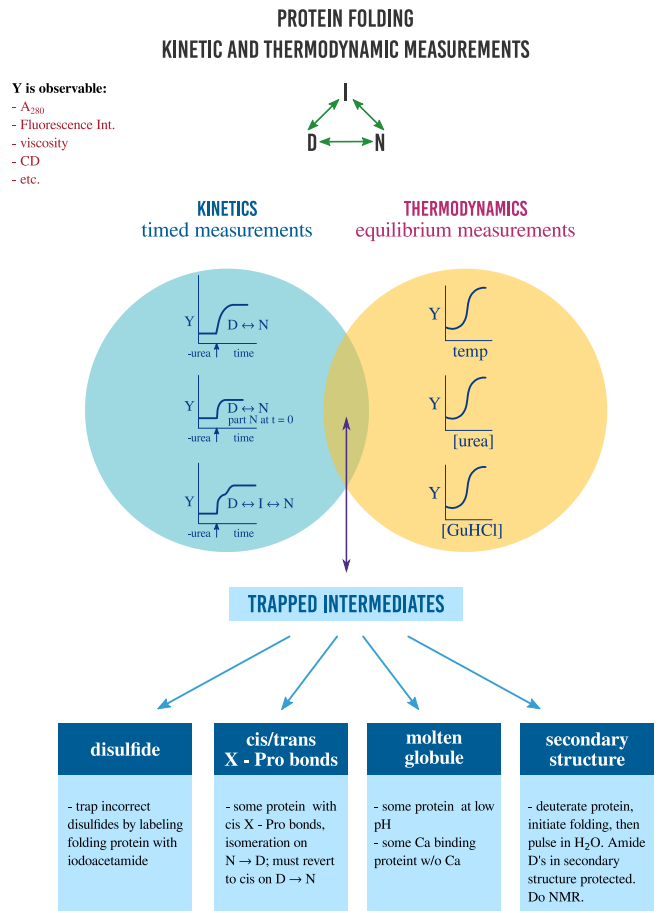


Figure 4.7.8.8: Kinetic and thermodynamic measurements of protein stability and folding. (CC BY-NC; Ümit Kaya via LibreTexts)

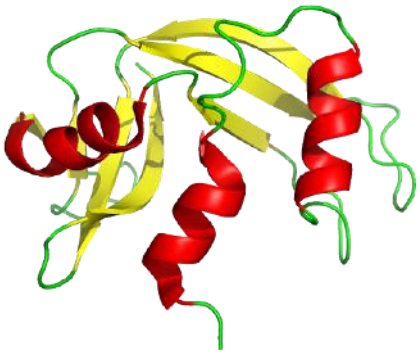
Folding appears to proceed not by an obligatory pathway but by a probabilistic or stochastic search of possible conformations. Evolution has surely selected sequences that can reach the folded state. Localized secondary structure motifs (like short alpha helices and beta turns) can form quickly (about 1 ms). The folding of small proteins, depending on their structure, occurs over a wide time frame (milliseconds to minutes). Most likely, a small number of amino acids coalesce into a core, which serves as a nucleus for folding into structures similar to the native state. Finally, packing interactions collapse the structure into the native state.

Generally, the more complex the backbone fold, the longer it takes the protein to fold. As complexity increases—i.e., the more interactions among distal regions of the polypeptide chain —the less likely it is that random interactions will lead to rapid protein folding. The folding mechanisms for larger proteins (greater than 100 amino acids) appear to proceed through intermediates, suggesting that different domains of the protein can fold independently.

4.7.8.2: Protein Folding In Vitro

Early protein folding studies involved small proteins that could be reversibly denatured and refolded. A two-state model, D ↔ N, was assumed. The denaturants were heat, urea, or guanidine HCl. Since denatured states are less compact than the native state, the solution viscosity can be used as a measure of denaturation/renaturation. Likewise, the amino acid side chains in different states would be in different environments. The aromatic amino acids Trp, Phe, and Tyr absorb UV light. After excitation, the electrons decay to the ground state through several processes. Some vibrational relaxation occurs, bringing the electrons to lower vibrational energy levels. Through radiative processes, some electrons can then fall to various vibrational levels at lower principal energy levels. The photons emitted are lower in energy and, hence, longer in wavelength. The emitted light is termed fluorescence. The wavelength of maximum fluorescent intensity and the lifetime of the fluorescence decay are very sensitive to the environment of the amino acids. Hence, fluorescence can also be used to measure changes in protein conformation. Other spectral techniques are

used, like CD spectroscopy and simple absorbance measurements. For small, single-domain proteins (such as RNase) undergoing reversible denaturation, graphs showing the extent of denaturation using each technique above, as shown in Figure 4.7.8.9, are superimposable, giving strong validity to the two-state model.



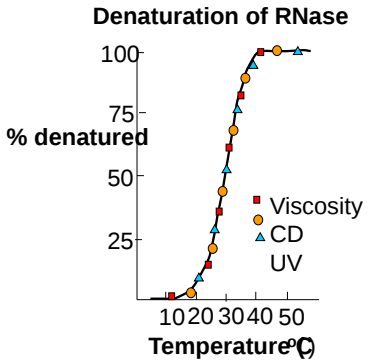


Figure 4.7.8.9: Denaturation of RNase - % Denaturation vs Temperature (after Ginsburg and Carroll, Biochemistry 4, pg 2159 (1965)

Proteins that fold without easily discernable, long-lived intermediates and follow a simple two-state model, D ↔ N are said to undergo **cooperative folding**. This simple model needed to be expanded as more proteins were studied. Some intermediates in the process were detected.

Some proteins show two steps in refolding studies, one slow and one fast, suggesting an intermediate. The longer a protein is denatured, the more likely it is to display an intermediate. One accepted explanation for this observation is that during an extended time in the D state, some X-Pro bonds might isomerize from trans to the cis state to form an intermediate. Alternatively, as in the case of RNase, which has a cis X-Pro bond in the native state, denaturation causes isomerization to the trans state. In the case of RNase, to refold, the accumulating intermediate I must reisomerize in a slow step to the cis state, followed by a quick return to the N state.

Some proteins containing multiple disulfide bonds that must reform correctly after reductive denaturation could refold into intermediates with the wrong S-S partner. Such intermediates can be trapped by stopping further S-S formation during refolding by adding iodoacetamide as shown in Figure 4.7.8.10.

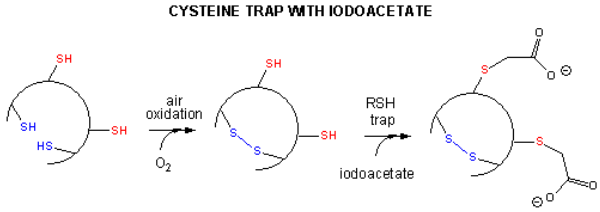


Figure 4.7.8.10: Cysteine trapping with iodoacetate.

By adding iodoacetamide at varying times along the folding pathway and separating starting (unfolded), trapped intermediates on the pathway, and the final native state by HPLC, the entire folding path can, in principle, be determined. This has been done for many proteins, including bovine pancreatic trypsin inhibitor (BPTI), a small protein with 58 amino acids, a molecular weight of

6512, and three sets of disulfide bonds (C5-C55, C14-C38, and C30-C51). The structure of BPTI with native disulfides is shown in Figure 4.7.8.11.

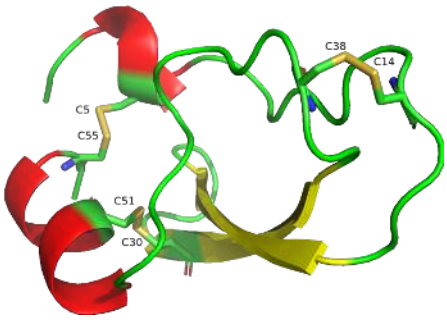


Figure 4.7.8.11: BPTI (2hex) showing 3 pairs of disulfide bonds

Given its characteristics, BPTI has been used as a model protein to study protein folding. These studies have shown that no non-native disulfides form as intermediates in the pathway. Two major intermediates form quickly, each with two of the three native disulfide bonds. These intermediates are named N', with disulfide pairs 14-38 and 30-51, and N*, with pairs 5-55 and 14-38. These two intermediates slowly form a common intermediate $N^{SH}{}_{SH}$, with disulfide pairs 5-15 and 30-51, which converts quickly to the native N state with three correct pairs (5-55,14-38, and 30-51). This folding pathway is shown in parts a and b of Figure 4.7.8.12.

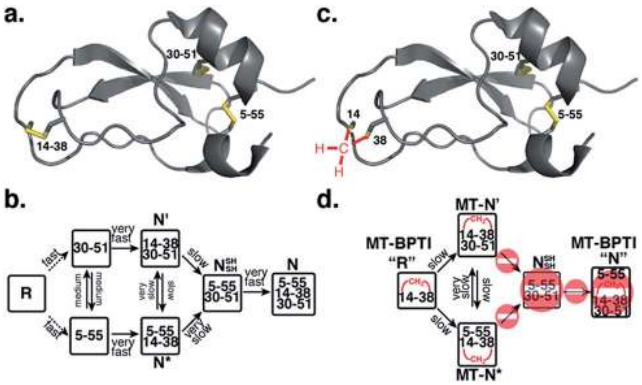


Figure 4.7.8.12: Folding Pathways for BPTI. Reem Mousa et al. *Chem. Sci.*, 2018, **9**, 4814-4820. DOI: 10.1039/C8SC01110A (Edge Article) *Chem. Sci.*, 2018, **9**, 4814-4820dkjfkdjkf. Creative Commons Attribution-Non Commercial 3.0 Unported Licence

In going from N' or N* to $N^{SH}{}_{SH}$, the **14-38** disulfide, which is very solvent-exposed (see Figure 4.7.8.11), must break before proceeding to the N (native state). Mousa et al. replaced that disulfide with a stable methylene thioacetal bridge (MT, alternative name methylenedithioether, S-CH2-S link) with the idea that once this stable (i.e., irreversible) 14-S-CH2-S38 bond formed in N' and N*, it would not break again. The formation of the methylenedithioether bond and its properties compared to disulfide bonds are shown in Figure 4.7.8.13.

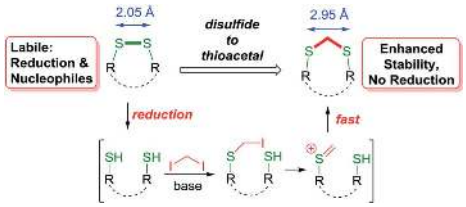


Figure 4.7.8.13: Formation and properties of the methylenedithioether bond. C. M. B. K. Kourra and N. Cramer, ***Chem. Sci.***, 2016,7, 7007-7012. **https://doi.org/10.1039/C6SC02285E.** C. M. B. K. Kourra and N. Cramer. Creative Commons. **Attribution-NonCommercial 3.0 Unported (CC BY-NC 3.0)**

Hence, the final intermediate, NSHSH, would not form, leaving just N' and N*. This is illustrated in Parts c and d in Figure 4.7.8.12.

In actuality, MT-BPTI **does** fold to the native form, whose crystal structure is virtually identical to native BPTI. Figure 4.7.8.12 shows "2D" folding plots that show the HPLC retention time of reactants, intermediates on one axis, and folding time on the other

for wild-type BPTI (part a) and MT-BPTI (part b). Hence, the folding scheme is more complicated than the one shown in Figure 4.7.8.14.

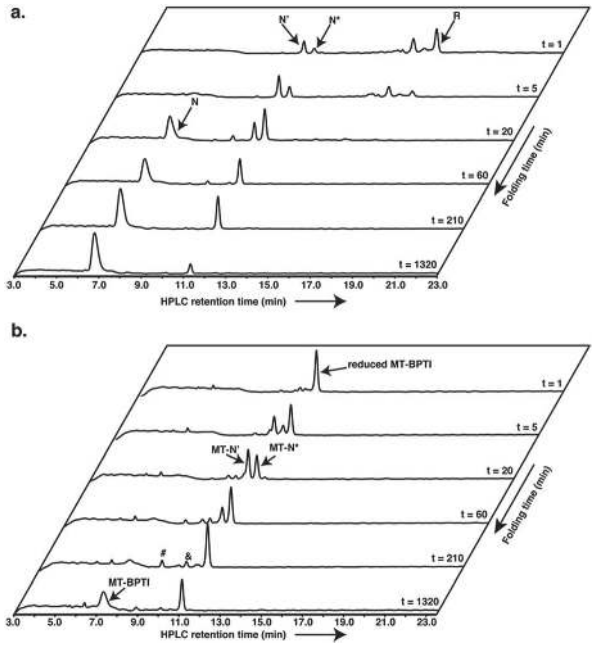


Figure 4.7.8.14: BPTI and MT-BPTI Folding Pathway In Vitro -

Some proteins form partially folded but stable intermediates under partially denaturing conditions. A good example is lactalbumin, which, under mildly acidic conditions (pH 4), low levels of guanidine HCl, or neutral pH and low ionic strength in the absence of calcium (which normally binds to the protein), forms a stable, isolatable intermediate (I) called the **molten globule (MG)**. Figure 4.7.8.15 shows the folded state with a bound calcium ion and circular dichroism spectra of the protein in various states.

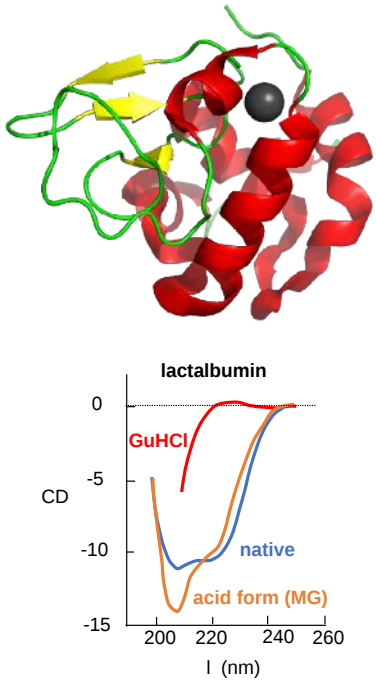


Figure 4.7.8.15: CD spectrum of Native, denatured (GuHCl) and molten globule (MG) forms of Lactalbumin (after Kuwajima. Biochemistry 24, pg 874,1985)

Data show that the MG is about 50% larger in volume than the N state. This compares to the denatured state, which can be 300% larger than the native state. Hence, it is more akin to the native state as studied by hydrodynamic techniques, but with greater solvent accessibility of hydrophobic side chains. The MG has a CD spectrum similar to that of the native state. Still, the aromatic side chains exhibit the same UV absorption and fluorescence characteristics as the protein in 6 M guanidine HCl, suggesting that the final tertiary state has not yet fully formed. The secondary structure in the MG may differ from that in the native state.

NMR can also be used to detect folding intermediates. Using this technique, proteins are unfolded in D_2O, which will cause the exchange of all Cs with ionizable protons, including the amide Hs. An amine is a weak base (pK_B around 3.5), so its conjugate acid, the protonated amine, has a pK_a of around 9.5. An amide or peptide bond would be a weaker base than an amine, since its lone pair is less available (due to delocalization through resonance) for protonation. The pKa for the conjugate acid of the amide (in which the amide N is protonated and has a plus charge) is much lower, around -0.5, than the pKa for the conjugate acid of an amine. At 2 pH units above its pKa, the amide N is close to 100% deprotonated. The pKa of the protonated group is important because the rate of H exchange depends on it, holding other variables constant. The pKa of an unprotonated amine (i.e., for the reaction $RNH_2 \rightarrow RNH^-$ is very high (30s), and hence, deprotonation of the RNH_2 is not likely under normal conditions. Figure **4.7.8.16** shows NMR D/H exchange/protection experiments for forming an alpha helix in a protein folding experiment.

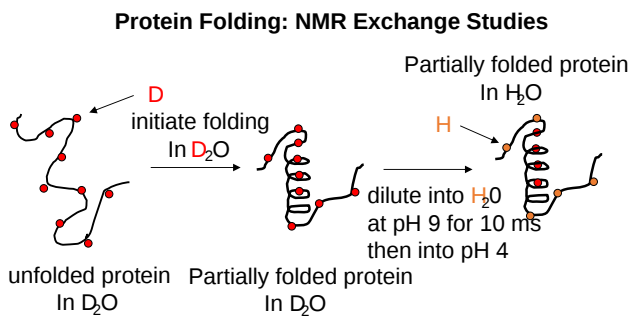
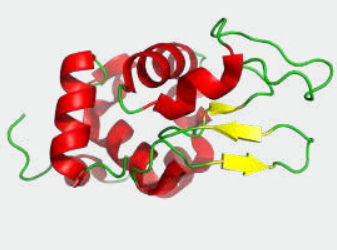


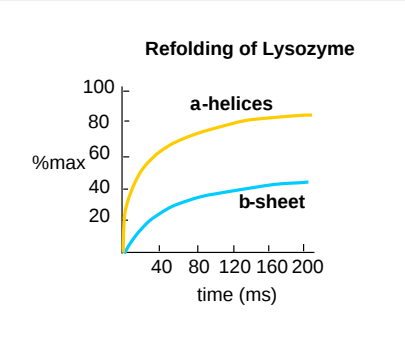
Figure **4.7.8.16**: NMR D/H exchange/protection in protein folding

Refolding is initiated by diluting the protein into a solution without the denaturant, but still in D_2O. As the protein folds and becomes more compact, the buried atoms are now sequestered from the solvent, and no longer readily exchange Ds. Then the protein is placed in H_2O at pH 9.0 for 10 ms, followed by dropping the pH to 4.0. D → H exchange is promoted at high pH and quenched for the amide Ds and Hs at low pH. Amide Hs that continue to exchange must be accessible to water. Those that aren't are usually buried in secondary structure.

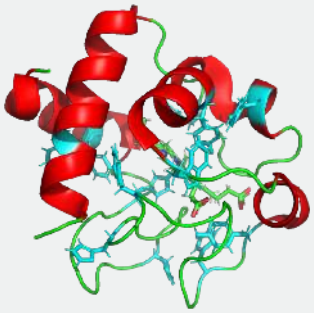
How would you interpret these structure and data?

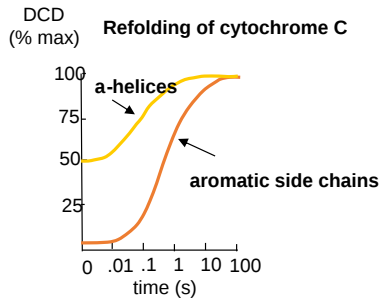
Hen Egg white lysozyme: Radford et al, Nature 358, pg 303 (1992) 2YVB





Cytochrome C: Elove et al, Biochemistry, 31, pg 6879 (1992). Cyan aromatic amino acids; Heme not shown. 5TY3





When the same techniques are applied to large, multidomain, or oligomeric proteins, only a few percent refold in vitro. Incorrect intermolecular interactions and heterogeneous aggregation appear to be the main problems that prevent correct protein folding in vitro

Here is a movie of a 6 us molecular dynamics simulation of the small protein villin.

The movie starts with the final crystal structure of villin and shows how it folds into its final structure.

With permission from the Beckman Institute for Advanced Science and Technology, National Institutes of Health //National Science Foundation, Physics, Computer Science, and Biophysics at the University of Illinois at Urbana-Champaign

4.7.8.3: The Denatured State

Although the structures of native and native-like states can be determined by X-ray crystallography and, in solution, NMR, little detailed information exists about the structures of denatured and intermediate states. Intermediate states are difficult to trap in a way that allows detailed structural analysis. In contrast to the "native" state, which consists of an ensemble of closely related states, intermediate and denatured states comprise many distinct states, making structural analysis more difficult. Religa and others from Fersht's lab have engineered a mutant of the Drosophila melanogaster engrailed homeodomain (En-HD) that enables such structural analyses. The mutation Leu16Ala (L16A) destabilizes the protein, making it denatured simply by changing the ionic strength. It is stable at high ionic strength and folds quickly under those conditions. However, it is "denatured" at physiological ionic strength, but contains significant alpha-helical structure and has nonnative contacts. It behaves like an early folding intermediate in that if placed in solutions of higher ionic strength, it rearranges to form the native state. It progressively "unfolds" to other states if placed in a lower ionic strength. Given the ambiguities in defining denatured and early folding intermediate states, Fersht's group suggests an "explicit" definition of the denatured state. They define the unfolded state (U) as the "maximally unfolded state of a protein, in which the backbone NH groups have little protection against 1H/2H exchange". They define the denatured state, D, as the "lowest energy non-native state under a defined set of conditions". In this scenario, the denatured state could also serve as a folding intermediate if placed under conditions that promote folding. Previous work from the group showed that the denatured state of En-HD has three helices protected from ^{2}H exchange and was 1 kcal/mol (4.1 kJ/mol) lower in energy than the unfolded state.

Section 4.6 discussed the added complexities to a simple 2-state D ↔ N model for protein folding. These include Silent Single-nucleotide polymorphisms, Metamorphic Proteins, and Intrinsically Disordered Proteins (IDPs).

4.7.8.4: Protein Folding In Vivo

Many differences exist between how a protein folds or unfolds in a cell and a test tube.

- The total concentration of all the proteins and nucleic acids in cells is estimated to be about 350 g/L, or 350 mg/ml. Most measurements in the lab are conducted in the range of 0.1 to 10 mg/ml.
- Proteins are synthesized in cells from an N to C-terminal direction. Hence, the nascent protein, as it emerges from its site of synthesis (the ribosome), might fold into intermediate structures since the entire protein sequence is not yet available for direct folding.
- Proteins are synthesized in the cytoplasm but must then find their final destination within the cell. Some end up in membranes. Some must translocate across one or even two membranes to reach specific organelles, such as the Golgi, mitochondria, chloroplasts (in plant cells), nuclei, lysosomes, and peroxisomes. Do they translocate in their native state?

Additional evidence suggests that protein folding/translocation requires assistance (i.e., catalysis) in the cell.

- Mutant cells defective in certain proteins can accumulate misfolded and aggregated proteins.
- Eukaryotic genes (found in higher cells with nuclei and internal organelles) can be expressed to produce proteins when transferred into prokaryotes (such as bacteria like as E. coli). Still, they often misfold and aggregate in the bacterial cells, forming structures called inclusion bodies.

Hence, recombinant proteins expressed in vivo have the same folding problems as larger proteins in vitro. In both cases, conditions favor the accumulation of nonnative proteins with exposed hydrophobic groups, leading to aggregation. Aggregation also occurs in vivo when a protein is overexpressed or expressed at a higher temperature than normal. Why? Mutant cells have been selected that suppress the formation of inclusion bodies in vivo. This effect was mediated by a class of proteins expressed by the bacteria and other cells when their temperature is raised. The function of these proteins, called **heat shock proteins (Hsp)**, was unknown until it was realized that they facilitate correct protein folding, in part, by binding to denatured proteins in the cells before they aggregate into inclusion bodies. Further studies discovered many proteins that facilitate protein folding and prevent aggregation in vivo. These proteins are now called molecular chaperones. Many are still named **Hsp#**, where # is the approximate molecular weight of the protein in kDa on a PAGE gel.

In a broader sense, a protein homeostatic environment exists within cells to maintain the proteome. This system includes chaperone proteins, which facilitate folding and refolding; the proteasome, which degrades ubiquitinated proteins; and the lysosome, which facilitates the autophagy of structures such as damaged mitochondria. Here, we will concentrate on chaperones, which are also involved in protein transport in the cell and in preventing aggregation. Some have used novel names to describe the various chaperone activities. These include holdases/translocases, unfoldases/foldases, and disaggregases, which process the various potential intermediates and end products. A simplified 2D free energy folding diagram showing the involvement of chaperones is shown in Figure 4.7.8.17.

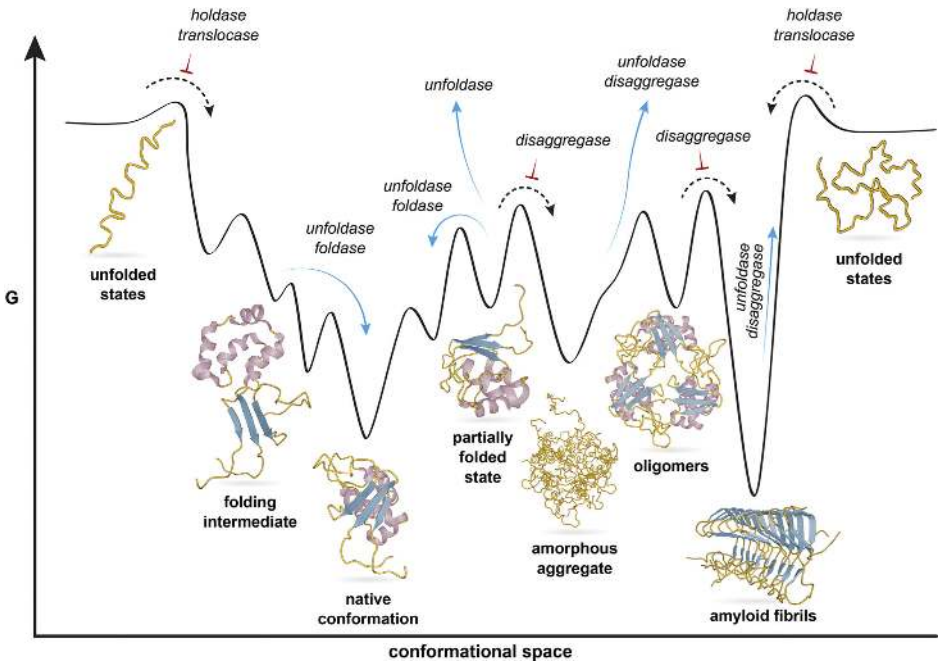


Figure 4.7.8.17: 2D Free Energy Folding Diagram with Chaperone Assist. Moseck et al. Front. Mol. Biosci., 14 June 2021 | https://doi.org/10.3389/fmolb.2021.683132 Frontiers in Molecular Biosciences 8, pg 514, 2021 https://www.frontiersin.org/article/...lb.2021.683132. **Creative Commons Attribution License (CC BY).**

The nomenclature used to categorize chaperones is confusing. In general, chaperones in all cells interact with unfolded or misfolded proteins and essentially catalyze their folding. They can do so by removing them from environments that would inhibit folding. Some can prevent aggregate formation or remove them. Some interact with proteins as they exit the ribosome during translation, while others escort damaged proteins to proteasomal degradation sites.

Many chaperones are involved in maintaining proteostasis. They can be classified by molecular size or by groups based on overall structure and mechanism. We'll use the second approach. Here are some of the key players

4.7.8.4.1: Cage Chaperonins: Oligomeric High Molecular Nanoparticles

Chaperonins, whose subunits can still be called chaperones, assemble into oligomeric, high molecular weight nanoparticle cages. These complexes are involved in the folding of larger proteins. The complexes consist of two stacked rings. **In**side is an inner cavity **in** which larger proteins can fold in isolation (an easy way to remember the name chaperon**in**). A somewhat silly analogy would be a person dressing alone in a closet. These complexes are ATPases, as ATP hydrolysis drives protein folding.

There are two main classes of chaperonins, Class I and Class II.

Group I: These are found in bacteria, in some archaea, and in mitochondria, which evolved through an endosymbiotic relationship with bacteria. Proteins in this group include **Hsp60**, which forms a homo 14-mer, and the co-chaperone **Hsp10**, which forms a homo 7-mer. In E. Coli, hsp60 is also called **GroEL**, while the small Hsp10 is called **GroES**. GroEL and GroES together form the GroEL/GroES Complex, which is shown in the interactive iCn3D model in Figure 4.7.8.18 (note: it loads slowly, given its large size).

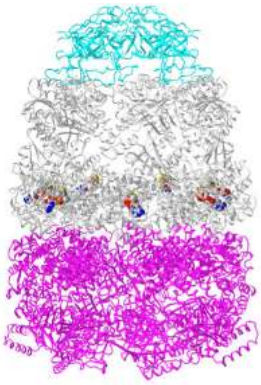


Figure 4.7.8.18: GroEL/GroES Complex (1pcq). (Copyright; author via source).
Click the image for a popup (which loads slowly) or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?JNESMy2hS9ZBTST66

The 14 GroEL subunits form separate 7-mer (magenta) and 7-mer (gray) hollow rings that stack on one another. The gray ring monomers have an ADP/AlF_3^- (spacefill) bound. Seven GroES monomers, shown in cyan, form a "lid" over the end of the complex. This hetero-21-mer displays C7 global symmetry. Lid binding requires ATP.

Figure 4.7.8.19 shows a mechanism for the GroEL/GroES protein folding cycle (shown in a linear arrangement).

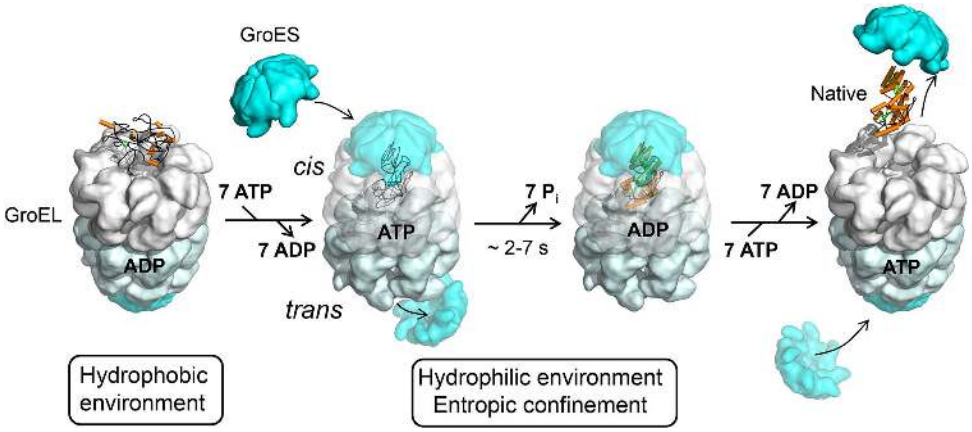


Figure 4.7.8.19: GroEL/GroES protein folding cycle in linear form. F. Ulrich Hartl Mol Biol Cell. 2017 Nov 1; 28(22): 2919–2923. doi: 10.1091/mbc.E17-07-0480. Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

In the left-most structure, a substrate protein (colored line drawing) binds in the cavity formed by the GroEL ring that does not have the GroES 7-mer cap (cyan). Rearrangement of the substrate protein and interactions with the interior wall trigger binding of ATP, binding of the top GroES cap on the cis end (same end as bound protein substrate), and release of the previously bound ADP, the distal trans-GroES cap, and a folded protein (not shown) from the trans end. Transient interactions and substrate protein conformation changes lead to the folding of the colored protein substrate within the time frame needed for effective hydrolysis of the 7-bound ATPs. The binding of ATP and a new GroES cap at the bottom triggers the release of the folded protein at the top, along with the top GroES cap. The cycle then repeats with the binding of a new target protein substrate, ATP, and a GroES cap, leading to the dissociation of ADP and the other GroES cap.

The sequestration of the bound protein substrate prevents aggregation of the folding protein. The binding of any polymer to a confined, small volume must be entropically disfavored, as the polymer's conformational flexibility is reduced. Somehow, an entropic activation energy barrier is reduced inside the cage, perhaps similar to the restraining effects of disulfides on protein folding, allowing less conformational space to be explored during folding. GroEL has also been shown to bind in its hydrophobic cavity a fluorescent CdS semiconductor nanoparticle, which can be released on the addition and cleavage of ATP

Group 2: These are found in Archaea and eukaryotes. The eukaryotic chaperonin in this group is **CCT** (cytosolic chaperonin containing TCP1), also known as **TRic** (tailless complex polypeptide one ring complex or TCP1-ring complex). These also contain two stacked rings, but this time, they are 8-mers of molecular weight 50-60K. This is similar to the Hsp60 monomers in the GroEL/GroES complex, but the monomers in Group 2 are not named with the beginning letters Hp. In addition, the complex does NOT have a cap, as in the GroES co-chaperone 7-mer in the GroEL/GroES complex. Rather, they have a built-in cap that closes on folding. The **CCT/TRic** chaperonin complex interacts with other "co-chaperones," including a **prefoldin 6-mer** complex, which inhibits aggregation and "delivers" the protein substrate to **CCT/TRic.** It also binds to phosducin-like proteins. Well-known substrate proteins for CCT/TRic include actin and tubulin (cytoskeletal proteins) and proteins involved in cell cycle control. Still, many other proteins (up to 10% of cytosolic proteins) might interact with it.

The genes encoding the eight monomers in CCT/TRic arose from gene duplication and subsequent mutation, so they are different from each other (technically, they are called paralogs). The subunits are named α, β, ξ, δ, ε, ζ, η, and θ. Figure 4.7.8.20 shows generalized structures of **CCT/TRic.**

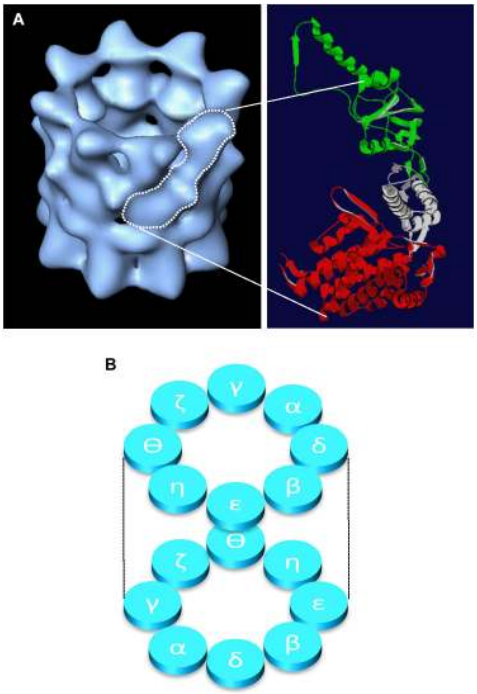


Figure 4.7.8.20: Low resolution models of the **CCT/TRic** chaperonin complex. Julie Grantham. Front. Genet., 19 March 2020 | https://doi.org/10.3389/fgene.2020.00172. **Creative Commons Attribution License (CC BY)**.

Panel A shows a cryoEM of the complex with one subunit outlined in white and an adjacent PDB structure of the alpha subunit (1A6D). The red domain binds ATP. The relative arrangements of the monomers are shown in B.

The Human TRiC/CCT complex (7LUP) with a protein substrate, reovirus outer capsid protein sigma-3, bound in the internal cavity, is shown in the interactive iCn3D model in Figure 4.7.8.21 (note: it loads slowly given its large size).

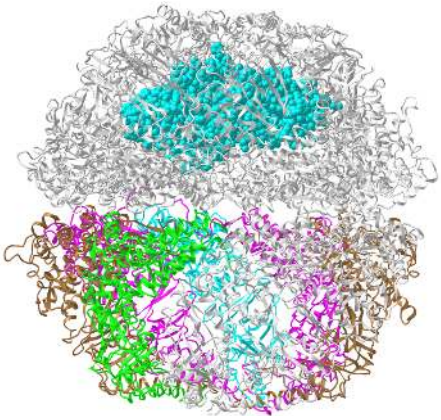


Figure 4.7.8.21: Human TRiC/CCT complex with reovirus outer capsid protein sigma-3 (7LUP) (Copyright; author via source).

Click the image for a popup (which loads slowly) or use this external link: https://structure.ncbi.nlm.nih.gov/i...Dxr3n6oy6zLsZ7

The cyan spacefill protein substrate is the encapsulated reovirus outer capsid protein sigma-3. The top ring 8-mers are in gray ribbon, while the bottom ones are in different colors. Each ring appears to bind 4 ATPs, for a total of 8, when the "attached lids" are closed.

4.7.8.4.2: Non-cage chaperones

Various chaperones not part of the nanoparticle chaperonin complex are also involved in protein folding in vivo. Some classify chaperones based on molecular weight into five classes: Hsp60 (which we discussed above as monomers in the chaperonin cage complexes), **Hsp70, Hsp90, Hsp104, and the small Hsps.** We'll focus mostly on Hsp70 and its "co-chaperone", Hsp 40, and their bacterial analogs, DnaK and DnaJ, respectively.

Hsp70/Hsp 40: Hsp70 binds to hydrophobic regions of proteins, which are more prevalent in unfolded and partially folded proteins, and helps refold them through repeated cycles of binding and release, dependent on ATP cleavage. It also helps unfold proteins through membranes and helps form/dissociate protein complexes. It's found in the cytoplasm and in various organelles, including chloroplasts, mitochondria, the nucleus, and the ER. There is a family of Hsp70 proteins (1, 1A, 1B, 2, 3, 4, 4L, 5–10, 12–18). Hsp70 proteins are made up of two regions. The **nucleotide-binding domain (NBD**) has intrinsic low **ATPase** activity and is located in the N-terminal region. The polypeptide **substrate-binding domain (SBD)** is at the carboxyl terminus. It transiently recognizes short peptides in the target protein. A flexible stretch of conserved amino acids connects the two domains. The substrate-binding domain contains a beta sheet where peptide substrates bind, and an alpha-helical region that acts as a lid.

From an evolutionary perspective, they are among the most conserved proteins. This shows the critical importance of protein folding to life. Some members of the Hsp70 family include **DnaK** (a bacterial Hsp70 widely studied with its Hsp40 **co-chaperone partner, DnaJ**), BiP (HSPA5), which assists protein folding in the ER, and alpha-crystallin in eukaryotes. Alpha crystalline comprises 30% of the lens proteins in the eye and, in part, functions to prevent nonspecific, irreversible aggregates. Some Hsp70 proteins are constitutively expressed, while others are induced by elevated temperatures or other stressors that affect protein structure, including radiation, inflammation, and exposure to heavy metals. They are key to maintaining proteostasis.

Hsp70 proteins:

- bind to growing polypeptide chains as they are synthesized on ribosomes.
- express activity as monomers.
- have ATPase activity - i.e., they cleave the phosphoanhydride ATP (which can drive reactions).
- bind short, extended peptides, which stimulate the ATPase activity
- release bound peptides after ATP cleavage

Hsp70s are highly flexible, so it is difficult to get detailed structural information when comparing bound (to target peptides) and free states. Hsp70 appears to clamp onto hydrophobic regions of target proteins, making transient interactions characterized by multiple binding and release steps until protein folding is complete. Hence, the affinity for the Hsp70 proteins must be low enough

(i.e., a high dissociation constant, K_D, for the target protein) for many binding and release steps. The affinity of the SBD for the target protein is regulated by the nucleotide-binding domain (NBD). When ADP is bound, the Hsp70 is in a **closed state** and has a high substrate affinity, so the bound target protein's release rate is slow. When ATP is bound, Hsp 70 is in the **open state** where the rate of association increases 100-fold, and the rate of dissociation increases 1000-fold, so the affinity is lowered significantly. This enables multiple cycles of release, folding, and rebinding.

As mentioned above, Hsp70s have intrinsic but low ATPase activity. When ATP is cleaved to ADP, Hsp70 returns to the closed, high-affinity state. Hence, ATP hydrolysis affects both the binding of protein substrate and the kinetics of folding. This activity increases upon binding of protein substrates and co-chaperones containing a J domain (as found in the bacterial co-chaperone DnaJ). Hence, considerable conformational changes and signaling occur between the NBD and SBDs in a process called allosterism.

The allosteric Hsp70 catalytic cycle is shown in Figure **4.7.8.22**.

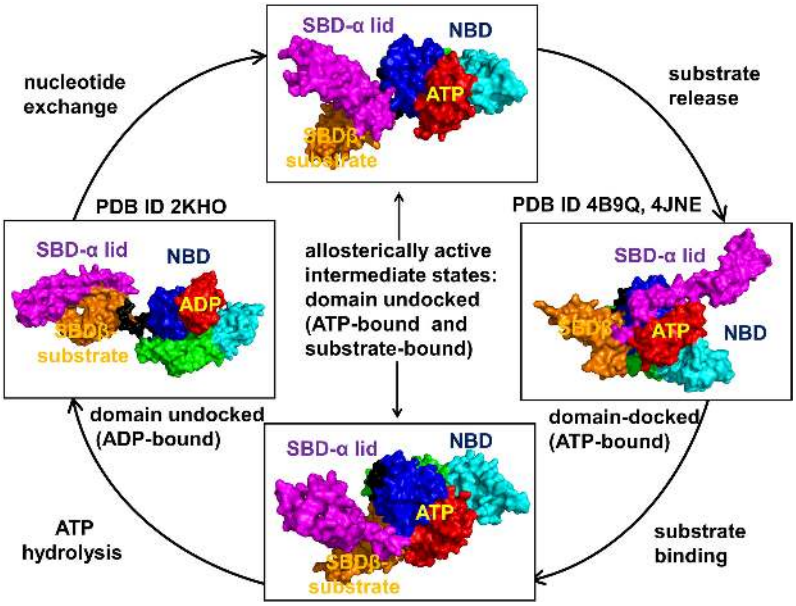


Figure **4.7.8.22**: Hsp70 Catalytic Cycle for Protein Folding. Stetz G, Verkhivker GM (2015) Dancing through Life: Molecular Dynamics Simulations and Network-Centric Modeling of Allosteric Mechanisms in Hsp70 and Hsp110 Chaperone Proteins. PLoS ONE 10(11): e0143752. https://doi.org/10.1371/journal.pone.0143752. Creative Commons Attribution License,

Let's explore the catalytic cycle, starting with the left structure. (Note that NBD subdomains are colored as follows: IA (in blue), IB (in red), IIA (in green), IIB (in cyan))

- Left structure (9 o'clock): ADP is bound, so Hsp70 is in the closed (high affinity for protein substrate)state. The SBD consists of two parts or subdomains: the α-lid (magenta) section and the β-sheet subdomain (orange), where the actual protein substrate binds (not shown here). The α-helical lid of the SBD helps to keep the protein substrate from dissociating. The PDB ID for this structure is 2KHO, but no protein substrate is evident if you explore this structure. The two domains of Hsp70, the SBD and the NBD, are connected by the flexible interdomain linker (black) and do not interact; therefore, they are shown in the **domain-undocked** state.
- Top (12 o'clock): This represents an "intermediate" in which ATP is exchanging for bound ADP, and the complex changes to the open form without releasing the protein substrate. Note that the α-lid no longer interacts as tightly with the β-sheet subdomain of the SBD.
- Right (3 o'clock): The protein substrate (again not shown) is fully released, the α-lid is fully disengaged from the β-sheet subdomain of the SBD, and Hsp70 is in the open (to substrate binding) conformation. The SBD β-sheet subdomain and the NBD domains are still engaged and are shown in the domain-docked conformation.
- Bottom (6 o'clock): Weak protein substrate binding leads to an "intermediate" in which the α-lid engages with the β-sheet subdomain of the SBD. Substrate binding promotes ATP hydrolysis, which occurs as Hsp70 returns to the full closed

conformation (left structure, 9 o'clock) with bound ADP

Figure 4.7.8.23 shows the incredible conformational change that occurs on movement from the ADP-bound higher-affinity closed form of E. Coli Hsp70 (DnaK) (2kho) to the ATP-bound lower-affinity open(to substrate binding) form (4jne). Note how the alpha-helical lid on the left-hand side of the RBD moves to engage with the NBD on the right-hand side.

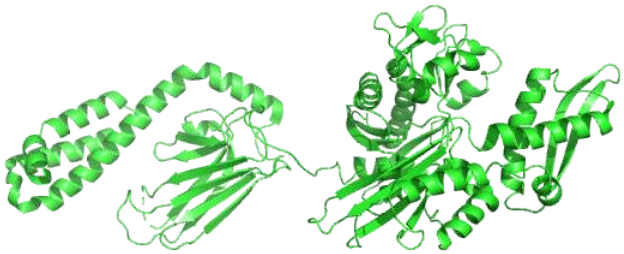


Figure 4.7.8.23: Animation of conformational changes in the receptor binding domain of Hsp70

Figure 4.7.8.24 shows how the chaperonins and Hsp70/40 (DnaK/DnaJ) work in concert.

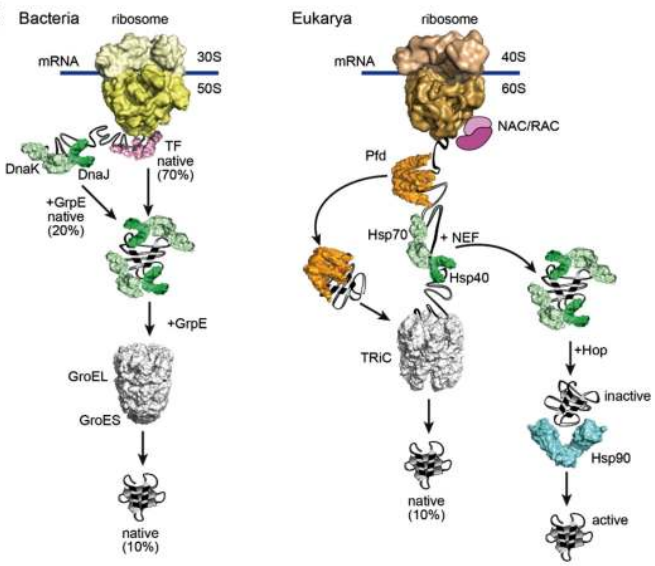


Figure 4.7.8.24: Comparison of folding pathways in bacteria and eukaryotes F. Ulrich Hartl Mol Biol Cell. 2017 Nov 1; 28(22): 2919–2923. doi: 10.1091/mbc.E17-07-0480. Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (http://creativecommons.org/licenses/by-nc-sa/3.0).

A few other chaperones acting downstream from Hsp70/Hsp40 and DnaK/DnaJ are shown in the image:

- TF (Trigger factor) and NAC/RAC: These are ribosome-binding chaperones
- Prefoldin (Pfd) binds unfolding proteins and shuffles them off to TRic.
- Hop: mediates interactions between Hsp70/Hsp40 and Hsp90
- Hsp90: This chaperone targets protein receptors (including steroid hormone receptors, protein kinases, and proteins with **n**ucleotide-binding site and **l**eucine-rich **r**epeat (NLR) domains as are found in inflammasomes) and has a cycle similar to Hsp70 with some differences. Hsp90 has N-terminal nucleotide-binding domains, middle domains that interact with protein substrates, and C-terminal dimerization domains, so the protein exists as a homodimer; when bound to ADP or in the absence of nucleotides, the homodimer forms an **open** V-shape conformation (as shown in the figure above). When ATP binds, a huge conformational change closes the V-shaped clamp. The protein has intrinsic ATPase activity, which reverses the conformational change. Many co-chaperones bind to the HSP90 dimer and modulate the cycle at different points.

Additional Proteins also catalyze protein folding at key steps in the process. Here are two examples.

- Protein Disulfide Isomerase (PDI) - converts wrongly-paired disulfides to their correct pairing. The active site consists of 2 sets of the following sequence - Cys-Gly-His-Cys, in which the pKa of the cysteines is much lower (7.3) than normal (8.5). How would this facilitate disulfide isomerization?
- Peptidyl Prolyl-Isomerase - catalyzes X-Pro isomerization by a mechanism that probably involves bending the X-Pro peptide bond. How would this facilitate the process?

Many proteins have been found to possess PPI activity. One class is the immunophilins. These small cytoplasmic proteins bind anti-rejection drugs to prevent tissue rejection after transplantation. The immunophilin FK506 binding protein (FKBP) binds FK506, while the protein cyclophilin binds the anti-rejection drug cyclosporin. The complex of cyclophilin:cyclosporin or FKBP:FK506 binds to and inhibits calcineurin, an important protein (with phosphatase activity) in immune cells (T cells) required for T cell function. In this case, immunophilin, a drug that binds to calcineurin, inhibits T cell activity, preventing immune attack on the transplanted tissue and rejection. The immunosuppressant drugs (FK506 and cyclosporin) inhibit the PPI activity of their respective immunophilin. The extent to which cyclophilin's PPI activity is required for its function is unclear, but it appears important for some of its biological effects.

Given the complexity of protein folding and the large number of chaperones involved in in vivo folding, it should not be surprising that chaperone proteins bind to protein substrates through more than hydrophobic interactions. Some chaperones (Spy and Trigger Factor) bind to charged regions on the surface of proteins, while the ER proteins calnexin and calreticulin bind to carbohydrates on glycoproteins. GroEL/GroES and TRiC/CCT also interact with protein substrates via electrostatic interactions. Nucleoplasm, a chaperone protein found in the nucleus, binds to histones, which are positively charged DNA-binding proteins.

4.7.8.5: Unfolded Protein Response

As the site responsible for the folding of membrane proteins and proteins destined for secretion, as well as the major site of lipid synthesis, the endoplasmic reticulum (ER) must maintain homeostatic conditions to ensure proper protein folding. Plasma cells that synthesize antibodies for secretion as part of the immune activation show large increases in protein chaperones and ER membrane size

The **unfolded protein response (UPR)** signaling pathway is the main pathway controlling ER biology. If demand for protein synthesis in the ER exceeds capacity, unfolded proteins accumulate. This ER stress condition activates IRE1, a transmembrane Ser/Thr protein kinase (which phosphorylates proteins). IRE1 activates a transcription factor that controls the transcription of many genes associated with protein folding in the ER. Another protein, ERAD (ER-associated degradation), returns unfolded proteins to the cytoplasm, where the proteasome degrades them. Proteins involved in lipid synthesis are also activated as lipids are needed for membranes, since the size of the ER increases. If the stress can not be mitigated, the signaling pathway leads to programmed cell death (apoptosis).

Schuck et al. investigated the specific role and importance of the UPR in ER homeostasis, as modeled in the yeast Saccharomyces cerevisiae. The UPR signaling pathway was analyzed using light and electron microscopy to visualize and quantify ER growth under various stress conditions. Western blotting was performed to determine chaperone protein concentrations after stress induction and to assess the association with ER expansion after the ER was exposed to various treatment conditions. The authors found that ER membrane expansion occurred through lipid synthesis, as stress induces the synthesis of proteins involved in lipid synthesis. Expansion failed when the proteins were absent, and the lipid concentration was low. In addition, these lipid synthesis proteins were activated by the UPR signaling pathway. By separating ER size control and UPR signaling, they found that expansion occurred regardless of chaperone protein concentrations. However, if lipid synthesis genes were unavailable, raising ER chaperone levels helped alleviate ER stress.

If unfolded proteins accumulate beyond the cellular machinery's ability to fold them, they can aggregate into structures toxic to the cells, triggering a **proteotoxic stress response** **(PSR)**. This occurs in a type of immune cell, a cytotoxic T cell, which expresses the CD8 protein on its surface. These cells normally kill virally-infected and tumor cells. In contrast to normal cells, which decrease protein synthesis in response to stress, CD8 cytotoxic T cells increase global protein synthesis in response to chronic viral exposure and in cancer. Protein aggregates form, along with stress granules, which contain precipitated aggregates of translation factors, mRNAs, RNA-binding proteins (RBPs), and other proteins, leading to the formation of exhausted T cells (T_{ex}). T_{ex} can be made artificially by growing normal effector T cells with an amino acid similar to proline called AZC, as shown in Figure **4.7.8.25** below.

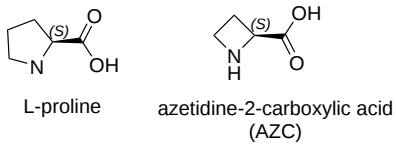


Figure 4.7.8.25: Structural comparision of Proline and AZC

Given their structural similarity, AZC is incorporated into proteins instead of proline, leading to protein misfolding and subsequent aggregation. This causes T_{ex} formation in the absence of viral challenge. In T_{ex}, there is also an increase in the expression of specialized chaperones, such as gp96 (also called GRP94) from the Hsp90b1 gene and of the chaperone BiP. BiP is an HSP70 chaperone in the ER lumen, where it binds to nascent proteins as they are translocated in the ER. It functions to prevent oligomers and aggegrates, allowing time for folding.

Finally, conversion of normal to T_{ex} cells involves the persistent activation of a special signaling pathway involving the membrane protein AKT. (For more information on AKT, go to Chapter 28.3).

These findings suggest that inhibiting the persistent activation of AKT and the specifically activated chaperones could enable the normal function of cytotoxic T cells and allow them to kill tumor cells.

4.7.8.6: Redox Chemistry and Protein Folding

In general, we envision the interior of a cell as a reducing environment. Cells have sufficient concentrations of "beta-mercaptoethanol"-like molecules (used to reduce disulfide bonds in proteins in vitro), such as glutathione (g-Glu-Cys-Gly) and reduced thioredoxin (with an active site Cys), to prevent disulfide bond formation in cytoplasmic proteins. For disulfide bonds to occur in a protein, a free sulfhydryl reacts with another one on the protein to form the more oxidized disulfide bond. This reaction occurs more readily if one of the Cys side chains has a lower pKa (due to its immediate environment), making it a better nucleophile. Most cytoplasmic proteins contain Cys with side-chain pKa > 8, which would minimize disulfide bond formation because Cys are predominantly protonated at that pH.

Disulfide bonds in proteins are typically found in extracellular proteins, which help maintain the structure of multisubunit proteins as they become diluted in the extracellular milieu. These proteins destined for secretion are cotranslationally inserted into the endoplasmic reticulum (see below), which presents an oxidizing environment to the folding protein and where sugars are covalently attached to the folding protein and disulfide bonds are formed (see Chapter 3D: Glycoproteins - Biosynthesis and Function). Protein enzymes involved in disulfide bond formation contain free Cys residues, which form mixed disulfides with their target substrate proteins. The enzymes (thiol-disulfide oxidoreductases, protein disulfide isomerases) have a Cys-XY-Cys motif and can promote disulfide bond formation or their reduction to free sulfhydryls. They are especially redox-sensitive since their Cys side chains must cycle between free and disulfide forms.

Intracellular disulfide bonds are found in proteins in the periplasm of prokaryotes and the endoplasmic reticulum (ER) and mitochondrial intermembrane space (IMS) of eukaryotes. For these proteins, the initial stage of protein synthesis (in the cytoplasm) is temporally and spatially separated from disulfide bond formation and the final folding site. Disulfide bonds can be generated in a target protein by concomitant reduction of a disulfide in a protein catalyst, leaving the net number of disulfides constant (unless an independent process reoxidizes the enzyme). Alternatively, a disulfide can be formed by electron transfer to oxidizing agents, such as dioxygen.

In the ER, disulfide bond formation is catalyzed by proteins in the disulfide isomerase family (PDI). To function as catalysts in this process, the PDIs must be in an oxidized state, capable of accepting electrons from the protein target to form disulfide bonds. A flavoprotein, Ero1, recycles PDI back to an oxidized state, and the reduced Ero1 is regenerated by passing electrons to dioxygen to form hydrogen peroxide. In summary, on the formation of disulfides in the ER, electrons flow from the nascent protein to PDIs to the flavin protein Ero1 to dioxygen (i.e., to better and better electron acceptors). The first step is a disulfide shuffle, which, when coupled with subsequent steps, leads to de novo disulfide bond formation.

In the mitochondria, disulfide bond formation occurs in the intermembrane space (IMS) and is guided by the mitochondrial disulfide relay system. This system requires two important proteins: Mia40 and Erv1. Mia40 contains a redox-active disulfide bond, Cys-Pro-Cys, and oxidizes Cys residues in polypeptide chains. Erv1 can then reoxidize Mia40, which, in turn, can be

reoxidized by the heme in cytochrome c. Reduced cytochrome C is oxidized by cytochrome C oxidase of the electron transport chain through the transfer of electrons to dioxygen to form water. The importance of IMS protein oxidation is less well understood, but it is believed that oxidative stress caused by dysfunction can lead to neurodegenerative diseases.

A recent review by Riemer et al compares the ER and mitochondrial processes for disulfide bond formation:

- Many more diverse proteins form disulfides in the ER than in the IMS. Most in the IMS have low molecular mass and have two disulfide bonds between helix-turn-helix motifs. These protein substrates include chaperones that facilitate the localization of inner-membrane proteins and proteins involved in electron transport in the inner membrane.
- There are many PDIs in the ER, probably reflecting the structural diversity of protein substrates there. However, Mia40 appears to be the only PDI in the IMS.
- "De novo" disulfide bond formation is initiated by Ero1 in the ER and Erv1 in the IMS. Convergent evolution led to similar structures in both—a 4-helix bundle that binds FAD with two proximal Cys.
- The mitochondrial pathway leads to the formation of water during the reduction of dioxygen, not hydrogen peroxide, thereby minimizing the formation of reactive oxygen species in the mitochondria. The peroxide formed in the ER is presumably converted to an inert form.
- The IMS is in closer contact with the cytoplasm via outer membrane proteins called porins, which allow some glutathione to enter the cytoplasm. The IMS presents a more oxidizing environment than the cytoplasm (with more glutathione). The ER, without a porin analog, would be more oxidizing.
- The reversible formation of disulfides in the ER regulates protein activity.

Disulfide Bond Regulation in the Periplasmic Space of Bacteria

The redox sensitivity of the Cys side chain is important in regulating protein activity. In particular, the thiol group of the amino acid Cys, an important nucleophile often found in the active site, can be modified to control protein activity. The formation of a disulfide bond or the oxidation of free thiols to sulfenic acid, and ultimately to sulfinic or sulfonic acid, can block protein activity. In a reduction reaction, the E. Coli periplasmic protein DsbA (disulfide bond A) converts adjacent free thiols into disulfide-linked Cystine. DsbB is reoxidized by DsbA back to its catalytically active form. What about a periplasmic protein like YbiS with an active site Cys? Since the environment of the periplasm is oxidizing, YbiS is protected from the oxidative conversion of the free Cys to either sulfinic or sulfonic acids, causing the protein to become inactive. The mechanism involves two periplasmic proteins, DsbG and DsbC, similar to thioredoxin. These two proteins can donate electrons to the unprotected thiol, preventing its oxidation and allowing YbiS to remain active in the periplasm. DsbG and DsbC are reduced by another periplasmic protein, DsbD, to maintain activity.

4.7.8.7: Protein Transport Across Membranes

How does a protein "decide" its final location after synthesis? Protein synthesis occurs in the cytoplasm, but proteins may end up outside the cell, in cell membranes, internalized into various organelles, or remain in the cytoplasm. How is the decision made? There must be signals in the protein that target proteins to various sites in a cell where processing can occur. Proteins destined for secretion or plasma membrane insertion typically have a signal peptide at the N-terminus that binds to a signal recognition particle in a cotranslational process, temporarily arresting translation and nascent folding. This complex docks to signal recognition complex docking sites in the endoplasmic reticulum membrane, where translation continues as the nascent polypeptide extends through a protein pore in the ER membrane. Transport across the ER membrane can also occur, partially or fully, post-translationally if nascent proteins are partially or fully folded through interactions with cytoplasmic chaperones such as Hsp70/40. Figure 4.7.8.26 shows the cotranslational (a) and post-translational (b) pathways for uptake into the ER lumen.

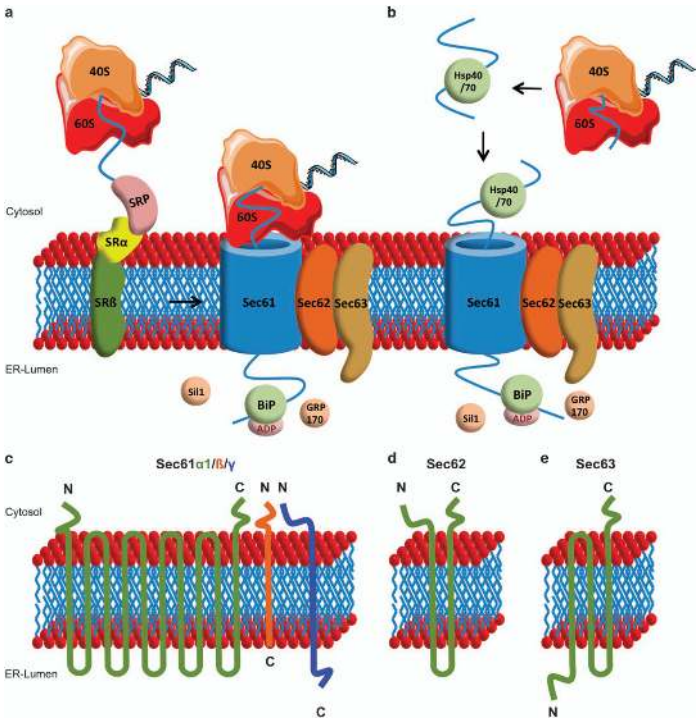


Figure 4.7.8.26: Co- and Post-translational transport of proteins across the ER membrane. Linxweiler, M. et al. *Sig Transduct Target Ther* **2,** 17002 (2017). https://doi.org/10.1038/sigtrans.2017.2 Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

In both (a) and (b), the protein is shown during synthesis as it is bound to the ribosome (40S/60S) nanoparticles. The cytoplasmic signal recognition particle (SRP) binds to the hydrophobic signal sequence of the nascent protein. The signal is typically at or near the N-terminus of the growing protein. Either co- or post-translationally, the nascent protein is delivered to Sec61, a ribosome receptor and gated pore that allows passage of the target protein. Sec61 is part of a larger ER translocon complex including Sec62 and Sec63. The membrane topology and subunit structure of the Sec proteins are shown in (c). Additional proteins, including the chaperones BiP and Hsp70/40, are also shown.

Protein transport across the endoplasmic reticulum membrane. Mechanism of (**a**) co-translational and (**b**) post-translational transport of precursor proteins through the Sec61 channel. (**c**) Topological domains of Sec61α1/ß/γ, (**d**) Sec62 and (**e**) Sec63. We note that (i) Sec63 interacts with Sec62 involving a cluster of negatively charged amino-acid residues near the C terminus of Sec63 and positively charged cluster in the N-terminal domain of Sec62, (ii) Sec62 interacts with the N-terminal domain of Sec61α via its C-terminal domain, (iii) BiP can bind to ER luminal loop 7 of Sec61 α via its substrate-binding domain and mediated by the ATPase domain of BiP and the J-domain in the ER luminal loop of Sec63, (iv) Ca^{2+}-CaM can bind to an IQ motif in the N-terminal domain of Sec61α and (v) LC3 can bind to a LIR motif in the C-terminal domain of Sec62. 40S, 40S ribosome subunit; 60S, 60S ribosome subunit; SR, heterodimeric SRP receptor; SRP, signal recognition particle.

Figure 4.7.8.27 shows an interactive iCn3D model of the Sec Complex from yeast (6ND1).

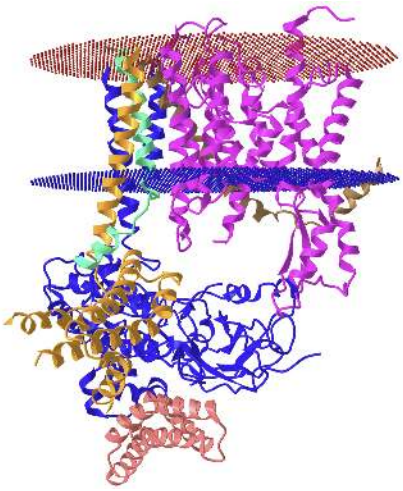


Figure 4.7.8.27: Sec Complex from yeast (6ND1). (Copyright; author via source).

Click the image for a popup (which loads slowly) or use this external link: https://structure.ncbi.nlm.nih.gov/i...6VoV8gTiUCQko8

The model shows Sec 61 and Sec 63 from Figure 25 above. Sec 61 is the ER protein-conducting channel (the analog in prokaryotes is Sec Y).

The model is colored as follows:

- Protein transport protein SEC61 (B)- magenta
- Protein translocation protein SEC63 (A) - blue
- Protein transport protein SSS1 - darker brown
- Protein transport protein SBH1 - green
- Translocation protein SEC66 - gold
- Translocation protein SEC72 - salmon

If the signal sequence of the protein to be imported is not very hydrophobic, it doesn't bind the signal recognition particle and, hence, is imported post-translationally. The Sec61 channel requires additional proteins, Sec62 and Sec63. These also required the BiP Hsp70/40 ATPase for import. Sec63 opens a gate on Sec61, leading to a wide opening that allows proteins into the lipid bilayer.

If destined for secretion, a protein enters the lumen of the ER. Proteins destined for insertion into the cell surface membrane get "stuck" in the ER membrane and, through vesiculation, merge with the Golgi and eventually with the cell surface membrane. Protein import into organelles, such as mitochondria, occurs post-translationally and requires the assistance of protein chaperones. Final protein folding occurs inside the organelle. In both cases, nonnative proteins pass through the membrane and then undergo final folding.

Amino acid sequences encoded in a protein help localize it to the ER and nucleus, as described above (for example, the signal peptide for ER translocation). However, the sequences required to direct proteins to many different organelles and specific molecular condensates remain unclear. A new AI language model, ProtGPS (in analogy to the GPS in your phone), trained on protein sequences from known organelles and compartments, can predict the locations of sequences not included in the training dataset and identify aberrant localization in mutated sequence codes. Hence, there is not only a folding code (we can generally predict a protein's 3D structure from its sequence) but also a localization code. This should be no great surprise, but it took AI to discover the code. Other factors, including the chemical microenvironment of a condensate, are also important.

An intriguing question is how the decision is made to keep a protein either in the membrane or allow it to pass through completely (in the case of proteins destined for secretion). Hessa et al. investigated this "decision-making" process by studying the eukaryotic membrane pore protein complex, Sec 61 translocon (shown in the above figures), whose activity must be closely regulated with the folding of the growing protein. In studying this process, they considered three regions of a membrane: the hydrophobic region, comprising the nonpolar acyl tails of membrane lipids; the interfacial region, in the vicinity of the polar head groups; and the aqueous regions (bulk water) on each side of the head groups. A 19-amino-acid peptide was used as the experimental model protein and added to the translocon. This length was chosen since it is long enough to span the hydrophobic part of the membrane if the

peptide were in an alpha-helical conformation (which is common in membrane-spanning proteins). They varied the proportion of amino acids that tend to partition into each of the three regions and studied the disposition of the peptide after interaction with the membrane and translocon. To test whether the results were consistent with thermodynamics of amino acid partitioning into nonpolar environments (rather than kinetic considerations), they used the Wimley-White hydrophobicity scale, based on the free energy of transfer of amino acid side chains into nonpolar environments, to predict target peptide disposition within the membrane. The table below shows the propensity of amino acids to be in each region at equilibrium based on this hydrophobicity scale.

Table: Amino Acid Partitioning Into Membrane Regions

Region	Amino Acids
Bulk water	Arg, Asn, Asp, Gln, Glu, His, Lys, Pro
Bulk water + interfacial	Ala, Cys, Gly, Ser, Thr
Interfacial	Tyr
Hydrophobic	Ile, Leu, Met, Phe, Trp, Val

Their experimental results were consistent with the predictions of the above scale. If a polyalanine 19 mer was used, no insertion was observed. With five leucines in the peptide, almost 90% was inserted into the membrane. The results would be modeled using a two-state equilibrium: Peptide inserted ↔ Peptide translocated.
They then substituted each of the twenty amino acids into a given position in a target peptide. They used the results to develop an empirical scale for membrane transfer, rather than one based on simple transfer to a nonpolar medium. This new scale matched the hydrophobicity scale, suggesting that insertion and transfer decisions were based on the thermodynamics of side chain partitioning. They also varied the position of the varied amino acids in the test peptide. If the amino acid favored the bulk and/or interfacial region, the peptide would be inserted if that amino acid were at the end of the peptide, not the middle. The peptide had to be amphiphilic for translocation, with one face polar and the other nonpolar.

They developed a simple equilibrium model to show the processes involved, as shown below in a top-down view of the membrane in Figure 4.7.8.28.

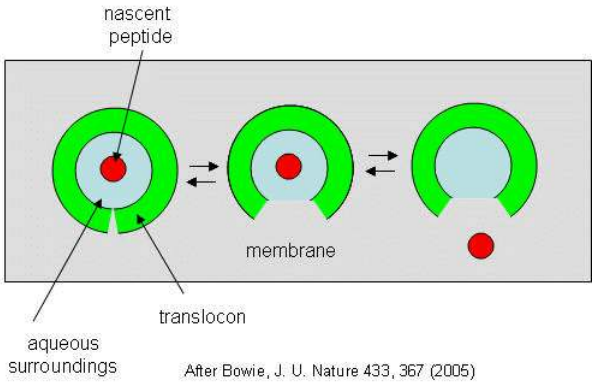


Figure 4.7.8.28: Translocon Equilibrium Model

The translocon, shown in green, has a water-filled pore and a sideways opening toward the membrane interior. The target peptide enters the pore. Transient conformational changes in the pore expose the peptide to the nonpolar membrane core. Based on the above transcription, the target peptide samples both the aqueous and nonpolar environments and partitions into them. If it partitions more favorably into the hydrophobic core, it will, causing the peptide to become membrane-bound. Otherwise, it will pass through to the other side. This can be modeled as an equilibrium process if the translocation rate is slow relative to the rates of translocon conformational change and the peptide's environmental sampling. The process becomes more complicated if the target is a large protein.

Bacterial toxin proteins also have evolved ways to pass through a cell membrane, again in a nonnative state, via a membrane protein channel. Krantz et al have recently discovered how the anthrax toxin protein moves through eukaryotic cell membranes.

Three anthrax proteins are involved. One is a "prepore" protein that binds to specific proteins on the cell membrane, where it is activated by limited proteolysis to form a pore protein that assembles into a homoheptamer in the membrane. Two other proteins secreted by the bacteria, lethal factor and edema factor, bind to the heptamer complex, and the whole assembly is then taken up into the cell by invagination, forming a vesicle with the pore complex in the membrane. This vesicle fuses with a lysosome in the cell, and upon acidification, a conformational change in the pre-pore complex triggers its activation. The lethal and edema factors unfold partially, possibly to a molten globule state, and are then passed through the pore into the cell, where they exert their toxic influences. An electrochemical potential gradient (which we will discuss later in the semester) is required to pass the factors through the membrane. The active pore further unravels the factor protein, facilitating transport.

Krantz et al. studied the pore protein by mutating two amino acids, Phe427 and Ser429, on each monomer to Cys. They then modified the Cys with [2-(trimethylammonium) ethylmethanethiosulfonate and observed effects on ion conductance of the pore and pore conformations. They noted that ion conductance was blocked when both residues were mutated and chemically modified, suggesting that these side chains were localized in the narrowest part of the channel. When Phe 427 alone was mutated to smaller side chains (Ala), ion conductance increased, but peptide transfer from the factor proteins was inhibited. This suggested that an aromatic ring in the narrow part of the channel opening participates in the translocation of bacterial proteins across the membrane. They then analyzed the transport of various small molecules with varying hydrophobicity through the wild-type pore. Their results were consistent with the binding of the molecules through hydrophobic and aromatic π-π interactions. They suggest a transport mechanism, consistent with their data, in which the unfolded protein "ratchets" through the pore, promoting the unfolding of the factor protein and exposing more hydrophobic groups to the nonpolar aromatic ring. This mechanism is similar to how the chaperone complex GroEL/GroES unfolds proteins in its large central cavity, a process that requires ATP hydrolysis rather than a transmembrane potential. In addition, the Sec61 translocon in the inner membrane of bacteria and eukaryotic ER membranes also has a pore containing a ring of hydrophobic groups (Ile).

4.7.8.8: Summary

In this chapter, we explore the dynamic nature of protein folding through the lens of thermodynamics and molecular kinetics. Unlike static images of folded proteins or lipid aggregates, protein folding is an inherently dynamic process that unfolds in three dimensions, navigating an expansive free-energy landscape. We begin by contrasting a simplified “funnel” view—where an unfolded protein smoothly funnels into a single, low-energy native state—with a more realistic rugged landscape featuring multiple local energy minima and intermediates. These intermediates, which may be transiently trapped, illustrate how proteins sample a vast array of conformations on the way to their final structure.

The discussion introduces the Levinthal paradox, which highlights the astronomical number of possible conformations that a protein could theoretically explore. Yet, proteins fold within seconds by efficiently funneling through native-like intermediates rather than random conformational sampling. This efficiency is underscored by the idea that the protein’s amino acid sequence inherently encodes the necessary information for rapid and correct folding.

A key experimental model is Anfinsen’s classic RNase refolding experiment, where a denatured protein—unfolded by urea and reducing agents—is allowed to refold in an oxidizing environment. This experiment demonstrated that the primary sequence contains all the information required for the protein to achieve its native conformation, despite the potential for myriad disulfide bond combinations (illustrated via the “socks” analogy). Moreover, the chapter covers various experimental techniques such as fluorescence, circular dichroism, and NMR D/H exchange that have been used to monitor folding transitions and probe the existence of intermediates.

Finally, the chapter contrasts in vitro folding—conducted under highly controlled conditions—with the complex, chaperone-assisted folding processes that occur in vivo. In cells, proteins fold as they are synthesized on ribosomes and often require assistance from molecular chaperones to prevent misfolding and aggregation. This integrated view of protein folding emphasizes that both intrinsic sequence information and the cellular environment contribute to the efficiency and fidelity of the folding process.

Overall, this chapter provides a comprehensive overview of the kinetic and thermodynamic principles underlying protein folding, the experimental methods used to study this process, and the biological mechanisms that ensure proper folding in vivo.

This page titled 4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.8: Protein Folding and Unfolding (Denaturation) - Dynamics** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.9: Protein Stability - Thermodynamics

Learning Goals (ChatGPT o1, 1/27/25)

1. **Thermodynamic Foundations of Protein Stability**
   - Explain the concepts of free energy (ΔG), enthalpy (ΔH), and entropy (ΔS) as they relate to protein folding and unfolding.
   - Describe how state functions and thermodynamic cycles (e.g., Hess's law) can be applied to understand and calculate stability contributions.
2. **Noncovalent Interactions in Protein Folding**
   - Identify and differentiate between the roles of hydrogen bonds, van der Waals interactions, and the hydrophobic effect in stabilizing the native state of proteins.
   - Evaluate how long-range (e.g., ion-ion, dipole-dipole) and short-range (e.g., van der Waals) forces contribute to the overall ΔG of protein folding.
3. **The Marginal Stability of Proteins**
   - Discuss why native proteins are only marginally stable under physiological conditions and how small energetic differences (e.g., equivalent to 2 H bonds) can dictate folding/unfolding.
   - Analyze how the balance between favorable interactions (hydrophobic effect, hydrogen bonding, van der Waals) and unfavorable factors (chain conformational entropy) leads to the observed marginal stability.
4. **Experimental Approaches to Assess Protein Stability**
   - Interpret denaturation data from techniques such as differential scanning calorimetry (DSC), circular dichroism (CD) spectroscopy, and fluorescence spectroscopy.
   - Explain how melting temperature (T_m) and changes in heat capacity (ΔCp) are used to quantify protein stability and folding energetics.
5. **Model Systems and Thermodynamic Cycles**
   - Use model compounds (e.g., N-methylacetamide for peptide bonds and benzene for hydrophobic groups) to illustrate how hydrogen bonding and the hydrophobic effect are quantified.
   - Construct and analyze thermodynamic cycles to determine the free energy changes associated with transferring groups from water to nonpolar environments.
6. **Effects of Additives and Environmental Conditions**
   - Predict how changes in pH, ionic strength, and the presence of denaturants (e.g., urea, guanidine hydrochloride) affect protein stability by altering electrostatic and hydrophobic interactions.
   - Describe the role of the Hofmeister series in modulating protein solubility and stability.
7. **Site-Directed Mutagenesis and Protein Stability**
   - Evaluate experimental data from mutagenesis studies that probe the contributions of specific side chains (e.g., Asn→Ala, Leu→Ala) to protein stability.
   - Compare the relative energetic contributions (in cal/mol/$Å^3$) of burying hydrogen-bonding versus hydrophobic groups in the protein interior.
8. **Protein Stability in Extremophiles**
   - Discuss the structural adaptations (e.g., increased salt bridges, tighter packing, disulfide bond formation) that enhance protein stability in thermophilic organisms.
   - Analyze how evolutionary pressures shape the amino acid composition and structural features to withstand high-temperature environments.

Each of these learning goals will help you integrate principles from physical chemistry with experimental and computational studies to develop a deep understanding of how proteins achieve and maintain their native structures under a variety of conditions.

4.7.9.1: Introduction to Protein Stability

This section is challenging but understandable using concepts from Introductory Chemistry!

This material is not easy and is perhaps the most intellectually challenging. Much of the organizational framework for this section is based on a classic article by Ken Dill, Biochemistry, 29, 7133-7155 (1990), which clearly defined the factors that contribute to protein stability.

Extrapolating from the results of studies of the transfer of small molecule H bond donors/acceptors and hydrophobic molecules from water to nonpolar solvents, it would appear that H bond interactions (as well as ion-ion interactions) do not drive protein folding per se. Rather, the largest contributors to the stabilization of the native state are the hydrophobic effect and van der Waals interactions among the tightly packed, buried atoms of the protein. However, using actual data from wild-type and mutant proteins of known structure, it appears that H bonds contribute significantly to protein folding and stability and may contribute more to the stability of the native state than the hydrophobic effect. The main factor opposing folding is chain conformational entropy, as folding proceeds from highly populated denatured states to a "single" folded state. These positive and negative factors sum to a small negative ΔG favoring protein folding, implying marginal stability of the native protein at normal temperatures.

What types of noncovalent forces might act within a protein and between proteins and solvent molecules that would cause a protein to fold spontaneously to a unique 3D structure? These forces can be long-range (ion-ion, ion-dipole, or dipole-dipole) or short-range (van der Waals repulsive and attractive forces). The interactions can be local (between adjacent amino acids in the linear sequence) or nonlocal (between sequences separated in the linear sequence but brought close together in 3D space). Clues as to what stabilizes the tertiary structure of a native protein can be gained by subjecting proteins to agents that unfold or denature proteins. Such agents include pH extremes, high concentrations of some salt solutions or organic solvents, and temperature extremes. Such experiments show that native proteins are only marginally stable (about 0.4 kJ/mol per amino acid, or about -10 kcal/mol, or -42 kJ/mol) for a protein of molecular weight 10,000 (about 100 amino acids). This is equivalent to the stability of 2 H bonds. We will consider the different types of noncovalent interactions (ion-ion, H bonds, van der Waals), as well as the stabilizing influence of the hydrophobic effect, and ask whether each is a significant driving force for protein folding. Figure **4.7.9.1** shows the relative contributions to the ΔG for protein folding.

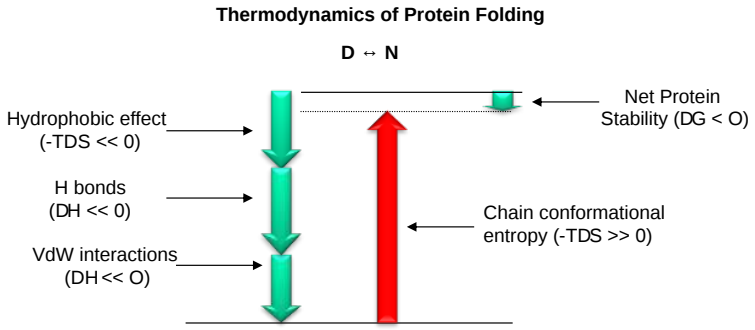


Figure **4.7.9.1**: Relative contributions to ΔG for protein folding

Most of this chapter will deal with H bonding and the hydrophobic effect. A theme of any biochemistry course is that if you can understand the interactions among small molecules, you can apply that knowledge to understanding larger molecules like proteins. To understand if H bonds within proteins, often buried in the more hydrophobic interior of the protein, drive protein folding, we will first examine the thermodynamics of H bond formation of a small molecule, N-methylacetamide, in water and a nonpolar solvent. To understand if the hydrophobic effect, mediated by the burying of nonpolar side chains within the more nonpolar center of the protein, drives protein folding, we will examine the thermodynamics of benzene solubility in water. Most recent studies involve the creation of specific mutants at amino acid positions to reveal the contributions of H-bonding and the hydrophobic effect to folding and protein stability.

4.7.9.2: Ion - Ion Interactions

These could be investigated by altering pH or ionic strength. Why is that?

a. General Charge Interactions - Proteins denature at either low or high pHs when they have a maximal positive or negative charge, respectively. Small proteins that seem to fit a simple two-state folding model (F ↔ U) have a characteristic melting temperature (T_M) at which 50% of the proteins in the population are unfolded. The higher the T_M, the more stable the protein. The graphs of the T_M for different proteins at different pH values are shown in Figure **4.7.9.2**.

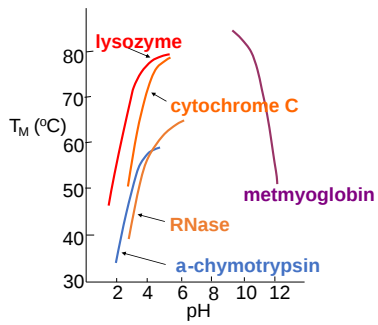


Figure **4.7.9.2**: T_M for small proteins at different pHs. After Dill, Biochemistry, 29, 7133-7155 (1990)

Under extremes of pH (but not so great as to catalyze peptide bond cleavage, electrostatic repulsions would cause the protein to denature. In the folded, compact state, charge density (q/Vol) increases at pH extremes, which could be alleviated by unfolding to a less dense state. But what about specific charge pair interactions? In contrast to the general charge interactions, these might stabilize a protein. Are they the predominant factor that determines stability?

b. Specific Charge Interactions (charge pairs) - If ion pairs are the source of protein stability, you would expect that high salt could disrupt them and lead to denaturation. Although some salts denature proteins, others stabilize them. Other evidence argues against this idea. Ion pairs are not conserved in evolution. In addition, the number of ion pairs in proteins is small (approx. 5/150 residues, with one of those on average buried). Also, the stability of a protein shows little dependence on pH or salt concentration (at low concentrations) near the isoelectric point, the pH at which a protein has a net charge of zero.

The overall charge state affects not only a protein's stability but also its solubility. Proteins are most insoluble at their isoelectric point, pI, since at that pH value (where they have a net zero charge), the proteins experience the least electrostatic repulsion. They are, hence, most likely to aggregate and precipitate. Low salt concentration also promotes insolubility. Mutagenesis studies show that solubility can be increased by replacing nonpolar groups on the surface with polar ones. Pace (2009) cites studies on RNase S, in which Thr 76 is maximally exposed. If it is replaced with aspartic acid, solubility increases to 43 mg/mL; if it is replaced with tryptophan, solubility decreases to 3.6 mg/mL. His, Asn, Thr, and Gln have lower solubility near their pIs than Ala, a surprising result. Similar results were obtained with Arg and Lys compared to Ala. Smaller side chains, Asp and Ser, at position 76 increased solubility compared with Ala.

4.7.9.3: Hydrogen Bonding

Linus Pauling first suggested that H bonds (between water and the protein and within the protein itself) would play a dominant role in protein folding and stability. It would seem to make sense since amino acids are dipolar, and secondary structure is common. Remember, however, that H bonds are found not only in the native state but also in the denatured state. Likewise, there are H-bonds between water and proteins in both the native and denatured states. Do H bonds in proteins contribute differently to the stability of the D vs N states? Many experimental and theoretical studies have investigated helix ↔ (random) coil transitions in small peptides. Remember all the intrachain H bonds in the helix? Are they collectively more stable than H bonds between water and the peptide in a (random) coil?

In this section, we will explore the effects of H bonds on small molecules and infer from them the likely contributions of H bonds to protein stability. After all, if we can't understand small molecules and their interaction, how can we understand the same interactions in big ones like proteins? However, when we use modern site-directed mutagenesis tools to explore H bonds in proteins, we reach different conclusions about their relative roles. So, consider this immediate section an exercise in model-building and critical thinking.

In thinking about conformational studies involving small peptides, it is useful to apply Le Chatelier's Principle to the equilibrium below:

random coil ↔ helix

Anything potential "perturbant" (small molecules, solvent, etc) that would preferentially interact with the helical form would push the equilibrium to the helical form.

Early models assumed that intrachain H bonds were energetically (enthalpically) more favorable than H bonds between peptide and water. But to form a hydrogen bond requires an entropy payback since a helix is much more ordered (lower entropy) than a random coil (higher entropy). At low temperatures, enthalpy predominates, favoring helix formation in solution. At high temperatures, the helix is disfavored entropically. Imagine the increased vibrational and rotational states permitted to the atoms at higher temperatures. Theoretical studies on helix-coil transitions predicted the following:

- as the chain length increases, the helix gets more stable;
- increasing the charge on the molecule destabilizes the helix since the coil, compared to the more compact helix, has a lower charge density;
- solvents that protonate the carbonyl oxygen (like formic acid) destabilize the helix and
- solvents that form strong H bonds compete with the peptide and destabilize the helix. In contrast, solvents such as $CHCl_3$ and dimethylformamide (a nonprotic solvent) stabilize the helix. Likewise, 2-chloroethanol and trifluoroethanol, which form none or weaker H bonds to the peptide than water, stabilize the helix. (In the case of trifluoroethanol, molecular dynamics simulations have shown that TFE preferentially interacts with (solvates) the peptide, which inhibits H bonds from the peptide backbone to water, stabilizing the intrahelical H bonds.

These helix-coil studies suggest that H bonds are important in stabilizing a protein.

But do they? Why should these H bonds differ from those in water? It's difficult to determine whether they are, since there are so many possible H bonds (between protein and water, water and water, and protein and protein), and their strength depends on their orientation and the dielectric constant of the medium in which they are located.

If intrachain H bonds in a protein are not that much different in energy than intermolecular H bonds between the protein and water, and given that proteins are marginally stable at physiological temperatures, then it follows that the folded state must contain about as many intramolecular hydrogen bonds within the protein as possible intermolecular H bonds between the protein and water; otherwise the protein would unfold.

To resolve this issue and determine the relative strength of H bonds between various possible donors and acceptors, many studies have been conducted to compare the energies of H bonds between small molecules in water with those in a nonpolar solvent. The rationale goes like this. Suppose the interior of a protein is more nonpolar than water (a lower dielectric constant than water). In that case, intrastrand H bonds in a protein might be modeled by considering the H bonds between small molecules in nonpolar solvents and asking whether the free energy change for the following process is negative.

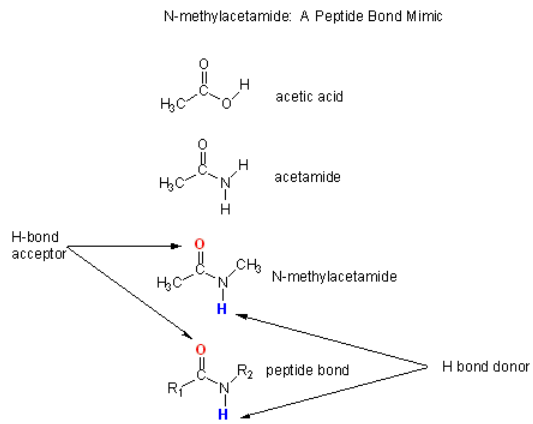
$$Dw + Aw \leftrightarrow (DA)n,\ \Delta G^o,\ K$$

D is a hydrogen bond donor (like the H in N-H), A is a hydrogen bond acceptor (like the O in C=O), w is water (i.e., donor and acceptor are in water), and n is a nonpolar solvent. ΔG^o and K are the standard free energy change and the equilibrium constant, respectively, for forming an H-bond in a nonpolar solvent from a donor and acceptor in water. This reaction simulates H-bond contributions to protein folding, where a H-bond mimics a buried H-bond in a nonpolar solvent. The reaction above is based on a thought experiment since setting up the necessary conditions for the measurement would be hard. However, we can calculate ΔG^o for this reaction since it is a state function, and it doesn't matter how one accomplishes this process.

In introductory chemistry courses, you learned about state functions, so this should not be a new concept. Remember, a state function is one in which the variable describing two connected states does not depend on the path between the states. A simple example involves potential energy and height. Imagine you climb a mountain to the peak. It doesn't matter which path you take up the mountain since the height difference (Δ height) is the same if you start at the same

point and end up at the peak. Your change in potential energy is also the same. ΔG^o for protein folding is also a state function for a given set of conditions (temperature, pressure, solvent)

A specific example is the formation of H bonds between two N-methylacetamide (NMA) molecules in water and in a nonpolar solvent. N-methylacetamide is a good mimic for the H bond donors and acceptors of the peptide bond of a polypeptide chain, as shown in Figure \(\PageIndex{3}\) below.



(\PageIndex{3}\): N-methylacetamide as a mimic for the peptide backbone

The reaction scheme shown in Figure **4.7.9.4** describes a set of reactions (a thermodynamic cycle) involving the formation of H-bonded dimers of NMA. A and B are molecules of NMA, in either water (w) or a nonpolar solvent (n). N-methylacetamide is a good mimic for the H-bond donors and acceptors of the peptide bond of a polypeptide chain.

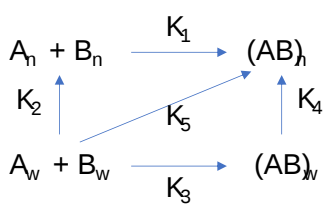


Figure **4.7.9.4**: Thermodynamic cycle for forming H bond in nonpolar solvent from donor and solvent in water.

K_1 is the equilibrium constant for the dimerization of NMA in a nonpolar medium. This can be readily determined and is >1, implying that $\Delta G^o < 0$. (Remember, ΔG^o = -RTlnKeq) For the dimerization of NMA in CCl_4, ΔG^o_1 = -2.4 kcal/mol (-10 kJ/mol).

K_2 is the equilibrium constant (think of it as a partition coefficient) for transferring two NMA molecules from water to a nonpolar solvent (again easily measurable). For NMA transferring from water to CCl_4, ΔG^o_2 = + 6.12 kcal/mol (+25.6 kJ/mol).

K_3 is the equilibrium constant for the dimerization of NMA in water. This can be readily determined and is <1, implying that $\Delta G^o > 0$. For the dimerization of NMA in water, ΔG^o_3 = +3.1 kcal/mol (+13 kJ/mol).

K_4 is the equilibrium constant (think of it as a partition coefficient) for transferring a hydrogen-bonded dimer of NMA from water to a nonpolar solvent. You try to think of a way to measure that! I can't. This is where thermodynamic cycles come in so nicely. You don't have to measure it. You can calculate it from $K_{1\text{-}3}$ since ΔG^o is a state function!

$$\begin{aligned}
&\Delta G^0_2 + \Delta G^0{}_1 = \Delta G^0_3 + \Delta G^0_4 \text{ or} \\
&-\mathrm{RTlnK_2} + -\mathrm{RT}\ln \mathrm{K_1} = -\mathrm{RT}\ln \mathrm{K_3} + -\mathrm{RT}\ln \mathrm{K_4} \\
&\mathrm{lnK_2} + \ln \mathrm{K_1} = \ln \mathrm{K_3} + \ln \mathrm{K_4} = \ln(\mathrm{K_2\ K_1}) = \ln(\mathrm{K_3\ K_4}) \text{ or} \\
&\mathrm{K_2\ K_1} = \mathrm{K_3\ K_4}
\end{aligned} \tag{4.7.9.1}$$

For NMA transferring from water to CCl_4, ΔG^o_4 = + 0.62 kcal/mol (+2.6 kJ/mol).

(Note: Biochemists like to talk about thermodynamic cycles, which may seem new. However, believe it or not, you have seen them before in introductory chemistry in the form of Hess's Law!)

From K_{1-4} and the corresponding ΔG^o values, we can now calculate ΔG^o_5 for forming H-bonded NMA dimers in a nonpolar solvent from two molecules of NMA(aq). This reaction, which we hope simulates the formation of buried intrachain H bonds in proteins on protein folding, is:

$$Dw + Aw \leftrightarrow (DA)n, \text{ for which } \Delta G^o_5 = +3.72 \text{ (i.e. disfavored).}$$

If this model is a good mimic for studying the formation of H bonds during protein folding, then one would conclude that the formation of buried H bonds does **not** drive protein folding.

However, suppose the transfer of D and A (from a large protein) from water to the nonpolar medium (modeled by K_2) is driven by other effects (such as the hydrophobic effect). Then, the negative value of ΔG^o_1 will significantly favor the formation of buried H-bonds. So, if this happens in proteins, it is clear why so many intrachain H bonds form since K_1 is favored. H bonds may not facilitate the collapse of a protein, but they do favor internal organization within a compact protein. That is, H bonds don't drive protein folding per se, but form so that too many unsatisfied H bonds would not destabilize the folded protein.

This simple model has potential problems. The interior of a protein is not homogeneous (i.e., the effective dielectric within the protein will vary). H-bond strength is also very sensitive to geometry. There are many H bonds within a protein, so slight errors in estimating H bond strength would lead to large errors in determining protein stability.

Another argument against H-bonds being the determining factor in protein folding and stability comes from solvent denaturation studies. If intrachain H bonds are so important, should solvents that bind to the backbone not denature the protein? Shouldn't water (55 M) act as a denaturant? It doesn't, however. Dioxane (a 5-membered heterocyclic ring with an O) has only a H-bond acceptor and wouldn't be expected to denature proteins, yet it does. H bonds also form more readily in nonpolar solvents. Peptides, which have random structures in water, can be induced to form helices when placed in alcohol solutions (trifluoroethanol, for example), which are more nonpolar than water, as explained above in the helix-coil studies. If H bonds are the dominant protein stability factor, then alcohols stabilize proteins. At low alcohol concentrations, proteins are destabilized.

4.7.9.4: Hydrophobic Interactions: Introduction

We have studied the role of the hydrophobic effect (the favorable entropic release of caged water molecules about solvent-exposed hydrophobic groups) in driving the solubility of nonpolar molecules in Chapter 2.5. Does this also drive protein folding? To explore this question, we will study the thermodynamics of small nonpolar molecules, especially benzene, with water and ask whether the thermodynamic parameters associated with benzene solubility are similar to those associated with protein stability. If this analogy holds, anything that promotes benzene solubility in water will increase the exposure of hydrophobic amino acid side chains to water and promote protein denaturation. What is the evidence to support this?

a. Crystal structures: PDB crystal structures show that most nonpolar side chains are buried inside a protein, which is tightly packed and excludes water. Studies show that as the surface area of amino acid side chains increases, the free energy of transfer of amino acids from water to ethanol becomes more negative, as shown in Figure 4.7.9.5.

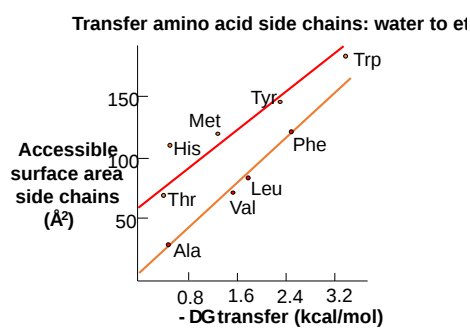


Figure 4.7.9.5: Transfer free energy of amino acids from water to ethanol

b. Low-temperature denaturation of proteins - It has been observed that proteins can denature at low temperatures (less than 0 ^{o}C), suggesting that nonpolar residues become more "soluble" in water at low temperatures (i.e., they move from the more hydrophobic interior of a protein to the more polar outside). Compare the solubility of nonpolar gases like CO_2 or N_2, which are more soluble at low temperatures. As you heat solutions of nonpolar gases in water, the gases become less soluble, as evidenced by bubble formation (i.e., phase separation of dissolved gases as they become more insoluble, as you have observed with a bottle of soda). If protein behavior is governed by this same behavior (greater solubility of nonpolar groups at low temperatures), it would suggest that proteins might denature at low temperatures (leading to increased exposure to water of the nonpolar side chains). This phenomenon has been observed.

c. Protein stability affected by different salt species - Over 100 years ago, Hofmeister determined the effectiveness of different cations and anions of salts to precipitate blood serum proteins in the 0.01 - 1 M concentration ranges. The series is shown below:

$$\text{Cations: } NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+} > \text{guanidinium}$$

$$\text{Anions: } SO_4^{2-} > HPO_4^{2-} > \text{acetate} > \text{citrate} > Cl^- > NO_3^- > ClO_3^- > I^- > ClO_4^- > SCN^-$$

- A salt of pairs of the first ions in these series, for example, $(NH_4)_2SO_4$, when added to aqueous solutions of proteins, precipitates the native form of the protein. We must account for the fact that it precipitates the protein in the native, not denatured state. More on why it precipitates proteins in a moment. The first ion in each series increases the surface tension of water (making it harder to make a cavity in the water to fit the nonpolar molecule). This decreases the solubility of nonpolar molecules. These "salt-out" nonpolar molecules promote dissolution in water, followed by aggregation and phase separation. By analogy, they will stabilize the native state, since the buried hydrophobic side chains would have a reduced tendency to move into the aqueous environment.
- The last ions in the series have a lesser effect on surface tension, increasing the **solubility** of nonpolar molecules ("salt-in"). By analogy, they will destabilize the native state since the buried hydrophobic side chains would have an increased propensity to move out into the aqueous environment.

The Hofmeister Series and its effects on the chemical properties of water and solutes are shown in Figure **4.7.9.6**.

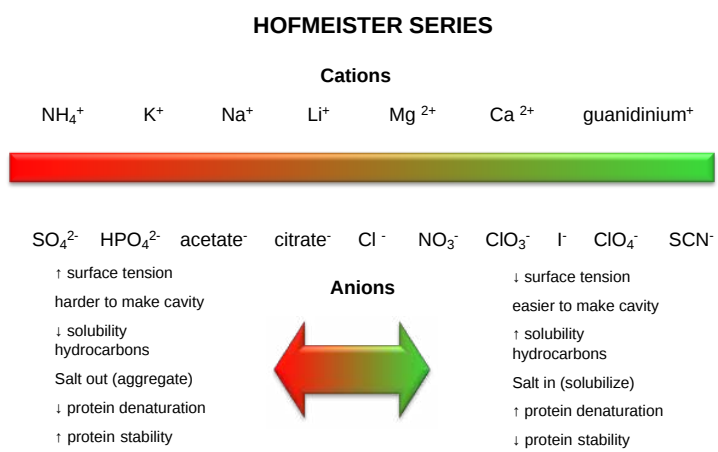


Figure **4.7.9.6**: Hofmeister Series

The solubility of benzene in aqueous salt solutions of this series increases from left to right, just as native protein stability decreases from left to right (i.e. the protein's nonpolar core residues become more "soluble" in water, leading to its denaturation).

d. Conservation of hydrophobic core residues - These residues are highly conserved and correlated with structure.

e. Urea denatures proteins - Another additive, urea ($H_2N(C{=}O)NH_2$), at high concentrations, is often used to denature proteins. People used to think that urea competed with the intrachain H bonds and hence unraveled the protein. The arguments above with H-bonds dispute this contention since water should denature the protein. How does urea denature proteins? It has been shown that the free energy of transfer of nonpolar amino acids into 8 M urea becomes increasingly negative as the side chains become larger and more nonpolar. This is also true for guanidine hydrochloride-induced denaturation. Both also increase the solubility of nonpolar molecules, proportional to their surface area. The structure of urea and guanidinium, along with the side chains of arginine and the Type II diabetes drug metformin, are shown in Figure **4.7.9.7**.

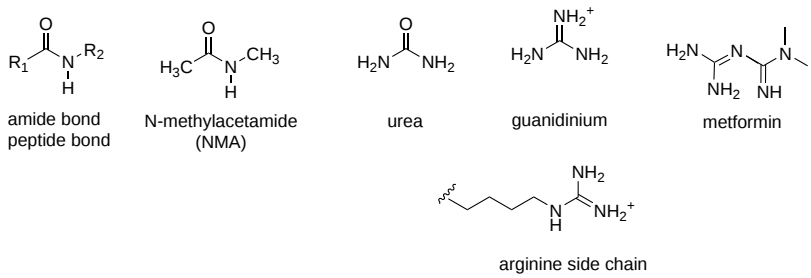


Figure 4.7.9.7 : Structures of urea, guanidine HCl, arginine, and metformin

Figure 4.7.9.8 shows the free energy of transfer of the nonpolar amino acids into 8 M urea and 6 M guanidine HCl as a function of the surface area of amino acid side chains.

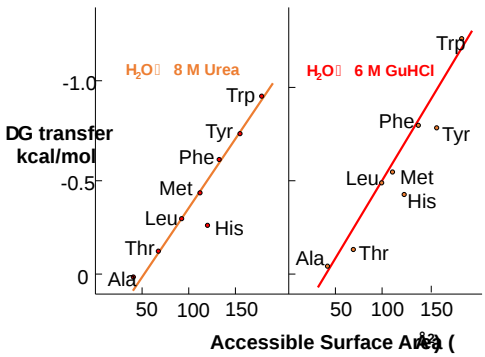


Figure 4.7.9.8: Free energy of transfer of the nonpolar amino acids into 8 M urea and 6 M guanidine HCl

4.7.9.5: Additives and Their Interactions with Protein Surfaces

Additives that stabilize the folded state of proteins often decrease their solubility. These additives are excluded from the protein's preferential hydration sphere (negative binding). Denaturants, in contrast, tend to increase protein solubility and interact preferentially with the protein surface. In their presence, proteins respond by increasing their surface area through denaturation. For stabilizers, proteins try to minimize their surface area by staying "native" and aggregating to form a precipitate, both of which minimize the surface area from which the stabilizing agent is excluded.

The main effect of dissolved ions on water structure has been thought to involve changes in H bonds (either enhancers/structure makers or inhibitors/structure breakers), which correlate with the salting-in or salting-out effects of various ions. Many techniques have been used to study these interactions:

- viscosity: inferential information on the structure
- diffraction (x-rays/neutrons): gives information on the coordination number of the solvation shell (static information)
- NMR: information on the average relaxation of bulk and hydration sphere water around ions (dynamic information)
- molecular dynamics simulations give insight into short but not long-range interactions between ions and water.

Recent studies have provided conflicting support for structure makers/breakers. Omta et al. (2003) used femtosecond mid-infrared pump-probe spectroscopy to study actual H-bonds between water molecules in salt solutions ($Mg(ClO_4)_2$, $NaClO_4$, and Na_2SO_4). In pump-probe spectroscopy, a sample is excited with a short pulse (pump) and, after a short time lag, with another pulse (probe) interacting with the excited state. Linearly polarized infrared pulses (pump) were used to excite OH groups in solution, followed by a probe pulse polarized at 45 degrees relative to the pump pulse. Only those excited OH groups that had rotated in the time interval between the pump and probe would be excited by the probe. Using this technique, the time frame for reorienting the OH groups, which is related to the "stiffness" of the H-bonds, can be determined. The salts did not affect the rotational motion of bulk water outside the first hydration shell, suggesting that they have no effect on the H-bond networks in bulk water. Mg^{2+} ions are considered structure makers, as they greatly increase water's viscosity, supposedly due to increased H-bonding among water molecules. This study does not support this model. The increased viscosity of Mg solutions must be attributed to ions directly interacting with water molecules. The solution can be modeled as bulk water with small rigid spheres representing the ion and its

first hydration sphere. Much more experimental and theoretical work must be performed to gain structural insight into the role of salts on water structure. Until then, we will try to understand the effects of different salts on water structure in descriptive terms and using thermodynamic quantities.

Studies have shown that urea preferentially binds to protein surfaces, increasing their surface area and hydrophobic exposure and denaturing proteins. However, the figure below shows that glycerol, a bigger polar but uncharged molecule, stabilizes the native state. This pair of uncharged additives similarly affects protein stability, as does the charged guanidine HCl/ammonium sulfate pair.

Figure 4.7.9.9 shows how solution additives might interact with the protein's surface.

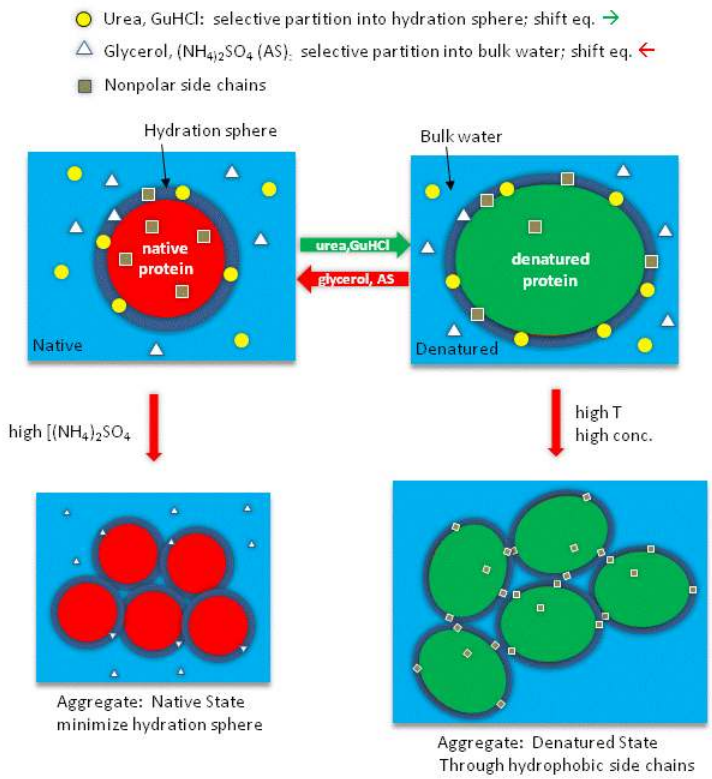


Figure 4.7.9.9: Interaction of solution additives with the protein surface.

Figure 4.7.9.10 shows a thermodynamic cycle for urea denaturation of proteins

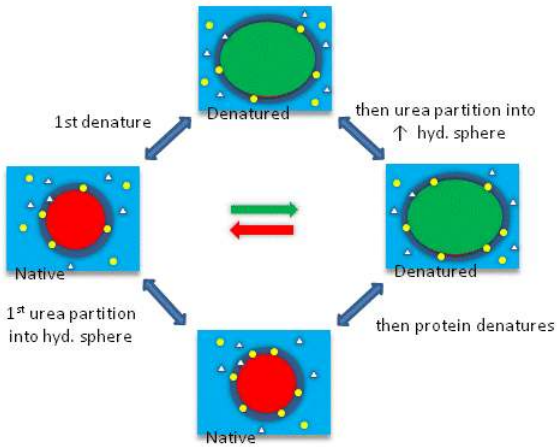


Figure 4.7.9.10: Thermodynamic cycle for urea denaturation of proteins

Figure 4.7.9.11 summarizes surface changes resulting from protein denaturation.

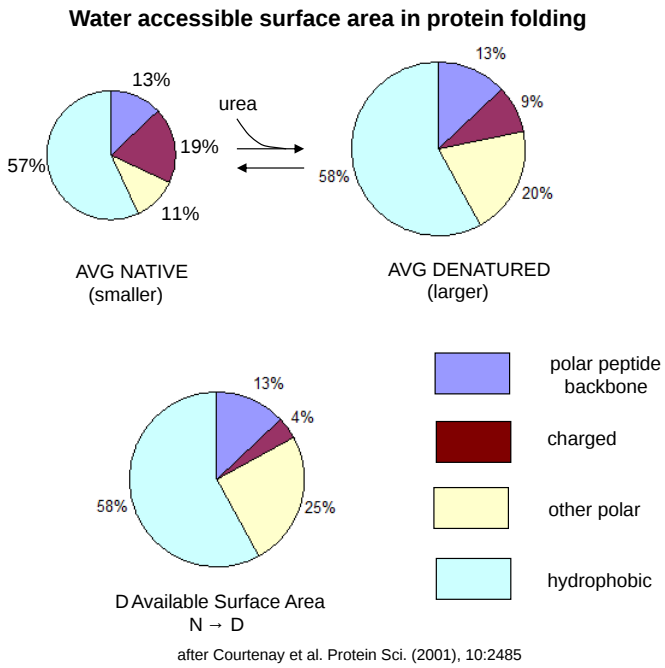


Figure 4.7.9.11: Water Accessible Surface Area in Protein Unfolding

4.7.9.6: The Hydrophobic Effect and Change in Heat Capacity

Our understanding of hydrophobic interactions has changed dramatically in the last several years. This is not reflected in most textbooks. The hydrophobic effect means different things to different people. Some refer to the transfer of nonpolar solutes/solvents to aqueous solution. Some refer to the same phenomenon only if the effects have a unique temperature dependency. Others refer to the ordering of water around nonpolar residues. The most recent explanation centers around the unique temperature dependencies of the transfer reactions. Before we can understand it, here is an interesting bit of data. If you dissolve one mole of methane in hexane, the volume of 1 L of hexane changes to 960 mL. However, if done in water, the water volume changes to 937 ml, indicating that water molecules pack more efficiently around nonpolar molecules than in their absence.

Let's now consider the thermodynamic aspects of the hydrophobic effect, as we did for micelle and bilayer formation. In summary, we found that the free energy of transfer of a hydrocarbon or alcohol from an aqueous solution to the pure hydrocarbon or alcohol, for example, was disfavored enthalpically (unexpectedly) but favored entropically (also unexpectedly until we included solvent in our model). These experiments were conducted at a single temperature, providing our first understanding of the hydrophobic effect. We will expand on this view by looking at the enthalpic and entropic contributions to the transfer of benzene into water as a function of temperature. This will lead us to a more modern view of the hydrophobic effect. To do this, we must consider the thermodynamics of mixing two substances with different properties, such as polarity, which affects solubility.

If you mix two substances A and B that aren't very soluble in each other, two opposing thermodynamic factors are relevant.

- An increase in entropy drives the tendency to mix.
- The mixing is usually opposed by enthalpy.

The latter makes "intuitive" sense since you might expect that van der Waal interactions between A-B might be less than those of A-A and B-B (i.e., the adage "like dissolves like"). If AA and BB self-interactions are stronger, A will not dissolve in B and vice/versa. You would also expect no significant changes in entropy and enthalpy as a function of temperature in this ideal mixing.

The most modern understanding of hydrophobic interactions involves mixing A and B, which is characterized by a unique temperature dependence of the changes in entropy and enthalpy. At room temperature, if one corrects the entropy changes for effects due solely to mixing, the "excess" entropy is what principally opposes the tendency for a nonpolar molecule to dissolve in water. Enthalpy changes are small. We have modeled this effect using structured water around the nonpolar residues. We will further our understanding of the hydrophobic effect by using benzene transfer from water as a model system (much as we did with N-methylacetamide for H bonding).

Before discussing entropy and enthalpy changes accompanying protein folding/unfolding, let's learn about the thermodynamic aspects governing benzene solubility in water. What happens to benzene solubility in water and the corresponding thermodynamic parameters as you raise the temperature? The graph in Figure **4.7.9.12** shows the change in G, H, and -TS for taking benzene from pure benzene to water, $B_B \rightarrow B_W$. This is real data. Note that the graph shows -TΔS, not just ΔS. When the values of each of these terms, ΔG, ΔH, and -TΔS, are negative, the transfer of benzene into water is favored. The blue highlighted region shows more values in a more physiologically relevant temperature range (0-100^0 C).

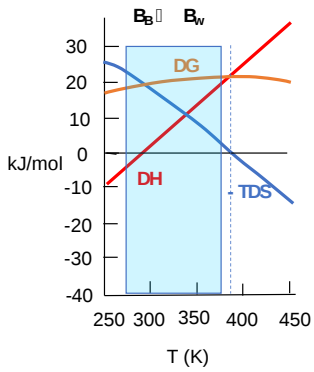
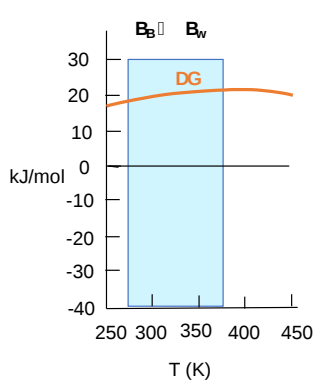


Figure **4.7.9.12**: ΔG, ΔH and -TΔS for transfer of benzene from pure benzene to water.

Two things to note for right now. ΔG for $B_B \rightarrow B_W$ is always positive and hence always disfavored. Another key feature is shown in the vertical dotted line at the temperature at which benzene has its greatest "aversion" to being in the water, where the slope of the ΔG vs T curve (dΔG/dT) = 0. At this point, entropy plays no role in benzene solubility in water. This goes against the explanation of the hydrophobic effect we discussed in Chapter 2.5. This maximum aversion to water is completely driven by enthalpy.

If these axes showed three lines describing the number of ducks, geese, and lunes in a lake as a function of lake temperature, no further explanations might be needed. However, many years of teaching experience and research show that thermodynamic parameters are threshold concepts that students have struggled with for a long time. A graph showing three simultaneously might be uninterpretable to many. So, let's deconstruct this graph into a series of stepwise graphs with targeted explanations to facilitate your understanding. These graphs with explanations are shown below.

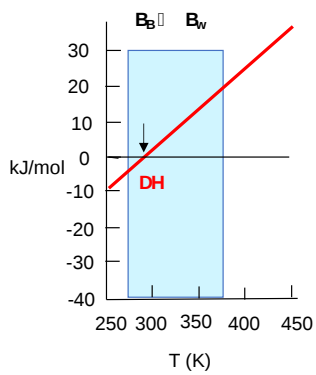
A graph of ΔG vs T for $B_B \rightarrow B_W$



ΔG for $B_B \rightarrow B_W$ is positive and **disfavored** over the entire temperature range. However, note several important features:

- There is a temperature (around 375^{0}K) when ΔG is at a maximum (i.e., in the language of math, the slope of the curve ΔG/ΔT or dΔG/dT is a maximum.
- As the temperature decreases from T_{max}, the transfer of benzene becomes less disfavored. It would be favorable if you extrapolated the curve to very low, but unrealistic, temperatures.

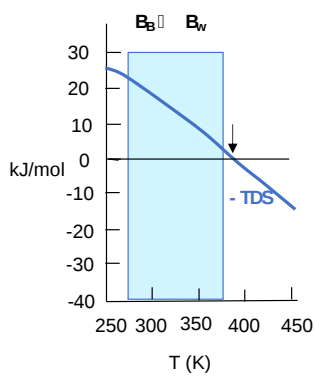
A graph of ΔH vs T for $B_B \rightarrow B_W$



The ΔH vs. T for or $B_B \rightarrow B_W$ graph is linear with a positive slope. It crosses the T axis at the point indicated by the arrow.

- At temperatures below the point indicated by the arrow, the transfer of $B_B \rightarrow B_W$ is **enthalpically** favored. We saw this unexpected finding in Chapter 2.5 while considering the transfer of butanol, for example, from pure liquid alcohol to water at just one temperature, 25 ^{0}C = 298 K. That data is consistent with this curve.
- At temperatures above the point indicated by the arrow, the transfer of $B_B \rightarrow B_W$ is increasingly disfavored.
- The slope of this line, ΔH/ΔT, is expressed in kJ/mol/K, which is the heat capacity of the system.

A graph of -TΔS vs T for $B_B \rightarrow B_W$



The graph of -TΔS vs T for $B_B \rightarrow B_W$ is almost linear with a negative slope, which shows that it becomes less entropically disfavored. The curve crosses the T-axis at the point shown by the arrow.

- At temperatures below the point indicated by the arrow, the transfer of $B_B \rightarrow B_W$ is entropically **disfavored.** We saw this in Chapter 2.5 for the transfer of alcohols from the pure liquid into water. We called this the **"hydrophobic effect,"** the main thermodynamic drive to move organic molecules **out of water.** We attribute this to the formation of a cavity in water for the nonpolar parts of the alcohol, leading to a more ordered set of H bonds than in bulk water and reducing the number of microstates available to the associated water molecules.
- However, at the temperature indicated by the arrow, there is **no entropic barrie**r to move benzene into water. Above those temperatures, it is favored. This seems to destroy our simple definition of the hydrophobic effect.

Let's add more to our molecular interpretation of this seemingly anomalous favorable entropy at higher temperatures. As the temperature is raised, the available positional and thermal entropy of water increases significantly. It would seem logical that putting a nonpolar residue into this water system would become easier than putting it into more structured water (characterized by fewer microstates and lower positional and thermal entropy) at a lower temperature! (Remember from our thermodynamics review that if the T_{surr} is high, a given heat transfer to or from the surroundings will have a smaller effect on the ΔS_{surr}. Conversely, if the T_{surr} is low, the effect on ΔS_{surr} will be greater. In a recent General Chemistry, Atkins uses the analogy of a sneeze in a library compared to one in a crowded street. An American Chemistry General Chemistry text uses the analogy of giving $5 to a friend with $1000, compared to giving $5 to someone with just $10.

But look at the other temperature anomaly. From an enthalpic perspective, it becomes increasingly difficult to dissolve benzene in water. At high temperature, -TΔS becomes zero, and there is no entropic barrier to dissolving benzene in water. The barrier is completely enthalpic. This is why a more nuanced definition of the hydrophobic effect has emerged.

If you sum ΔH and -TΔS at each temperature, you get the curve for the total ΔG to take benzene from pure benzene to water. Notice that it is always positive, so it is always disfavored. The ΔG function is curved. It increases at low temperatures and decreases at very high temperatures, implying that there is a temperature at which benzene in water has minimum solubility (a

maximum in ΔG). The minimum benzene solubility (the max. positive ΔG) occurs when dΔG/dT = 0. This occurs when ΔS = 0, and the maximal aversion of benzene to water is entirely enthalpic, a statement not consistent with our initial understanding of the hydrophobic effect.

For those who have taken physical chemistry ...

For just PV work, dG = VdP - SdT, so that dG/dT = V dP/dT - S. At the maximum of the curve of G vs T, dG/dT = 0, so

V dP/dT = S. For this system, dP/dT = 0, so S = 0. This is observed in the top graph above.

Here is another equation for heat capacity that you derived in physical chemistry:

Cp = dH/dT = TdS/dT.

The last equality stems from Maxwell's relationships, which Physical Chemistry students should remember.

Let's review some more thermodynamics. (Molar) heat capacity, C_p, is the heat required to raise the temperature of a mole of a substance 1 ^oC. It has units of kJ (or kcal)/mol. The slope of the enthalpy curve ΔH vs. T for $B_B \rightarrow B_W$ has units of kJ/mol/K and is the **heat capacity,** Cp, where Cp = dH/dT.

To refresh your mind and demystify the term heat capacity, let's look at the heat absorbed as ice is converted to gas as a function of temperature (i.e., the phase transition for water). Figure **4.7.9.13** a graph of the heat absorbed, Q_{abs}/gram vs T (top), and a derivative plot of the top graph (slope of the top graph = heat capacity at each point on the T axis).

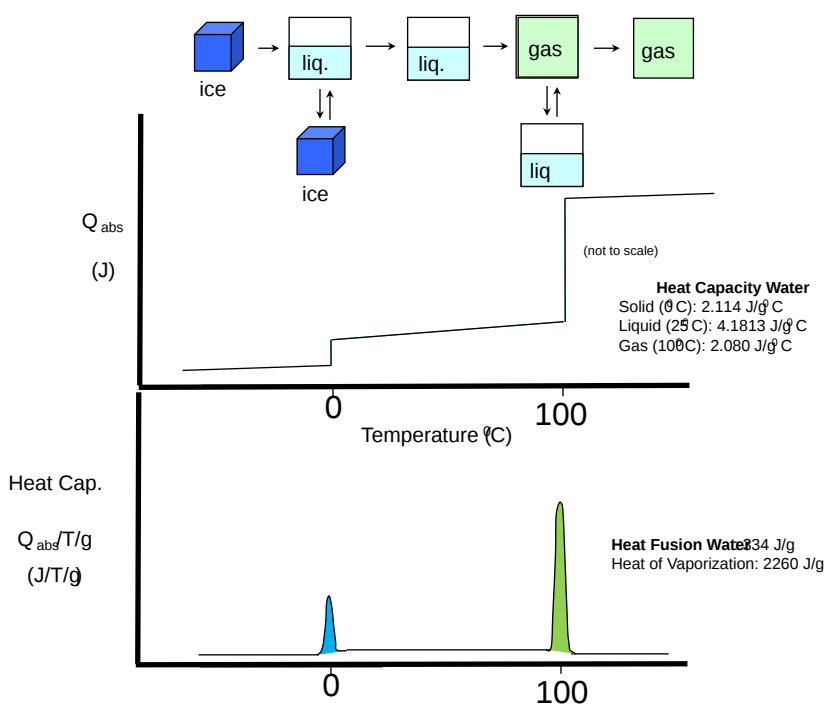


Figure **4.7.9.13**: Q vs T and Cp vs T plots for phase transitions in water.

The graphs for benzene transfer into water now teach us one final concept. The ΔH vs. T graph is linear for the transfer of benzene from the pure liquid to water. We now know that the slope of that graph, ΔCp, is a constant positive number for the process when a hydrophobe is exposed to water.

This positive C_p observed when a hydrophobic group is transferred into water is the signature of the hydrophobic effect.

A positive Cp occurs when both H and S depend on temperature, as observed when a hydrophobe is transferred from a less polar environment to water. Likewise, a negative Cp is observed when hydrophobes in water are transferred to a more nonpolar environment.

4.7.9.7: Hydrophobic Effect Applied to Proteins

Now, let's return to the world of proteins and see how we might apply what we learned from the transfer of benzene from the pure liquid to water. An analogous transfer experiment would be the denaturation of proteins in which buried hydrophobic amino acids are "transferred" to water. Figure **4.7.9.14** shows the derivative plot of heat absorbed vs. T (i.e., a plot of heat capacity Cp vs T) for the thermal denaturation of a protein obtained using differential scanning calorimetry.

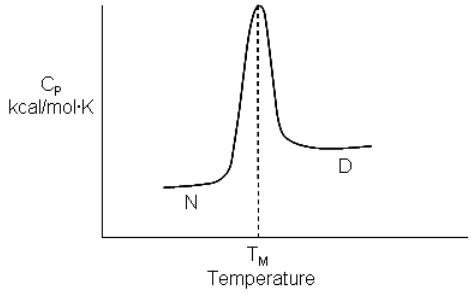
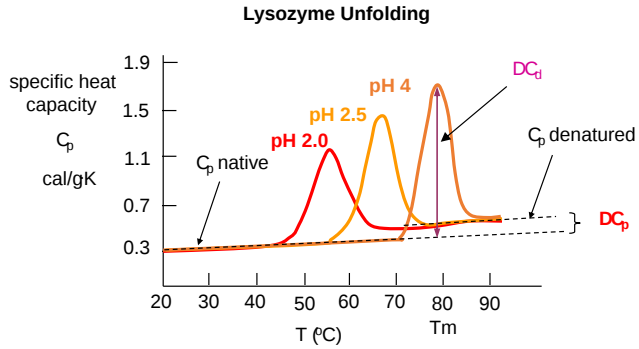


Figure **4.7.9.14**: Cp vs T for the thermal denaturation of a protein

Look at the graph in Figure **4.7.9.15**, which shows the heat capacity of the protein lysozyme at various pHs vs temperature obtained using differential scanning calorimetry.



Privalov and Khechinashvilli. J. Mol. Biol. 86:665 (1974)

Figure **4.7.9.15**: Heat capacity vs T for lysozyme at various pHs

As the protein is heated, it reaches a temperature where a large amount of heat is suddenly absorbed as it unfolds. The area under the curve represents the heat absorbed on denaturation (units of cal/g·K x K = cal/g). The midpoint temperature is the protein's Tm. (Why would the Tm depend on the solution's pH?) Notice that there are two ΔCps shown in the graph. One, $\Delta \mathbf{C}_d$, is associated with the actual denaturation process and is analogous to the change in heat capacity observed during phase changes, such as the solid-to-liquid transition of water. The other is ΔC_p, the difference in heat capacity between the denatured and native states. As was the case for the transfer of benzene to water, the ΔC_p for protein denaturation is also positive, suggesting that hydrophobes are transferred from the protein's interior to water in protein denaturation.

Figure **4.7.9.16** attempts to give a molecular description that applies to different regions of the denaturation curve and helps explain the positive ΔC_p observed in thermal denaturation curves of proteins.

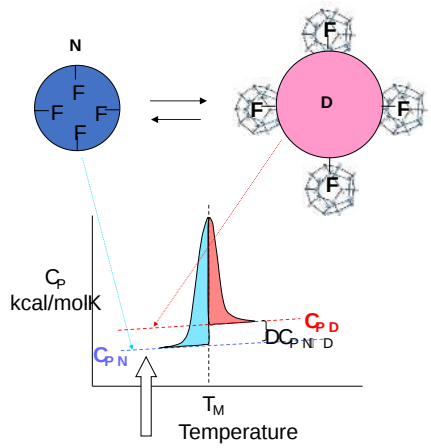


Figure 4.7.9.16: Molecular Descriptors associated with temperature denaturation of proteins

What is the molecular basis for this large heat capacity transfer change for benzene? One can show that the Cp is also proportional to the surface area of the nonpolar solute. Figure 4.7.9.16 shows the smaller, more compact native state, with buried Phe (F) side chains denaturing and moving to the more open D state with exposed F side chains. Since these are nonpolar, we can envision a "clathrate" or cage of ordered water around them. The heat capacity curves for both the native and denatured states are extrapolated into the region where T < Tm (even though there is very little denatured state in that region).

The "caged" water around the exposed F in the D state is low-energy due to the "ice-like" Hydrogen-Bond network. More heat would be absorbed (as the temperature is increased) to break up that cage compared to the same amount of heat applied to the N state. Hence, C_p D > C_p N. At room temperature, water molecules surrounding the nonpolar residue are low in energy (lots of H bonds) and low in entropy (thermal and positional, fewer available microstates). As the temperature rises, water populates higher-energy states (fewer H bonds) and higher entropy (thermal and positional; more available microstates). The increase in temperature causes the "melting" of the surrounding water structure in terms of energy and entropy. The two different energetic states of water provide an energy storage mechanism.

Here is another explanation that might help. Water molecules form an "iceberg-like" cage, or clathrate, around nonpolar molecules. The water is fully H-bonded (to itself, not to the nonpolar molecule) in a fashion analogous to ice, but the H bonds' geometry is nonideal. This structuring of water decreases its entropy. With increasing temperature, the structured water "melts," which produces the large heat capacity of a solution of a nonpolar molecule in water, just as the actual melting of ice showed a large heat capacity. This large heat capacity is the signature thermodynamic feature of the solution of a nonpolar molecule in water.

The following equations can be derived for the actual part of the graph in which denaturation occurs (N ↔ D).

$$\begin{gathered} C_p = C_d = \frac{dH}{dt} \\ dH_d = C_d dt \\ \int_{T_1}^{T_s} dH_d = < \Delta H_d > = \int_{T_1}^{T_2} C_d dT \end{gathered} \tag{4.7.9.2}$$

to help calculate the calorimetric enthalpy changes ($\Delta H_{cal} = \Delta H_d$) of denaturation from differential scanning calorimetry.

where $<\Delta H_d>$ is the average enthalpy change that occurs on denaturation, represented by the blue and red shaded areas in the curve above (between T1 and T2). Compare this to the van 't Hoff enthalpy discussed in Laboratory Determination of ΔG^o for Protein Folding/Unfolding.

For the calorimetric determination of ΔCd and ΔHd, the values are temperature dependent, and no assumption about the temperature independence needs to be made, as in the case of the ΔH_{vHoff} plot.

Water Clathrate Cages - An Extension

For more details on these "cages," click the link below.

Answer

Is there a more quantitative description of the ordered water than a cage? Sharp et al investigated the "structure" of the caged water around nonpolar and polar molecules in a theoretical analysis supported by molecular dynamics (Monte Carlo) simulations. The average bond angles and lengths of water-water H bonds in the first hydration sphere around a nonpolar molecule, such as benzene, decreased. In contrast, they increased around polar or charged species, such as potassium ions. The average changes observed arose from two types of H bonds compared to bulk water: shorter, more linear, and longer, more bent.

A review by Silverstein suggests that an immobile clathrate cage is **not** a good representation of water surrounding a hydrophobe. Although we like to envision molecular models that allow us to "explain" experimental thermodynamic findings, such models should be rigorously tested. An alternative explanation hinges on water's small size (compared to other solvents), tight packing, and high density. Consider the density of water compared to that of more nonpolar liquid solvents, as seen in the table below.

Table: Density of common solvents

Solvent	Volume (Emin) (A3) (Spartan)	Density (g/ml)
H_2O	19.17	1.00
methanol CH_3OH	40.66	0.791
ethanol CH_3CH_2OH	59.08	0.789
n-propanol $CH_3CH_2CH_2OH$	75.35	0.804
n-butanol $CH_3CH_2CH_2CH_2OH$	95.68	0.810
hexane	124.8	0.654

Let's consider the density of water surrounding an exposed nonpolar. If we envision the surrounding water as a clathrate, we might assume it is "ice-like". So what are the physical properties of ice and water that might give us insight into the water surrounding a nonpolar molecule?

Ice, of course, has a lower density than liquid water. The number of H bonds can't explain this, since experimental evidence shows that ice has an average of 4 H bonds per water molecule, whereas liquid water has an average of 2.4. Experimental data also show that, to accommodate water molecules within a rigid network of interacting waters with tetrahedral symmetry, the H-O-H bond angle increases from 104.5 to 106. With fewer packing constraints, liquid water molecules can self-organize to maximize packing and, hence, macroscopic density. Studies suggest that 10 water molecules solvate a buried methyl group, and infrared studies show that 4 of them exhibit significant rotational diffusion barriers, indicating they are effectively immobilized and hence "ice-like." Silverstein suggests that the water surrounding a nonpolar group in solution should be considered dynamically, with some water immobilized (as in ice) and the remaining water more fluid-like.

Now, let's review the benzene graph and apply it to protein unfolding one more time. The graph for $B_B \rightarrow B_W$ (Figure **4.7.9.12**) shows a maximum in benzene insolubility. As the temperature decreases from its maximum, benzene becomes more soluble in water. Conversely, as temperatures rise, the solubility of benzene decreases (just like nonpolar gases become increasingly insoluble with increasing temperature). If you extrapolate the ΔG curve in this range of decreasing temperature past the range shown on the graph, it would cross the X-axis and become <0, implying benzene would be favored to dissolve in water. Does the low-temperature behavior of benzene/water interactions (becoming more soluble as the temperature is decreased from the maximum

temperature for its insolubility) extend to and predict protein behavior at low temperatures? Figure 4.7.9.17 shows the analogy between benzene solubility in water at cold temperatures and cold temperature denaturation of proteins.

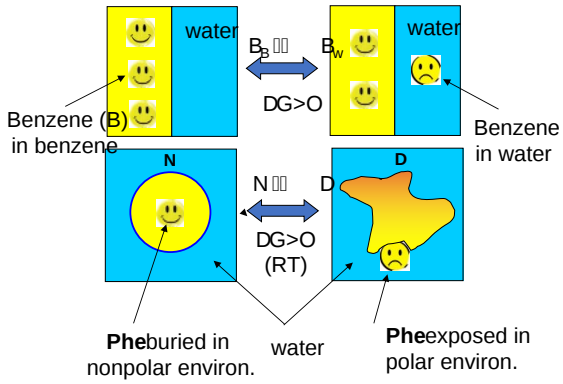


Figure 4.7.9.17: An analogy between benzene solubility in water at cold temperatures and cold temperature denaturation of proteins

In the figure, F stands for a Phe side chain, which can be buried, sequestered from water as it would be in the native state of the protein, and exposed to water, as it might be in the denatured state.) The answer to our question is yes, at low temperatures. The analogy to benzene being more soluble at low temperatures is the hydrophobic side chains in a protein becoming more likely to flip into water, denaturing the protein. The low-temperature behavior would predict low-temperature protein denaturation. This phenomenon has been observed. Note that it doesn't require a state change when the nonpolar side chains prefer water; just a shift in that direction might be enough to tip the balance and lead to denaturation of the marginally stable protein.

Please note that we are attempting to extrapolate the thermodynamic parameters associated with benzene solubility in water to the denaturation of a protein, **NOT TO THE SOLUBILITY OF A PROTEIN**!

What about high temperature? Proteins denature as the temperature increases in the range where the ΔG curve for benzene reaches a peak. Suppose the buried hydrophobic residues behave like benzene. In that case, they would increasingly "prefer" **not to** flip out into the water as the temperature rises to the maximum in the ΔG vs. T curve. Hence, the benzene predicts that the protein should become **more** stable. What, then, explains the observed denaturation of proteins at high temperatures? Another factor must account for it. What is it?

As the temperature increases, more protein conformational states become available and are occupied. At low temperatures, let's say that there is only one native state available and occupied, and (to pick a number) maybe 10 non-native states that are energetically accessible. At high temperatures, there is still only one native state, but perhaps 1000s of accessible nonnative states. More accurately, think of the protein existing in an ensemble of conformations. As the temperature increases, more non-native states can be populated than at lower temperatures, leading to an entropic driving force favoring unfolding. Which way would this **chain conformational entropy** drive the protein at high temperature? It would be driven to the most number of states - to the denatured state. Hence, a modern definition of the hydrophobic effect can explain low-temperature denaturation, but not high-temperature denaturation.

4.7.9.8: Summary of studies from small molecules (N-methyacetamide and benzene)

It is clear that proteins are not particularly stable, and many contributions of varying magnitudes must sum to give them their marginal stability under physiological conditions. The hydrophobic effect plays a major role in protein stability. Also, since proteins are so highly packed compared to a loosely packed denatured state, collective Van der Waals interactions must play a significant part. Remember that these interactions, especially induced dipole-induced dipole interactions, are short-range and become most significant under conditions of closest packing. Opposing folding is the chain conformational entropy just described. Since proteins are so marginally stable, even one unpaired, buried ionic side chain or 1-2 unpaired, buried H-bond donors and acceptors in the protein may be enough to "unravel" the native structure, leading to the denatured state.

4.7.9.9: Mutagenesis and Protein Stability

In the last decade, the contributions of the hydrophobic effect and H bonds to a protein's overall stability have been studied using site-directed mutagenesis. In this technique, the DNA coding sequence for a given amino acid in a gene can be altered so that the

resulting mutant protein differs from the normal (wild-type) protein by a single amino acid. To probe the hydrophobic effect, for example, a buried hydrophobic amino acid such as Ile could be replaced with Gly, which is much smaller and makes a lower hydrophobic contribution to the stability of the native state. The result of this mutation might leave a "hole" in the protein (not unlike the vacant holes in crystal structures of salts). This "hole" might be diminished in size by a subtle rearrangement of the nearby protein structure. Certain amino acids would not be used as replacements in such studies. For instance, an Ile would not be replaced by a positively charged Arg, as this would destabilize the protein. The extent of destabilization in mutant proteins can be determined by calculating the ΔG^o for the native-to-denatured transition using urea as the denaturing agent, as discussed in another section.

Previously, the following statistics were presented concerning the distribution of amino acids in a protein's tertiary structure. New values, shown below in red, are based on much more crystallographic data, as summarized in Pace's article.

- The location of side chains varies with polarity. Nonpolar side chains, such as Val, Leu, Ile, Met, and Phe, are nearly always (83%) in the protein's interior.
- Charged polar side chains are almost invariably on the surface of the protein. (54% of Asp, Glu, His, Arg, and Lys are buried away from water, which is startling!)
- Uncharged polar groups such as Ser, Thr, Asn, Gln, Tyr, and Trp are usually on the surface but frequently in the interior. If they are inside, they are almost always H-bonded (63% buried - Asn, Gln, Ser, Thr, Tyr, again startling).
- Globular proteins are quite compact, with water excluded. Like the NaCl crystal, their packing density (Vvdw/Vtot) is about 0.75, equal to the closest packing density of 0.74. Compare this to organic liquids, which have densities of about 0.6-0.7.

Two articles by Pace suggest that Dill's "influential review (from which much of the above derives) that concluded that hydrophobicity is the dominant force in protein folding" should be rethought concerning the effect of hydrogen bonds. Using site-directed mutagenesis to change Asn (which can H-bond through its side chain) to Ala (which can't) in a variety of proteins, he has shown that approximately 80 cal/mol/A^3 of stability is gained if a side chain (in this case Asn) can form buried H-bonds to buried amide links of the protein backbone. Similar studies of mutants replacing Leu with Ala and Ile with Val suggest that only 50 cal/mol/A3 is gained from burying a hydrophobic -CH2- (methylene) group. Extending these results to protein folding suggests that protein stability is determined more by the formation of buried H bonds than by the hydrophobic effect!

The investigators measured ΔG^o for the N ↔ D transition (presumably by varying the urea concentration and extrapolating the ΔG^o for unfolding to 0 M urea (see: 4.12: Appendix - Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation). For the reaction as written, $\Delta G^o_{unfolding} > 0$ at room temperature and 0 M urea. Since they are destabilized, the mutant protein would have a less positive value for $\Delta G^o_{unfolding}$ (They would also have a less negative value for folding since they are less stable). The difference in $\Delta G^o_{unfolding}$ between the wild type and mutant ($\Delta\Delta G$) is expressed as:

$\Delta\Delta G = \Delta G^o_{unfolding}$ wild-type - $\Delta G^o_{unfolding}$ mutant > 0

$\Delta\Delta G > 0$ since $\Delta G^o_{unfolding}$ wild-type > $\Delta G^o_{unfolding}$ mutant. The more positive the $\Delta\Delta G$, the more the mutant is destabilized compared to the wild type. The data for a series of mutants is shown below.

Analysis of Mutants: H Bonds in Protein Folding

mutation	ΔVol side chain (Å^3)	% buried	ΔΔG kcal/mol (kJ/mol) (destabilized)	ΔΔG cal/mol/Å^3 (J/mol/Å^3) (destabilized)
Asn to Ala	37.4	95	2.9 (12)	78 (326)
Leu to Ala	74.5	99	3.6 (15)	48 (200)
Ile to Val	25.8	100	1.3 (5.4)	50 (209)

What leads to protein stabilization/destabilization when Asn is changed to Ala?

One possible contributor to stability is the side chain conformational entropy. Since the Ala in the mutant would find itself in a larger "hole" and have greater freedom for motion, it would have more conformational entropy, stabilizing the mutant over the wild type. Hence, this effect can **NOT** explain the observed destabilization of the Asn to Ala mutant.

In the proteins he studied, only one of eight Asn to Ala mutations involved an Asn in a helix, so the average change could not be attributable to differences in helix propensities for the two amino acids.

In the mutants, assuming no rearrangement of the remaining side chains, there is an "unnecessary" and unoccupied 37 $Å^3$ cavity. The creation of this cavity is thermodynamically unfavorable (about 22 cal/mol/$Å^3$, based on values for hydrophobic mutations). If the same penalty were applied here, the Asn to Ala mutant would be destabilized by 0.8 kcal/mol (22 x 37.4). This is significantly less than the observed destabilization (2.9 kcal/mol, 12.1 kJ/mol), so this effect also could not account for the destabilization of the Asn to Ala mutants.

If compensatory changes were made to minimize the cavity size, this would only help stabilize the protein and could not account for the observed destabilization.

Possible Explanation of Destabilization of Asn to Ala Mutants

possible reasons	explanation	effect on the mutant	support observed destab. of mutant?
residue conformational entropy	Ala in a bigger hole: more freedom of motion; favored entropically	stabilize mutant	NO
free energy change excess cavity formation	energy penalty to make an unoccupied cavity approx. 0.8 kcal/mol (3.3 kJ/mol)	destabilize	yes but of insufficient size compared to the observed effect (2.9 kcal/mol, 12 kJ/mol)
free energy change protein conformational changes	rearrange protein to fill the cavity	stabilize mutant	NO

Hence, these alternative explanations for the mutant's destabilization can't account for the data, so we're left with the explanation that the native protein's stability over the mutant is accounted for by burying the amide group's H-bond donors and acceptors and by associated changes in van der Waals interactions.

Pace argues that burying the amide group of Asn is similar to burying the peptide bond of the main chain. Their sizes are very comparable. Free amide groups can form four H bonds, but peptide (amide) groups can only form three. Even if the value of 78 for the $\Delta\Delta G$ (cal/mol/$Å^3$) is adjusted for this, the new value of 62 is still larger than that for burying a methylene group. Analysis of 108 folded proteins has shown that hydrophobic groups contribute 118,200 $Å^3$ of buried volume, compared to 92,000 $Å^3$ for peptide groups. Multiplying these figures by 78 and 49 (from the above table) suggests that burying peptide groups contributes more to protein stability than burying hydrophobic groups.

Would electrostatic interactions of the buried peptide group with the surrounding environment destabilize a protein? Pace argues that this would be more than compensated for by favorable van der Waals interactions (short range) at the buried site. This can be illustrated by comparing the ΔG transfer of an amide from water to the vapor phase (11.2 kcal/mol, 47 kJ/mol) and from water to cyclohexane (7.6 kcal/mol, 32 kJ/mol). Transfer to the vapor is more unfavored (due to the desolvation required when it moves to the gas phase) than to cyclohexane, even though a cavity must be created in the cyclohexane (a process which would be unfavored entropically). Transfer to octanol is even more favored (1.4 kcal/mol, 5.9 kJ/mol), but all these values are still positive (disfavored). Similar experiments involving the transfer of a methylene group (-CH2-) are negative, given the hydrophobic effect and van der Waals interactions arising from close packing. These suggest that van der Waals interactions formed on burying an amide in any solvent are stabilizing. Now, consider the packing density of atoms for various substances:

Packing Densities

substance	packing density
water	0.36
cyclohexane	0.44
closest-packed spheres	0.71
protein interiors	0.75

This table shows that collective van der Waals interactions (short-range) will stabilize the protein interior more than the same groups in bulk water (or in the denatured state). Carbonyl groups are more polarizable than methylene groups, which should contribute to van der Waals interactions.

One other addition. It has been noted that Gly peptides are not very soluble in water. Even with the polar peptide bonds, the backbone appears to be solvophobic. If the backbone of any polymer can't interact well with the solvent - i.e., the solvent is "poor" - then the backbone interacts with itself, which drives collapse. If the backbone interacts well with a "good" solvent, it won't collapse as readily.

4.7.9.10: Protein Stability in Thermophilic Organisms

What kinds of modifications are made to the sequence of a protein as the temperatures at which the organism thrives increase? A study by Szilagyi and Zavodszky examined 93 structures of 25 proteins: 29 from organisms that live at elevated temperatures (thermophiles, >50 °C for optimal growth) and 64 at nominal temperatures (mesophiles). Here are their results:

- the number of H-bonds and secondary structure elements does not correlate with temperature, but the number of salt bridges does.
- in hyperthermophiles (>80 ^{o}C for optimal growth organisms) that thrive at very high temperatures (100 ^{o}C), few internal cavities were found.
- The surface had more polar residues in those that thrive at intermediate-high temperatures (45-80 °C).
- in general, there was an increase in weaker ion pairs (increased distance between the charged side chains) in the hot group, but increases in strong and weak ion-ion bonds in the very hot group.

Kashefi and Lovley reported the identification of a bacterium obtained from a hydrothermal vent in the northeastern Pacific Ocean. In a laboratory setting, the strain grew in water at 121 °C under high pressure. These are the same conditions used in autoclaves to produce sterile samples. Cell doubling occurred within 24 hours under these conditions. The authors suggest that this strain would be useful in determining the molecules and their properties necessary for such high-temperature growth.

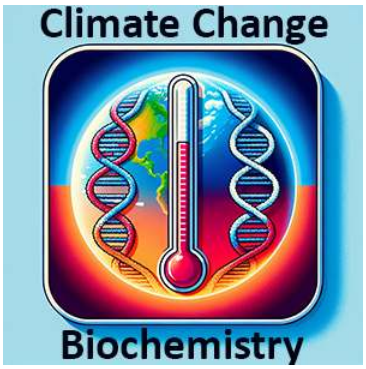
Using the computational program Rosetta Design, Korkegian et al. identified mutations in buried side chains of the homodimeric cytosine deaminase. Buried residues are presumably important in the stability of a protein and are targets for mutagenesis experiments that would increase the protein's melting temperature (Tm). In the program, a target sequence was "threaded" onto the template protein sequence (the wild-type protein), and changes were made to side chains in the random sequence. Energies were calculated, and those changes that resulted in lower energies were saved. Target residues (88) within 4 angstroms of the active site and the dimer interface were fixed to those in the wild-type template to minimize alterations in the deaminase's catalytic activity, which they chose to study. Remember, the study's goal was not to increase the enzyme's catalytic activity but its thermostability. The rest (65) were changed, and energies were calculated. 49% of the amino acids subjected to random change produced no change in the amino acid compared to the template (wild-type) side chain. Sixteen surface changes were ignored. Two sets of changes were observed, one involving amino acids packed between an alpha helix and beta strands, and the other set between two alpha helices. When prepared in the lab using recombinant DNA technology, these latter mutants were soluble at high protein concentrations and could be studied. Three different mutants (A23L, I140L, V108I) were generated, each increasing the T_M by about 2 degrees. However, a triple mutation had T_M values 10 degrees higher than the wild-type protein and a 30-fold longer half-life ($t_{1/2}$) at 50 degrees C. When the triple mutant was introduced into bacteria, the bacteria grew better at

higher temperatures. Crystal structures of both the wild-type and triple mutants show essentially an identical fold, with about 70 $Å^2$ of additional surface area buried in the mutant protein.

Beeby et al. analyzed sequence and structural data from P. aerophilum (archaea) and Thermus thermophilus (thermophilic bacteria) and found that disulfide bonds stabilized proteins from these species. Cytoplasmic proteins from eukaryotes don't have disulfides due to the presence of reducing agents (such as glutathione) in the cell. In thermophiles, a novel protein, protein disulfide oxidoreductase, was identified that catalyzes the formation of sulfide bonds. Finally, Berezovsky and Shakhnovich have also analyzed proteins from hyperthermophilic archaea and bacteria and compared them to analogous proteins from mesophilic bacteria. They found two types of stabilization of hyperthermophilic proteins, depending on the organism's evolutionary history. Proteins from cells that originally evolved in high-temperature conditions (Archaea) were very compact (maximizing van der Waals interactions). They had a high number of contacts per residue and a high percentage of hydrophobic residues. Still, they did not use specific structural stabilizing interactions (such as electrostatic interactions in salt bridges). In contrast, proteins from cells that originally evolved under mesophilic conditions but later adapted to hyperthermophilic conditions had evolved specific sequence features that stabilized electrostatic interactions (e.g., more charged residues, salt bridges).

4.7.9.11: Climate Change and The Effect of Increasing Temperatures on the Biosphere

Perhaps humanity's greatest challenge is the effects of global warming and climate change on the biosphere and its health. Follow the links below to see how increasing temperatures on our planet due to the production and use of fossil fuels affect the health of our biosphere.



32.11: Part 3 - A Warmer World: Temperature Effects On Chemical Reactions

32.12: Part 3 - A Warmer World: Temperature Effects On Proteins

4.7.9.12: Summary

Protein stability is governed by a delicate balance of noncovalent interactions and thermodynamic factors that together determine a protein’s ability to adopt and maintain its native three-dimensional structure. This chapter builds on fundamental concepts from physical chemistry to explore the forces that drive protein folding and stabilize the native state, as well as those that oppose folding, such as conformational entropy.

At the core of protein stability is the interplay between favorable interactions—such as the hydrophobic effect, hydrogen bonds, van der Waals forces, and specific ion-ion interactions—and unfavorable contributions, such as chain conformational entropy. Although studies of small molecule transfers (e.g., benzene from a nonpolar solvent to water) initially suggested that hydrogen bonds might not be the primary driving force for protein folding, experimental evidence from wild-type and mutant proteins indicates that hydrogen bonds can significantly stabilize the folded state. In many cases, the energetic gain from burying hydrogen-bonding groups within the protein interior can even outweigh the contributions of the hydrophobic effect.

The overall free energy change (ΔG) for protein folding is only marginally negative, reflecting that native proteins are only modestly stable under physiological conditions. This marginal stability implies that even small disruptions—such as a single unsatisfied hydrogen bond or the formation of a cavity due to an amino acid mutation—can trigger unfolding. Differential scanning calorimetry (DSC), circular dichroism (CD) spectroscopy, and fluorescence measurements are key techniques for quantifying protein stability, providing insights into parameters such as the melting temperature (T_m) and changes in heat capacity (ΔCp).

A significant portion of this chapter is dedicated to understanding how model compounds, such as N-methylacetamide (as a peptide bond mimic) and benzene (representing hydrophobic side chains), are used in thermodynamic cycles to estimate the contributions

of hydrogen bonding and the hydrophobic effect. These studies show that, while hydrogen bonds themselves may not drive the protein's collapse, they are critical for the proper internal organization of the folded state. In contrast, the hydrophobic effect, driven by the entropic release of ordered water molecules around nonpolar groups, is a major factor in the overall stability of proteins.

Mutagenesis studies further reinforce these ideas by demonstrating that even minor changes in amino acid side chains can significantly alter protein stability. For example, replacing a hydrogen bond–capable asparagine with alanine often results in a noticeable decrease in stability, suggesting that the formation of buried hydrogen bonds contributes more to stability than the hydrophobic effect alone.

Finally, the chapter discusses how proteins from thermophilic organisms adapt to extreme temperatures. These proteins often display features such as increased numbers of salt bridges, tighter packing, and, in some cases, disulfide bonds to enhance stability under high-temperature conditions. Comparative studies of proteins from mesophiles and thermophiles underscore that both specific interactions (like ion pairs) and global properties (such as compactness) play important roles in thermal stability.

Overall, this chapter challenges students to integrate thermodynamic principles with experimental observations to understand the marginal stability of proteins. This concept is central to both basic biochemical research and practical applications such as protein engineering and drug design.

This page titled 4.7.9: Protein Stability - Thermodynamics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.9: Protein Stability - Thermodynamics** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules

Learning Goals (ChatGPT o1, 1/26/25)

1. **Differentiate Normal vs. Aberrant Aggregation:**
   - Explain the difference between protein multimerization (e.g., dimers, trimers, filaments) in the native state and the formation of pathological aggregates such as amyloid fibrils.
   - Recognize the role of conformational flexibility and alternative folding in driving the formation of aggregates.
2. **Describe the Mechanism of Amyloid Fibril Formation:**
   - Illustrate how proteins that can adopt two stable or metastable conformations convert from a native (often α-helical) form to an abnormal (β-sheet-rich) conformation.
   - Explain the nucleation process and subsequent fibril elongation via templated conformational conversion.
3. **Analyze the Role of Molecular Chaperones:**
   - Discuss how molecular chaperones assist in proper protein folding and prevent aggregation.
   - Understand how chaperone malfunction or overload can contribute to the accumulation of misfolded proteins.
4. **Link Protein Aggregation to Disease Pathogenesis:**
   - Identify the connection between amyloid aggregates and neurodegenerative diseases (e.g., Alzheimer's, Parkinson's, familial amyloidotic polyneuropathy, prion diseases).
   - Discuss how mutations, environmental factors, or post-translational modifications can shift the balance toward aggregation and disease.
5. **Interpret Structural and Experimental Data:**
   - Utilize structural models (e.g., cryo-EM and NMR data) to understand the common β-sheet architecture of amyloid fibrils.
   - Evaluate experimental approaches (e.g., in vitro folding/refolding experiments and mutagenesis studies) used to investigate the aggregation process.
6. **Examine the Specificity and Regulation of Aggregation:**
   - Explain how specific amino acid sequences and even single-point mutations can program a protein to aggregate.
   - Analyze factors that control the transition from soluble protein states to reversible phase-separated granules or irreversible fibrils.
7. **Understand Phase Separation and Its Biological Implications:**
   - Describe how liquid–liquid phase separation leads to the formation of intracellular granules and condensates.
   - Compare these reversible assemblies with the irreversible aggregation seen in amyloid-related diseases.
8. **Evaluate Therapeutic Implications:**
   - Discuss potential therapeutic strategies aimed at modulating protein aggregation or stabilizing the native state.
   - Consider how insights from aggregate formation can lead to the development of diagnostics and treatments for neurodegenerative disorders.

These learning goals encourage a deep understanding of the thermodynamic, structural, and cellular factors that influence protein aggregation, linking molecular-level phenomena to clinical outcomes and potential therapeutic approaches.

4.7.10.1: Introduction to Protein Aggregates

We have studied different types of protein aggregation, including aggregation of the native state (to form dimers, trimers, multimers, and filaments). We've also studied how misfolded proteins can aggregate and how a whole family of molecular chaperones helps newly synthesized and misfolded proteins fold correctly. But what happens if a protein can fold two reasonably

stable or metastable structures of starkly different conformation? A class of proteins known as **prions** or **amyloid proteins** exhibits this characteristic. Their "alternative" conformation can bind to the "normal" structure, causing it to flip to the alternative conformation. This can then seed the continuation of the process, which ultimately forms large aggregates with a fibrillar structure. This process leads to aberrant cell function and, when it occurs in neurons, can lead to a variety of brain diseases such as Alzheimer's Disease.

Protein aggregates complicate the work of researchers studying protein folding in vitro and attempting to express human proteins in prokaryotes, such as E. coli, in vivo. This often results in the formation of large protein aggregates called inclusion bodies. Instead of viewing these aggregates as unwanted "junk", some study them avidly. It turns out that these aggregates are not as non-specific as was earlier believed. In addition, understanding how and when they form will give us clues into the etiology and treatment of some of the most debilitating and feared diseases.

4.7.10.2: Specificity of Aggregate Formation

In the early 1970s, it was demonstrated that chymotrypsinogen could not be folded in vitro without forming aggregates. An intermediate was presumed to have formed that, if present in high concentration, would aggregate irreversibly instead of folding to the native state. Refolding of tryptophanase showed that it aggregated only with itself, suggesting specificity. In the 1980s, a single amino acid folding mutant was found in a viral protein. Both the normal and mutant viral proteins unfolded at high temperatures. Still, only the mutant would aggregate at high temperatures, suggesting that aggregation could be programmed into or out of a gene. Also, a single amino acid change in bovine growth hormone prevented aggregation without affecting correct folding.

This knowledge of protein folding and aggregation was soon applied to understanding several diseases in which protein aggregates were observed, either initiating or being associated with the diseases. These protein aggregates were termed "**amyloid** deposits" and appeared to be associated with, and perhaps causative of, several neurodegenerative diseases. The term **amyloid** was first used by the German pathologist Rudolf Virchow. In 1853, he described waxy tissue deposits associated with eosinophils (a type of immune cell). These deposits resembled starch (made of amylose and amylopectin), so he termed them amyloid. However, all known amyloid deposits are composed of protein, not starch.

It now appears that these diseases are likely caused by improper protein folding and subsequent aggregation. Except in rare inherited diseases, amyloid deposits are composed of normal, wild-type proteins (not mutants) that undergo conformational changes to form monomers. These monomers then catalyze the conversion of normal monomers into the altered form, which polymerize into fibrils. Sometimes, in inherited conditions or when mutations occur in a specific protein, amyloid deposits consist of the mutant protein. The proteins in these deposited fibers are composed predominantly of β sheets perpendicular to the fiber axis. Sometimes, the protein's monomeric "normal conformation" has little beta sheet structure.

Figure 4.7.10.1 shows a simplified model of how a normal protein with a "normal" conformation, enriched in this case for illustration in alpha-helices, can form fibrils of abnormal monomers, which are highly enriched in beta sheets. These can self-associate to form large insoluble protein "amyloid" fibers. Note the green arrows representing beta-strands.

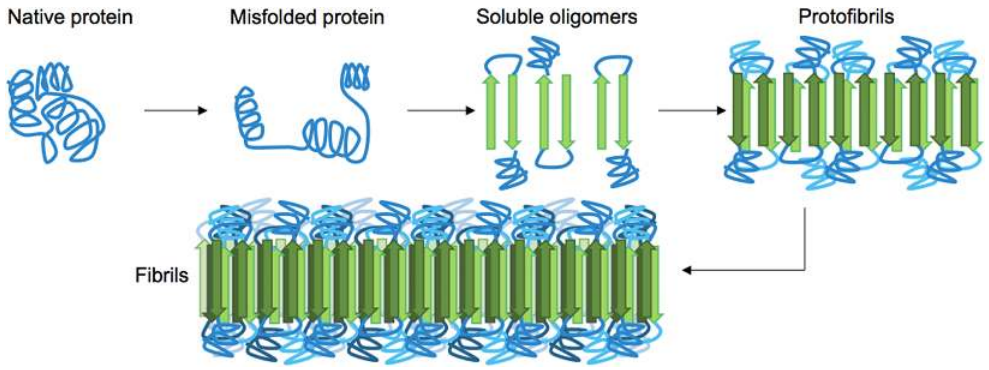


Figure 4.7.10.1: Simplified pathway toward amyloid fibril formation.Kaminski et al. *Neurophotonics* 3(4), 041807 (29 June 2016). https://doi.org/10.1117/1.NPh.3.4.041807 . Creative Commons Attribution 3.0 Unported License.

In the misfolded state, proteins have an increased propensity to oligomerize through interactions of their metastable beta-sheet domains. These can form more stable beta-sheet states, and the ensuing oligomers act as nuclei for subsequent elongation reactions, forming so-called protofibrils. The final amyloid fibril usually consists of many intertwined protofibrils.

4.7.10.3: Diseases of Protein Aggregates

We'll now describe a series of diseases caused by or highly associated with fibril formation from normal soluble proteins. For each one, we will present the best available structure of the amyloid fiber obtained mostly through cryo-EM. What's amazing is that, at first glance, all of the amyloid fiber structures have an astonishingly similar structural appearance. We present them not to be redundant but to illustrate how natural processes can yield a common structural outcome, often lethal, from a great diversity of protein structures. The amyloid fibers are so structurally similar that a common therapy to prevent their formation may be developed.

4.7.10.3.1: Familial amyloidotic polyneuropathy (FAP)

This affects 1/10,00 to 1/100,000 people. The monomer protein involved is **transthyretin** (147 amino acids, MW 15,887), which normally exists in the blood as a homotetramer (a dimer of dimers). Figure **4.7.10.2** shows the structure of the dimer (6fxu). Note that each monomer contains mostly beta structure. The protein binds L-thyroxine, and around 40% of blood plasma transthyretin is bound to retinol-binding protein.

Figure **4.7.10.2**: Structure of the transthyretin dimer (amino acids 21-147) (6fxu)

In mildly acidic conditions in vitro, the equilibrium between the tetramer and monomer is shifted toward the monomer, which can aggregate into fibrils. Monomer aggregation could be promoted by a possible transition to a molten globule-like state (as discussed previously for lactalbumin). This secondary but loosely packed tertiary structure has more exposed hydrophobic groups. If the concentration is high enough, the molten globules aggregate. In individuals with the disease, mutations in the protein destabilize the tetramer, shifting the equilibrium toward the monomer, which presumably increases molten globule formation and aggregation. Specifically, Val30Met and Leu55Pro mutations promote the dissociation of the tetramer and the formation of aggregates. Conversely, Thr119Met inhibits tetramer dissociation—amyloid aggregates deposit in the heart, lungs, kidneys, and other organs, leading to death. Figure **4.7.10.3** shows how the transthyretin dimer could be stabilized to form monomers or dimers, leading to fibril formation.

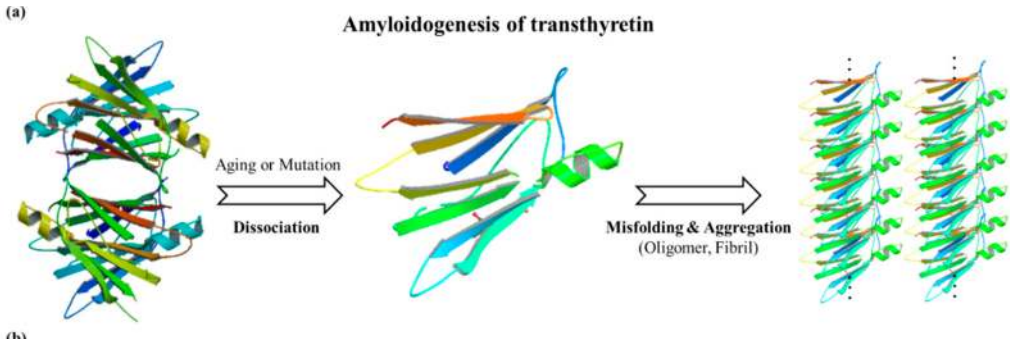


Figure **4.7.10.3**: Conversion of the transthyretin dimer to fibril formation. Park, G.Y.; Jamerlan, A.; Shim, K.H.; An, S.S.A. Diagnostic and Treatment Approaches Involving Transthyretin in Amyloidogenic Diseases. *Int. J. Mol. Sci.* **2019**, *20*, 2982. https://doi.org/10.3390/ijms20122982. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

Figure 4.7.10.4 shows the intramolecular hydrogen bonds bonds within monomers (-----) and intermolecular hydrogen bonds between monomers (-----) in the transthyretin dimer. We present this figure to remind readers that in all of the complex amyloid fibrillar structures presented below, the bulk of hydrogen bonds are "intermolecular" between adjacent monomers in their extended states (phi/psi angles consistent with beta-structure) within the fibrillar structure.

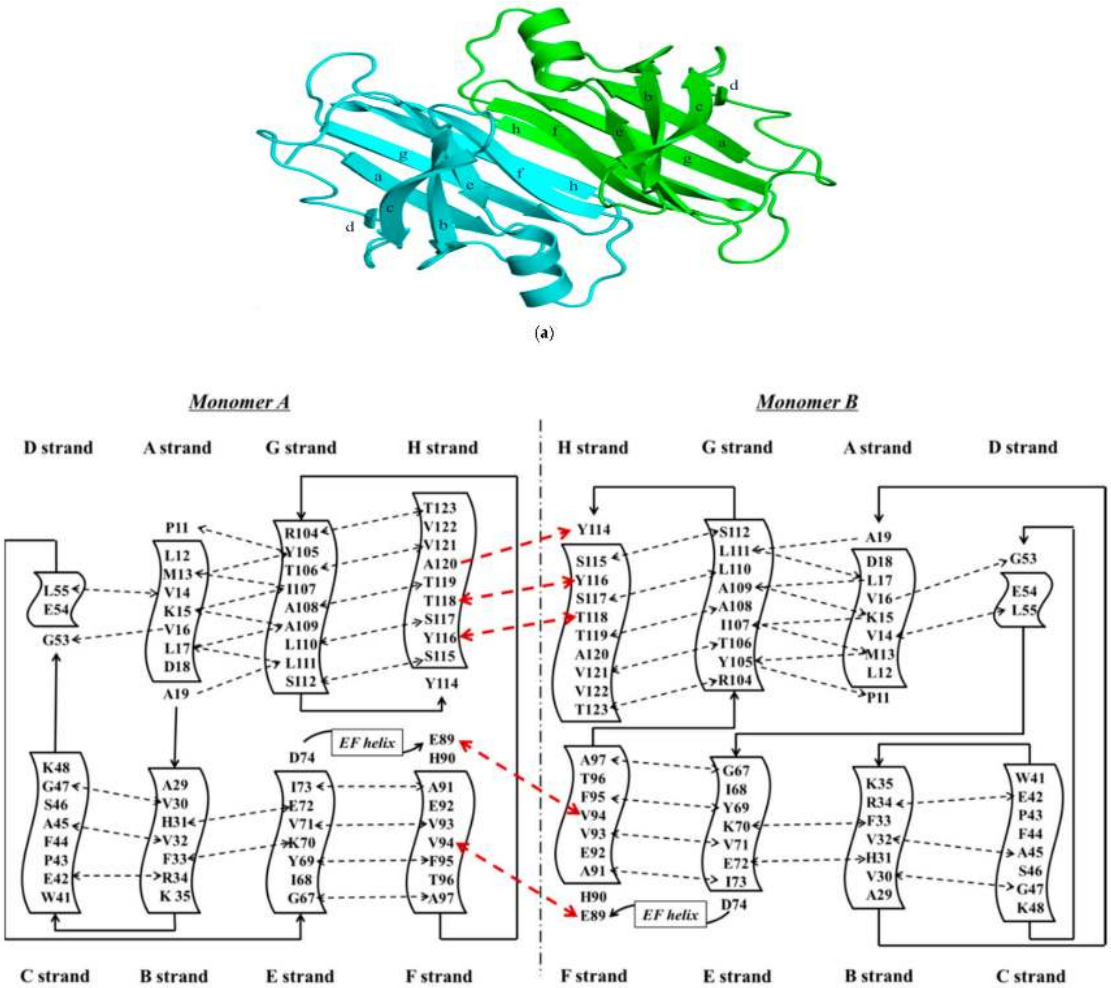


Figure 4.7.10.4: (**a**) 3-Dimensional structure of ATTRwt (PDBe ID code; 1BMZ [18]) in dimeric form; Hydrogen bonds within monomers (-----) and **between** monomers in the dimer (-----) in the **transthyretin dimer.** Park et al., ibid

Figure 4.7.10.5 shows an interactive iCn3D model of the Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis (6SDZ). Each separate monomer is shown in a different color. The static image below shows arrows indicating beta strands forming hydrogen bonds from the main chain amide Hs and carbonyl Os to another main chain atom on an adjacent monomer.

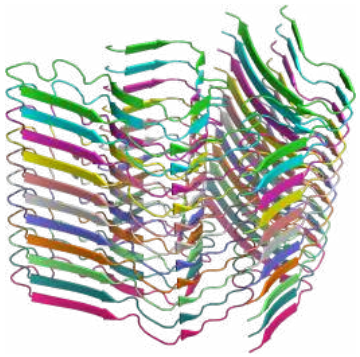


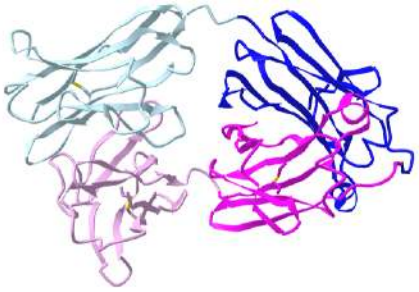
Figure 4.7.10.5: Click the image for a popup or use this external link: Cryo-EM structure of a transthyretin-derived amyloid fibril from a patient with hereditary ATTR amyloidosis (6SDZ). (Copyright; author via source) Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wMtaWWJ5KJkQ99

Note the beautiful but, unfortunately, deadly array of adjacent extended monomeric chains (each colored differently) that form hydrogen bonds with adjacent extended monomers to stabilize the fibrillar structure.

4.7.10.3.2: Light Chain Amyloidosis - AL amyloidosis (amyloidosis from the light chain)

The light chain (MW approx 25,000) is a normal component of circulating immunoglobulin antibody (protein) molecules. Each contains two light and two heavy chains. We will discuss the structure of antibodies in detail in Chapter 5.5. Needless to say, antibodies are very diverse molecules. The immune system can generate antibodies to recognize almost any foreign molecule. The light chains of antibodies are incredibly diverse and variable, despite all having two **immunoglobulin domains**. Each domain consists of approximately 110 amino acids, organized into two layers of β-sheets. Each layer contains 3-5 antiparallel β-strands, with a disulfide bond connecting the two layers. As a result, they exhibit significant β-structure in their monomeric form. The large diversity in light chains is partly generated by recombining gene fragments to produce individual light chains. Some variants of the light chains (λ1, λ2, λ3, λ6, and κ1) are associated with AL amyloidosis, a potentially fatal disease. Mutants in the light chain can cause a destabilization of the native state to a state similar to a molten globule, which then conformationally converts to a structure that aggregates into amyloid fibers. These can be deposited in various tissues.

As shown above, the pathway that a simple 25K monomer takes to produce such a complex but regular structure must start with a simple dimer formation between two monomers. Figure 4.7.10.6 shows an interactive iCn3D model of the normal conformation of a λ6a light chain **dimer** (6mg4). The λ6a light chain variant is more prone to aggregate to form fibrils.



NCBI iCn3D Figure 4.7.10.6: Full-length human lambda-6A light chain dimer showing IgG fold domains (6mg4). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...MNsiddm4RAdKn8

One light chain in the dimer is shown in blue, with one IgG fold domain in light blue and the other in dark blue. The other light in the dimer is shown in shades of magenta to show the two domains. Normally, light chains don't form dimers; instead, they associate with heavy chains to form full IgG antibodies. Free light chains will form dimers if not associated with a heavy chain, altering conformation and forming amyloid fibrils.

Figure 4.7.10.7 shows an interactive iCn3D model of the AL amyloid fibril from a lambda 3 light chain in conformation A (6Z1O). Each separate monomer is represented in a distinct color.

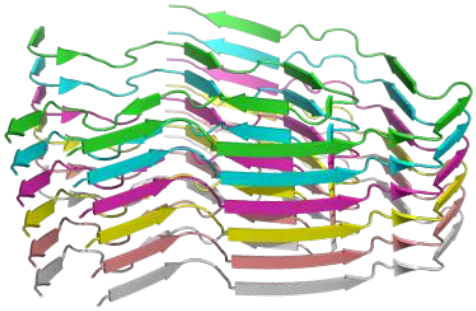


Figure 4.7.10.7: AL amyloid fibril from a lambda 3 light chain in conformation A (6Z1O). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...pSu2Nm714gvjh7

Figure 4.7.10.8 shows a comparison of the secondary structures of the native conformation of the light chain and those found in the two different fibrillar forms (A and B) of the amyloid fibrils (panel a). Remember that each light chain has two IgG domains, each having two sets of β-sheets containing 3-5 antiparallel β-strands.

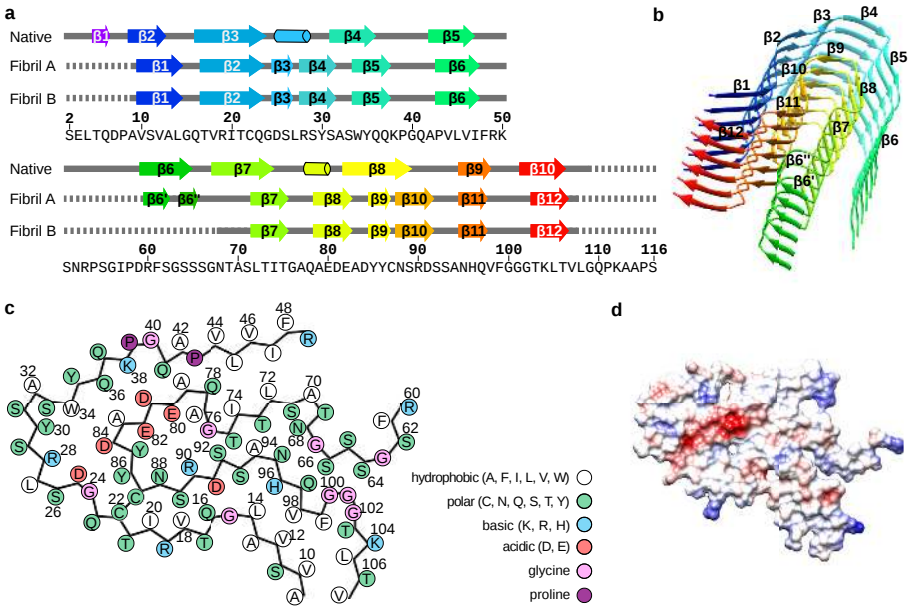


Figure 4.7.10.8: Secondary structure and folding of the fibril proteins. Lynn Radamaker et al. *Nat Commun* **12,** 875 (2021). https://doi.org/10.1038/s41467-021-21126-2 . Creative Commons Attribution 4.0 International License

Panel b shows the alignment of six fibrillar light chain proteins in conformation A. Panel C shows the backbones' trace and the side chains' local environment in a **single** light chain from the fibrillar A conformations. It shows cross-β-sheets interactions with parallel hydrogen bonds across strands. Note the disulfide that connects Cys 22 and Cys 87 on adjacent strands. The circles show surface side chains. Note that they are enriched in green (polar) and blue/red (basic/acidic) regions, with some hydrophobic patches (e.g., V10-L14). Some are pairs, such as D24 and R28. Panel D shows the electrostatic surface of a single light chain from the fibrillar A conformation, with red indicating negative and blue positive.

4.7.10.3.3: Alzheimer's Disease (AD)

This disease accounts for 60-80% of dementia cases. Brains in Alzheimer's patients have amyloid deposits of the **amyloid beta (Aβ) protein** and aggregates of a protein called **tau.** The origin of this disease remains unresolved. Some believe aggregates of the amyloid beta protein cause the disease. Others suspect that tau plays a role in forming tau bundles. Still others suggest an infectious disease agent (more on that later). Regardless of the underlying cause, the amyloid beta protein aggregates are neurotoxic and, at the very least, correlate with, if not cause, the disease.

The amyloid aggregates in Alzheimer's disease begin with a change in a monomeric protein normally found in neuronal membranes. The protein, called **β-amyloid precursor protein** (**BAPP** or simply **APP**), is a transmembrane protein. A slightly truncated, soluble form is secreted from cells and found in the extracellular fluid (cerebrospinal fluid and blood). The normal function of these APP proteins is not yet clear. An endoprotease cleaves a small 40-42 amino acid fragment from this protein, named the **amyloid beta (Aβ) protein**. Figure **4.7.10.9** shows the normal processing of the amyloid precursor protein APP (left) and the abnormal, amyloidogenic form (right).

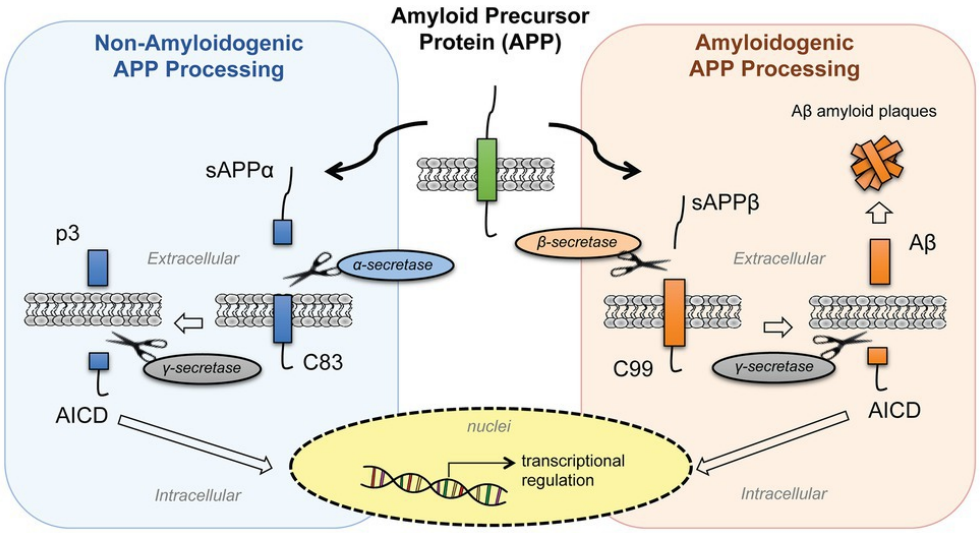


Figure **4.7.10.9**: Processing of amyloid precursor protein (APP) in normal and abnormal conditions. Ristori et al. Front. Physiol., 27 August 2020 | https://doi.org/10.3389/fphys.2020.01056. Creative Commons Attribution License (CC BY).

In a "normal" processing pathway, the proteases α- and γ-secretase release two variant peptides into the extracellular environment: soluble amyloid precursor protein cleaved by α-secretase (sAPPα) and p3 fragments. In the amyloidogenic processing pathway, β- and γ-secretases release soluble amyloid precursor protein cleaved by β-secretase (sAPPβ) and β-amyloid (Aβ) peptides. Both pathways release the same intracellular domain, AICD, which moves to the nucleus and acts as a transcription factor to regulate gene expression. The β-amyloid (Aβ) peptides aggregate to form the fibrillar plaque.

The amyloid beta (Aβ) protein, or a mutant form of it, aggregates to form beta-sheet-containing fibrils in Alzheimer's disease. The NMR solution structure of the monomer amyloid beta-peptide (1-42) is shown in Figure **4.7.10.10**. Note the absence of any beta structure.

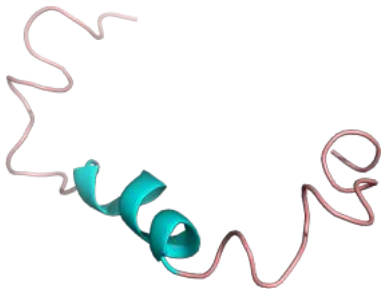


Figure **4.7.10.10**: Solution structure of the amyloid beta-peptide (1-42) 6SZF.

Several mutations in different proteins have been linked to Alzheimer's disease, but they all appear to increase the production or deposition of the amyloid beta protein. These deposited plaques are extracellular and have been shown to cause neuronal damage. They are located in brain regions essential for memory and cognition. The APP gene is located on chromosome 21, the same chromosome that is present in an extra copy (trisomy 21) in Down syndrome, which is characterized by symptoms including presenile dementia and amyloid plaques. Aggregate formation appears to be driven by increased APP expression and, hence, amyloid beta protein. Additionally, some mutants may serve to destabilize the amyloid beta protein, thereby increasing its aggregation.

Figure 4.7.10.11 shows an interactive iCn3D model of the prevalent amyloid-beta fibril structure from Alzheimer's disease brain tissue (6W0O).

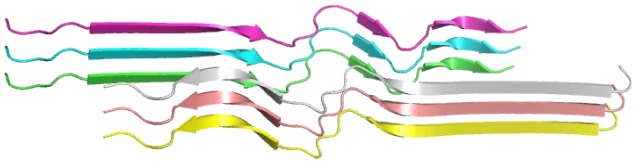


Figure 4.7.10.11: Amyloid-beta(1-40) fibril derived from Alzheimer's disease cortical tissue (6W0O). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...NZCX9sAhnFPiB6

The **tau** protein (758 amino acids, MW 78,928), much larger than the other proteins we discuss in this chapter, has also been implicated as a cause or factor in Alzheimer's Disease. It facilitates microtubule assembly and stability, as well as other neuronal functions (see Chapter 5). Since its C-terminus binds microtubules in axons and the N-terminus binds the plasma membrane, it might link both. The cytoskeleton of neurons is disrupted in Alzheimer's and in patients with other neurodegenerative diseases. When tau tangles are present, these diseases are also referred to as **tauopathies**. One tauopathy is **chronic traumatic encephalopathy (CTE)** caused by repetitive head impacts (from contact sports or physical abuse) or concussions arising from explosions in combat. No full-length structure of tau has been determined to date. The structures of a predicted model of solution phase monomeric tau and tau fibers from the brain of patients with neurodegenerative diseases (Alzheimer's, CTE, and Corticobasal Degeneration - CBD) are shown in Figure 4.7.10.12.

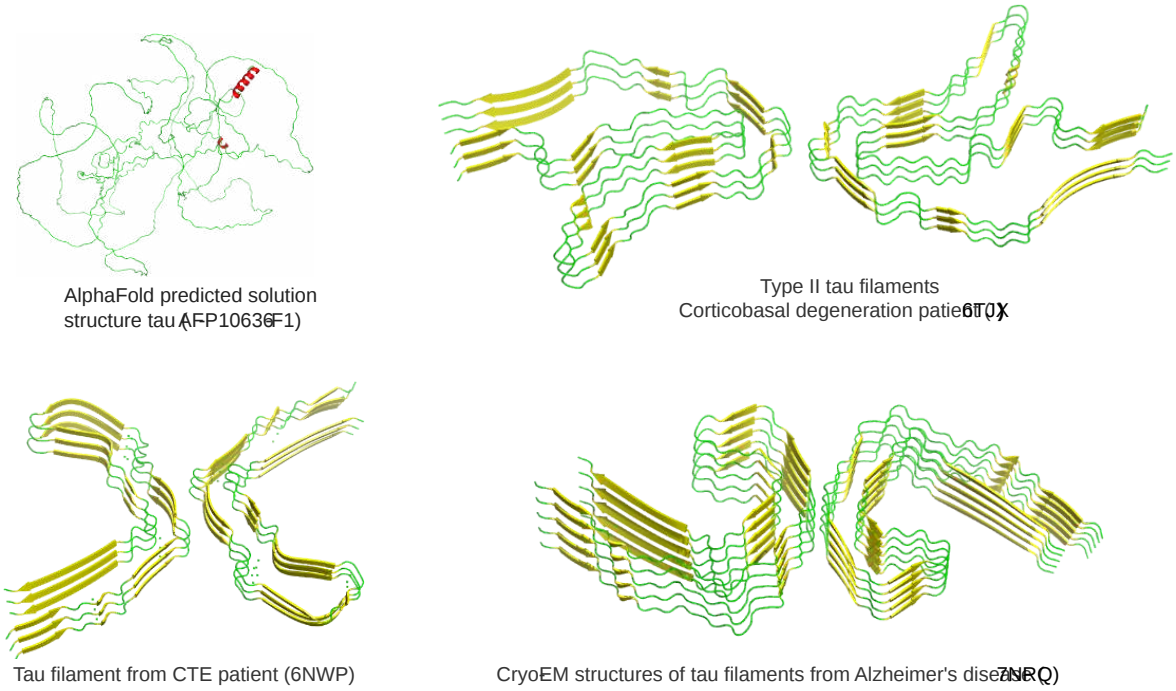


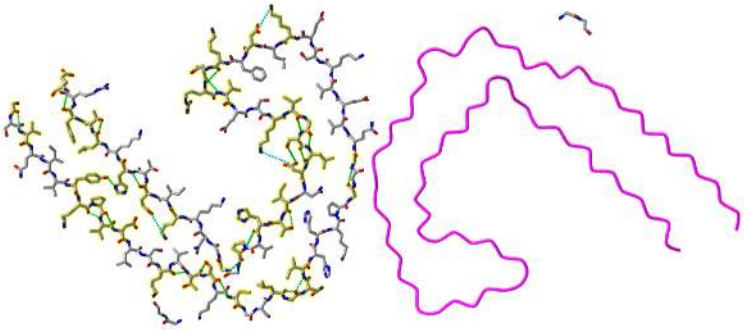
Figure 4.7.10.12: Structure of solution phase tau and tau filaments from various tauopathies.

The predicted structure of the monomeric protein (using AlphaFold) is almost completely devoid of secondary structure. The tau structures in other tauopathies are similar but distinct. In CBD tau fibers, there are four microtubule-binding repeats (4R). Pick's Disease, another tauopathy, has three repeats (3R), while taus in AD and CTE are 3R or 4R.

The fibril cores of tau in both CBD (Lys 274-Glu380, which contains the end of R1 and R2-R4) and Alzheimer's Disease (AD) contain around 13% glycine residue. These allow the main chain flexibility and intersheet packing, enabling the large conformational changes necessary for beta structure formation and fibril formation. Repeats of the PGGG motif allow sharp turns

or extended chains. Valine (around 10%) and isoleucine, leucine, and phenylalanine facilitate inter-sheet packing through induced dipole-induced dipole interactions and the hydrophobic effect. Certain tau fibrils in CBD contain hydrophilic pockets that bind molecules that have yet to be elucidated and that might seed the nucleation of fibrils. The cavity has three lysines and a histidine, so any ligands that bind are likely anionic in nature. In addition, recent evidence suggests that different combinations of post-translational modifications, including ubiquitinylation and acetylation of lysines 311, 317, 321, 343, and 353, may lead to distinct tau fibril structures. CTE tau protein in tangles appears to be hyperphosphorylated.

Figure **4.7.10.13** shows an interactive iCn3D model of a paired helical **tau** filament from Alzheimer's Disease human brain tissue (6VHL).

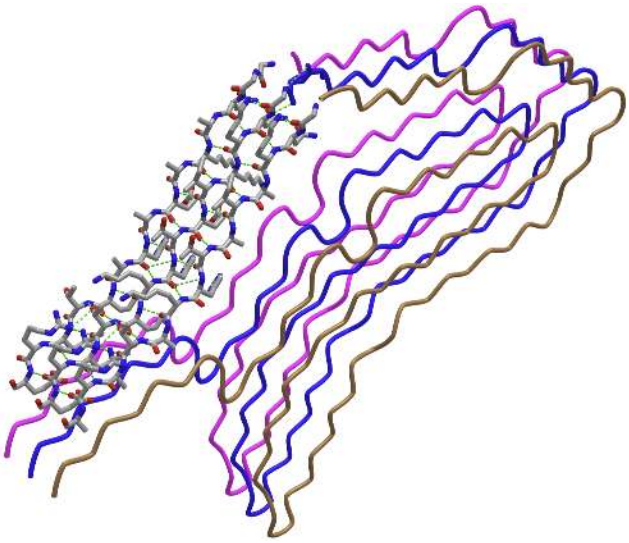


NCBI iCn3D Figure **4.7.10.13**: Paired helical filament of Tau (6VHL) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...NHsZdhDJNentg9

Zoom into the interactive images to see the H bonds within one of the filaments. Note that the left filament's hydrogen bonds (green dotted lines) are between side chains and not between backbone amide Hs and carbonyl Os. These are pointing above and below the backbone plane, where they could interact with other chains above and below to form the multi-chain fibers seen in the examples above.

Figure **4.7.10.14** shows an interactive iCn3D model of a singlet Tau fibril obtained from corticobasal degenerated human brain tissue (6VHA).



NCBI iCn3D Figure **4.7.10.14**: Singlet Tau fibril from corticobasal degeneration of human brain tissue (6VHA). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...Bbxr2Z1caGqnG8

The backbone of the C-terminal amino acid (367-380) of each of the three separate chains of the fibril is shown in CPK colors, and hydrogen bonds between the strands that form the parallel beta strands are shown as green dashes.

Progress has been incredibly slow on ways to treat Alzheimer's. Most methods focus on reducing amyloid beta production and aggregation by identifying small molecules that inhibit specific steps in its production, including secretase cleavage and the subsequent conformational changes necessary to form amyloid fibers. But what if the amyloid aggregates are secondary to the primary cause? What if the amyloid beta protein overproduction was the brain's response to defend against the cause?

Selection Bias and and Fraud in AD Research

Alzheimer's patients have aggregates of the amyloid beta and tau proteins. Do these aggregates cause the disease? Maybe. Maybe not! It's very easy to confuse correlation with causation. Also, it isn't easy to go against present beliefs. If you seek funding for an alternative cause and your funding derives from grant applications reviewed by experts in the field who are firmly in the amyloid beta and tau "camps", you are not likely to get funding.

Here is an interesting story that illustrates a source of bias we all face, including AD researchers. During World War II, when British Royal Air Force (RAF) planes returned from bombing missions, they were inspected for damage caused by German anti-aircraft fire. To protect pilots from harm, the planes were reinforced in the areas that had received incoming enemy fire. It made sense and was intuitive until a statistician, Abraham Wald, realized that the damaged planes had returned safely. Instead of looking at the sites hit by enemy fire, he recommended looking at the sites NOT hit. In those areas, it was likely the hits were fatal. The fatal sites were not included in the collected sample data, so the sample was biased. Selectively attending to certain data or ideas while ignoring other information is called **selection bias**. We all face this, and scientists studying and funding AD also do.

Another problem that has emerged in AD research is scientific fraud. Decades of research have yielded no treatments that significantly impact the disease. One reason, in addition to selection bias, is scientific fraud. Charles Piller has written about extensive fraud using falsified data collected by key and leading researchers in the field. This has had an enormous and deleterious effect not only on AD research but on the victims of this disease as well.

4.7.10.3.4: Lewy Body and Parkinson's Disease

α-Synuclein (140 amino acids, MW 14,460) is expressed in the brain and presynaptic terminals in the central nervous system, but also in more distal neurons, and is involved in the regulation of neurotransmitter release and in the synaptic vesicles that hold them. Its aggregation is a cause or consequence of Parkinson's Disease and Lewy Body Dementia. It's found in the cytoplasm and the nucleus and is also secreted. Figure 4.7.10.15 shows an NMR solution structure (left) and AlphaFold-predicted structure (right) for this protein, which in solutions is so disordered that no full crystal structure has been determined.

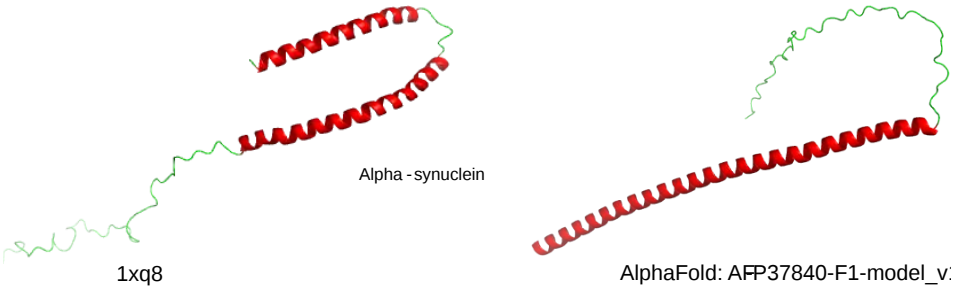


Figure 4.7.10.15: Alpha-synuclein structure (left, NMR solution structure, right, AlphaFold predicted structure)

Figure 4.7.10.16 shows an interactive iCn3D model of an amyloid fibril structure of alpha-synuclein determined by cryo-electron microscopy (6A6B). Two protofilaments with clearly Greek key topologies are shown.

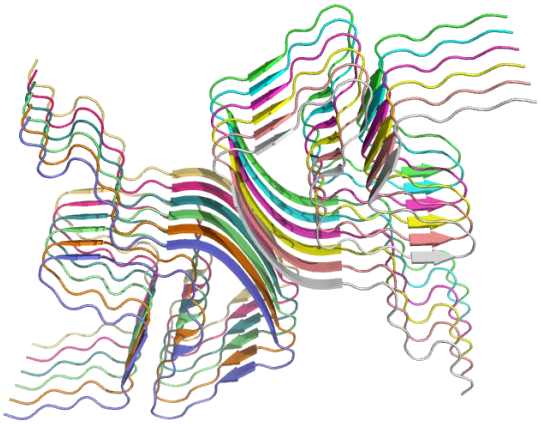


Figure 4.7.10.16: Amyloid fibril structure of alpha-synuclein determined by cryo-electron microscopy (6A6B). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SbQq7fhFpQ9K4A

? New insights into Parkinsons - can it start in the gut?

Click below for more information!

Answer

A substantial body of evidence suggests that Parkinson's Disease may have its origin in the gut or even the nose. Changes in the gut lining, caused by and/or leading to inflammation, as well as an altered gut microbiome (which can also cause inflammation), may lead to misfolding of gut alpha-synuclein. Sources of gut damage and inflammation include an altered gut microbiome, *Helicobacter pylori* infection, and the use of nonsteroidal anti-inflammatory drugs (NSAIDs). Altered synuclein formed in the gut could migrate through the vagal nerve from the gut to the brain. The processes involved are outlined in Figure 4.7.10.17*a* below.

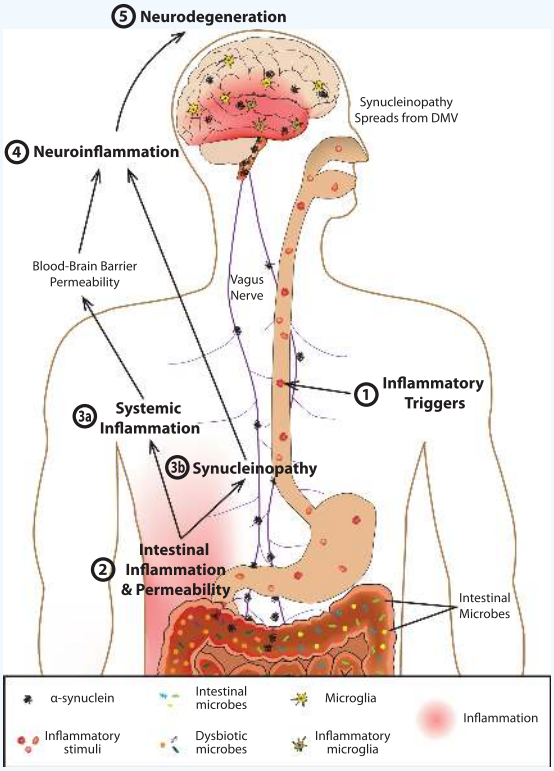


Figure **4.7.10.17*a***: **Model of gut-originating, inflammation-driven PD pathogenesis.** Houser, M.C., Tansey, M.G. The gut-brain axis: is intestinal inflammation a silent driver of Parkinson's disease pathogenesis?. *npj Parkinson's Disease* **3**, 3 (2017). https://doi.org/10.1038/s41531-016-0002-0. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

In a susceptible individual, inflammatory triggers (1) initiate immune responses in the gut that deleteriously impact the microbiota, increase intestinal permeability, and induce increased expression and aggregation of αSYN (2). Synucleinopathy may be transmitted from the gut to the brain via the vagus nerve (3b), and chronic intestinal inflammation and permeability can promote systemic inflammation, which, in turn, can increase blood-brain barrier permeability (3a). Intestinal inflammation, systemic inflammation, and synuclein pathology in the brain promote neuroinflammation (4), which drives the neurodegeneration that characterizes PD (5).

The vagal nerve is a very long nerve that consists of a bundle of sensory (80%) and motor (20%) neurons. It controls autonomic processes, such as breathing, blood pressure, heart rate, gut motility, and neural responses related to feeding and drinking. Sensory neurons of the vagal nerve are found in the digestive, cardiovascular, and pulmonary systems. Figure **4.7.10.17*b*** shows the vagal nerve and its branches.

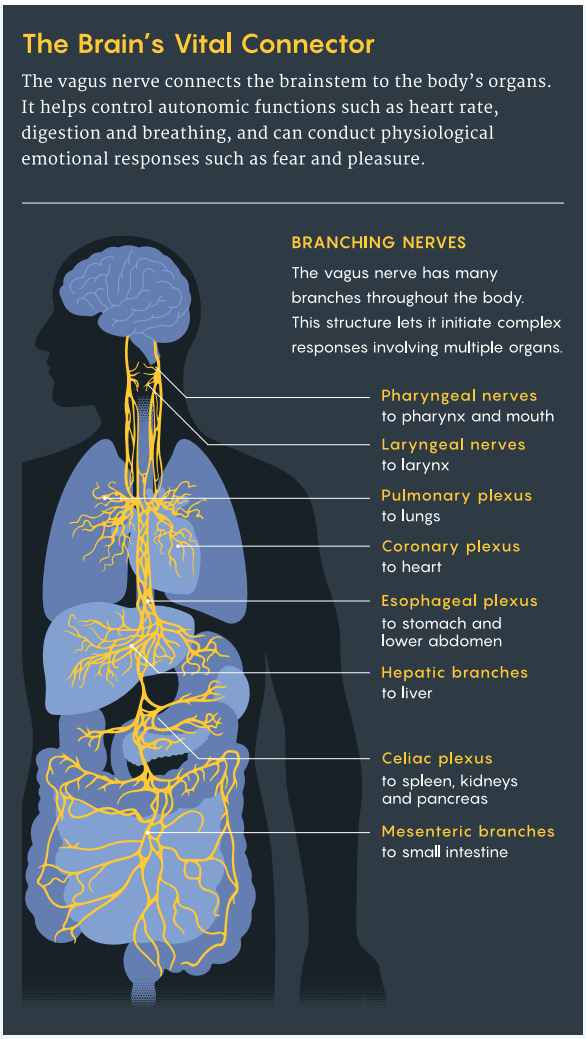


Figure **4.7.10.17b**: The Vagal Nerve. R. Douglas Fields. How Our Longest Nerve Orchestrates the Mind-Body Connection. QuantaMagazine. *August 26, 2024*. https://www.quantamagazine.org/how-o...tion-20240826/. With Permission from Quanta Magazine.

Another recent study suggests that altered alpha-synuclein might arise in the nose from the protein's interaction with environmental toxins. These include trichloroethylene (TCE) and perchloroethylene (PCE), used in dry cleaning and degreasing, paraquat (an herbicide), and air pollution.

4.7.10.3.5: Transmissible spongiform encephalopathies (TSEs) - Prion Diseases

Prion diseases are another set of brain diseases resulting from aggregates of monomers of the prion protein (PrP^c). As with the examples above, the aggregates form amyloid beta fibrils. The prion diseases include scrapie in sheep, bovine spongiform encephalopathy (mad cow disease), and in humans, Creutzfeldt-Jakob Disease (CJD), Fatal Familial Insomnia (FFI), Gerstmann-Straussler-Scheinker Syndrome, and Kuru (associated with cannibalism). In these fatal diseases, the brain, on autopsy, resembles a sponge with holes (hence the name spongiform). In contrast to the diseases described above, these can be transmitted from one animal to another, but typically not between species. (However, consider the controversy surrounding mad cow disease.) Also, the infectious agent can self-replicate in vivo. The logical conclusion is that a virus (slow-acting) is the causative agent. However, the infectious agent survives radiation, heat, chemical agents, and enzymes designed to kill viruses and their associated nucleic acids. Mathematical analyses suggested that the infectious agent in such diseases could be nothing more than a protein. Stanley B.

Prusiner, in the '80s, isolated just such a protein, which he named a **prion**, for a **proteinaceous infectious agent**. In October 1997, he was awarded the Nobel Prize in Medicine.

The normal monomeric prion protein, PrP^c (253 amino acids, MW 27,661), is highly conserved in mammals and is widely expressed in embryogenesis. Expression is highest in the central nervous system. The protein's normal function remains unclear. It is a physiological substrate to a particular membrane receptor (the Gpr 126 G protein-coupled receptor). Knocking out the gene shows that the normal protein is involved in synapse structure/function, myelination of neurons, and circadian rhythms, probably by acting as a transcription factor. It also helps regulate Cu^{2+} and Zn^{2+} levels in the central nervous system. The protein is cleaved, and a 209-amino-acid fragment is bound to the extracellular side of the neuronal membrane, anchored by the attachment of a lipid (GPI) anchor.

The N-terminal residues (23-124) are flexible, followed by residues 125-231, which are mostly alpha-helical. There is a disulfide between Cys179 and Cys214. The PrP^c (without the PI link) is water soluble, a monomer, protease-sensitive, and consists of around 45% alpha helix and 3% beta sheet. Given its highly disordered structure, no full-length crystal structure of the protein has been determined to date. The solution NMR structure of residues 125-231 has been determined, and the structure of the full protein has been modeled with AlphaFold. These structures are shown in Figure **4.7.10.18**.

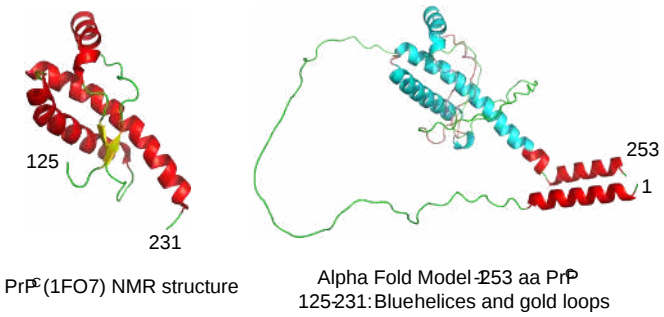
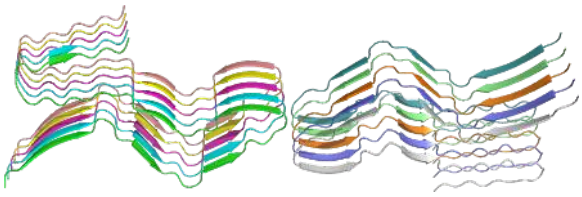


Figure **4.7.10.18**: Structure of the soluble form of PrP^C (left NMR, 1FO7, NMR structure, right Alpha Fold predicted model)

The blue helices and gold loops in the computer model consist of the same amino acids (125-231) as the NMR solution model on the left-hand side of Figure **4.7.10.2**.

The problem in transmissible spongiform encephalopathies (TSEs) is that neurotoxic amyloid-like protein aggregates form. The protein found in plaques (in cases other than those inherited) has the same primary sequence as PrPc but a different secondary and, presumably, tertiary structure. The protein found in the plaques, called PrPsc (the scrapie form of the normal protein), is insoluble in aqueous solutions, protease-resistant, and has a high beta-sheet content (43%) and a lower alpha-helix content (30%) than the normal version of the protein, PrPc.

Figure **4.7.10.19** shows an interactive iCn3D model of the cryo-EM structure of an amyloid fibril formed by full-length human prion protein (6LNI). Each line represents a PrP^{sc} chain from amino acids 170-229, which is the core of the fibril.



NCBI iCn3D Figure **4.7.10.19**: Cryo-EM structure of an amyloid fibril formed by full-length human prion protein. (6LNI) (Copyright; author via source).
Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/i...ByT4rL33brHRY7

The aligned zig-zag lines indicate beta strands that are hydrogen-bonded to the adjacent strands. The two alpha helices in the C-terminal domain become beta strands.

A genetic, inheritable form of the disease also exists, in which a mutant form of the PrP^c occurs, whose normal structure is destabilized by the mutation. The aggregates caused by the mutant form of the disease are understandable in light of the other diseases we discussed above. How does the normal PrP^c form PrP^{sc}? Evidence shows that if radiolabeled PrP^{*c} from scrapie-free cells is added to unlabeled PrP^{sc} from scrapie-infected cells, the PrP^{*c} is converted to PrP^{*sc}! It appears that the PrPc protein has two forms, not much different in energy, one composed mostly of alpha helices and the other of beta sheets. A dimer of PrP^{c}·PrP^{sc} might form, destabilizing $PrP^{*c,}$ causing and driving a conformational shift to the PrP^{sc} form, which would then aggregate. Exposure to the PrP^{sc} form would then catalyze the conversion of normal PrP^c to PrP^{sc}. Hence, it would be transmissible by contact with just the PrP^{sc} form of the protein. Likewise, species specificity could be explained if only PrPc dimers were present. PrP^{sc} could form from proteins of the same species. The inherited form of the disease can be explained by the mutant form of the normal protein, which forms the beta structure more readily, leading to aggregation.

The same mutation in PrPc, Asp178Asn, can cause two distinct diseases: CJD and FFI. Which disease you get depends on whether you have 1 of two naturally occurring, nonharmful variants at amino acid 129 of the normal PrPc gene. If you have a Met at that position and acquire the Asp178Asn mutation, you get CJD. If, on the other hand, you have a Val at amino acid 129 and acquire the Asp178Asn mutation, you get FFI. This disease was first observed in 1986 and has been reported in five families worldwide. It manifests in the late 50s, and it affects men and women equally. It is characterized by a progressive loss of sleep and disrupted circadian rhythms. The brain shows neuronal losses. It is known that amino acids 129 and 178 occur at the start of alpha helices, as predicted from propensity calculations. Chronic exposure to micromolar levels of synthetic fragment 106-126 of PrP^c kills hippocampal neurons. This peptide also has the greatest tendency to aggregate synthetic PrP^c peptides.

Kuru killed many members of the Fore tribe in New Guinea until the cannibalistic practice of eating dead relatives was stopped. Analysis of the genes for the prion protein in the Fore tribe and other ethnic groups worldwide shows two versions differing by just one amino acid in all people (remember that a single gene is represented on both the maternal and paternal chromosomes). That these two forms exist worldwide suggests that they have been selected for by evolution and confer some biological advantage. People with just one form of the protein are more susceptible to developing prion diseases. Mead and Collinge have shown that about 75% of older Fore women (who had lived through cannibalistic practices) had two different prion genes, compared to about 15% of women from other ethnic groups. This high percentage suggests that these women were protected from the disease, leading, through natural selection, to a high proportion of heterozygotes in this defined population. The presence of two forms of the prion gene (which likely protect against prion disease) suggests that cannibalism might have been widespread among our early ancestors.

There appears to be a main difference between the formation of amyloid fibers from prion proteins and those from other proteins, such as mutant lysozymes. If you add mutant lysozyme to normal lysozyme, the amyloid fibers contain only the mutant protein. However, if you incubate mutant prion proteins with normal prions, the normal proteins become pathological.

Recently Added: 10/11/25

4.7.10.4: Polynucleotide Repeat Mutations

Deleterious protein aggregates can also occur in proteins whose gene contain anomalous expansions and repeats of triplet nucleotide-encoding key amino acids like glutamine and glycine. We will discuss two below.

4.7.10.4.1: Huntingtin Protein and Huntingon's Disease

A mutant version of the **huntingtin** protein causes the fatal neurological Huntington's Disease.

- The normal huntingtin gene has fewer than 35 CAG repeats (encoding a string of less than 35 polyQ in the protein). The polyQ section of the protein can form short intramolecular beta-strand structures that can interact with other huntingtin proteins to form some amyloid structure, but they may also stay extended.
- The huntingtin mutation causes more than 35 CAG repeats, leading to an expanded polyglutamine (polyQ) stretch in the N-terminal (encoded by the first exon of the HTTex1 gene) of the protein. This tract of polyGln forms a beta-hairpin connecting ~20-residue-long β-strands in a tight turn. These beta strands on one huntingtin can, through intramolecular hydrogen bonding, form amyloid fiber aggregates and undergo liquid–liquid phase separation.

The antiparallel β-sheets from the polyQ core in the mutant lack the usual twist found in amyloid fibers. Interactions between adjacent eight-residue glutamine strands are shown in Figure **4.7.10.*xx*** below.

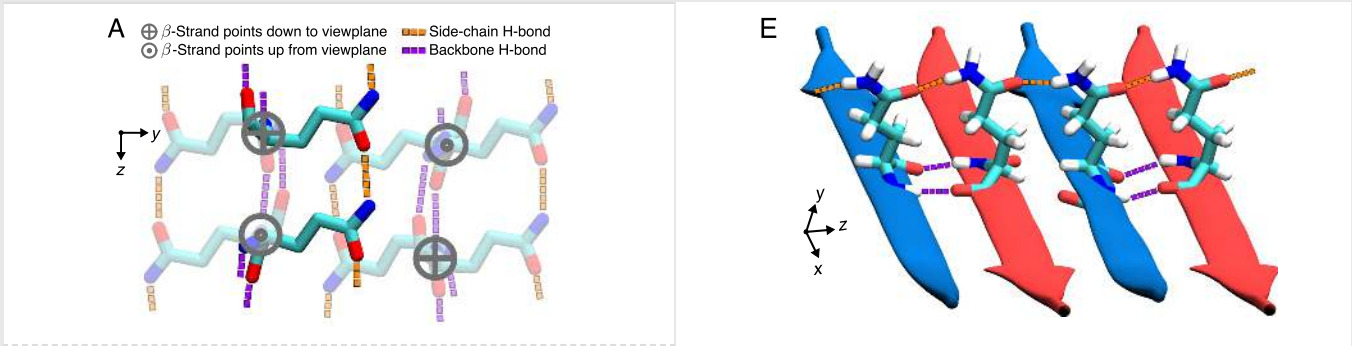


Figure **4.7.10.*xx***. Structure of the polyQ beta structure hydrogen bonds in huntingtin. Bagherpoor Helabad, M., Matlahov, I., Kumar, R. *et al*. Integrative determination of the atomic structure of mutant huntingtin exon 1 fibrils implicated in Huntington's disease. *Nat Commun* **15**, 10793 (2024). https://doi.org/10.1038/s41467-024-55062-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel **A:** The eight-Gln building block used to generate polyQ core candidates. Six residues are shown faded out to bring the two non-faded in focus: these two are on adjacent *β*-strands of an antiparallel *β*-sheet, stabilized by backbone hydrogen bonds (purple), and continuous chains of side-chain hydrogen bonds (orange). The latter are crucial for packing the polar glutamines into the waterless amyloid core. When generating the core candidates, all $\chi 1$ and $\chi 3$ dihedral angles were independently rotated to explore all possible hydrogen-bond networks.

Panel **E:** Illustration of the inter-side-chain hydrogen-bond ladders (orange) in *M2*

The mutant polyQ fibrils form an amyloid core composed of many stacked β-sheets, with water excluded. Envision it as a series of glutamine "ladders" that form two distinct conformers, A and B. On either side of the polyQ in the primary sequence is an N17 (N-terminal end) and a proline-rich domain PRD at the C-terminal end. The PRD section is quite disordered and protects sites in the N17 region from post-translational modification (ubiquitination and phosphorylation), which could have led to protein degradation and reduced amyloid formation.

MATLEKLMKAFESLKSFQ_{44}PPPPPPPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPPGPAVAEEPLHRP

Here are two models from molecular dynamics simulations of the aggregates. Figure **4.7.10.*xx*** shows an atomic model of the water-facing surface of the polyQ amyloid.

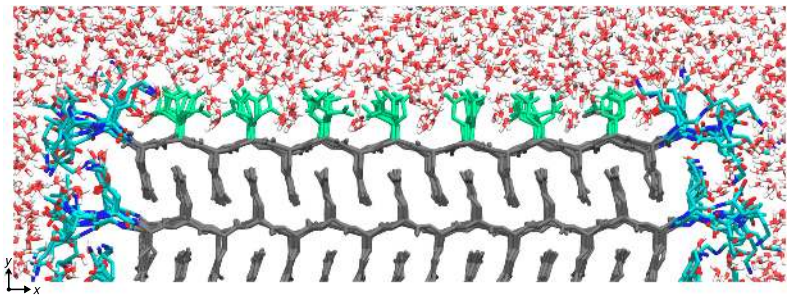


Figure **4.7.10.*xx***: Atomistic MD snapshot of the $D_2Q_{15}K_2$ peptide fibril's polyQ surface in contact with water. Exposed and buried Gln residues are colored green and gray, respectively. Note that the Gln side-chains within the amyloid core are well-ordered, while the water-facing side-chains exhibit greater mobility. Bagherpoor Helabad, M. et al., ibid

Figure **4.7.10.*xx*** shows a structural representation of the Q44-HTTex1 amyloid fibril.

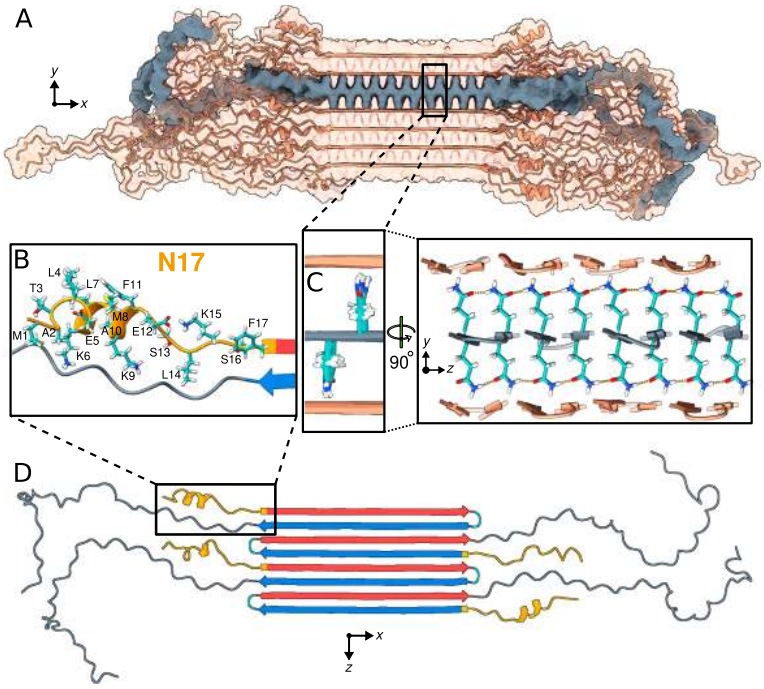


Figure 4.7.10.xx: Structural representation of the Q44-HTTex1 amyloid fibril. Bagherpoor Helabad, M. et al., ibid

Panel A: A graphical depiction (after 5-μs simulation in Amber14SB[86]) of the HTTex1 fibril. The region shaded in gray denotes a single sheet within the fibril's architecture.

Panel B: An atomic view of the N17 domain within the fibril, naming the specific amino acids.

Panel C: An atomic depiction of the glutamine side-chains within the fibril. The high stability of the fibril structure is primarily due to extensive hydrogen-bonding interactions among the glutamines, as depicted in the right panel.

Panel D: Top view representation of the β-sheet highlighted in panel (**A**). A quartet of HTTex1 monomers is visible. The polyQ is color-coded for the type "a" (red) and "b" (blue) strands; the tight β-turn is cyan. The N17 and PRD domains are orange and gray, respectively. Note the structural variation between different monomers in the same fibril sheet, including in particular the range of helical content in the N17 domain.

Download this amazing movie simulation of the Q44-HTTex1 fibril structure.

4.7.10.4.2: GGC repeat expansions and Fragile X-associated tremor/ataxia syndrome (FXTAS)

Expansions of another trinucleotide repeat, GGC, cause Fragile X-associated tremor/ataxia syndrome and neuronal intranuclear inclusion disease. The GGC repeats are translated into **a polyGly** sequence in protein when they are in the coding part of a gene, but can also be translated into different proteins if they are in the untranslated region near another site on the DNA. This non-canonical protein synthesis (translation) mechanism is called **RAN (R**epeat-**A**ssociated **N**on-AUG) translation. Transcription of the mRNA derived from the DNA sequence does not start with an AUG codon for the amino acid methionine.

Sometimes nucleotide repeats are 4 or 5 bases long. Given that translation is not regulated, multiple reading frames can be used to decode nucleotide repeats, allowing the synthesis of multiple proteins, including polyGly, polyAla, polyGln, polyPro, and polySer sequences.

PolyGly proteins (as well as the other poly sequences), given the small size of the side chain (-H), are inherently flexible and are very disordered. PolyGly and other polyAminoAcid proteins can form amyloid proteins, bind to important proteins (RNA-binding proteins, chaperones, etc), and generally disrupt RNA metabolism, leading to cell dysfunction and death. When genes for polyG proteins were engineered into cell lines, the resulting proteins formed long sequences that aggregated in the nucleus. They also recruited other proteins naturally enriched in glycine. The human protein with the most glycines is FAM98B, a subunit of the enzyme tRNA ligase complex (tRNA-LC). This is used to splice together tRNA exons after intron removal. When polyG proteins

were expressed, they bound FAM98B, leading to nonfunctional tRNA-LC in the nucleus. In mice, this cause severe neurological disorders in the brain.

To be added after resolution of the **Government Shutdown**: FAM98B: AF-Q52LJ0-F1: undefined. https://www.uniprot.org/uniprotkb/Q5...ntry#sequences and iCn3D model

4.7.10.5: Misfolding and Aggregation Summary

Recent work has shown that proteins considered harmless can form misfolded intermediates that aggregate into pre-fibril structures, which are toxic to cells. This process is usually prevented in the cell by the interaction of nascent protein forms with chaperones, which sequester exposed hydrophobic patches and prevent aggregation. (Obviously, prion proteins and the others mentioned above are exceptions.) Amyloid fibers (characterized by subunits with an abnormal amount of beta-structure) can be made from many different types of proteins, as noted above. Is this property specific to just a handful of proteins, or is it more common than expected from the limited examples noted so far? The new studies show that when a bacterial protein, HypF, is incubated at pH 5.5 in the presence of trifluoroethanol, aggregates (but not fibrils) form with enhanced beta structure. These aggregates slowly form into fibrils characteristic of amyloid protein fibers. The early aggregates (formed before fibril formation) were cytotoxic. Similar results were seen with dimers and trimers (prefibril states) of the amyloid-β peptide released from cultured neurons.

A diverse group of proteins that do not share significant secondary or tertiary structures can form amyloid-like protein aggregates. Even though their monomer forms share little in common, the insoluble amyloid aggregates have a common structure in which the monomers within the aggregates exhibit significant beta structure, with the strands running perpendicular to the aggregate axis. Since it has recently been shown that almost any protein, under the "right" set of conditions, can form such aggregates, the stabilizing feature of protein aggregates may be present in any protein. Evidence suggests that the polypeptide backbone, and not the side chains, is key in forming stable interstrand H-bonds in beta-secondary structures in amyloid aggregates. In contrast, native, nonamyloid forms of normal proteins arise through specific interactions between unique side chain sequences and structures, which outcompete nonspecific interactions among backbone atoms found in amyloid structures. Nonspecific aggregation becomes more prevalent when buried hydrophobic side chains and main chain atoms become more solvent-exposed. Such exposure occurs when native proteins form intermediate molten globule states under altered solvent conditions or when destabilizing mutants of the wild-type protein arise. Some mutations may alter protein folding cooperativity, thereby increasing the fraction of non-native protein states. Other mutations that decrease the protein's charge or increase its hydrophobicity might enhance aggregation. In addition, chemical modifications to proteins (such as oxidation or deamination) may destabilize the native state, leading to the formation of the molten globule state. Once formed, this state may aggregate through sequestration of exposed side-chain hydrophobes or inter-main-chain H-bond formation. Aggregate formation appears to proceed through the initial formation of soluble units (which may or may not be more toxic to cells than the final aggregate). Aggregates are kinetically stable species. Since amyloid aggregates are cytotoxic and almost any protein can form them, albeit with different propensities, nature, through evolutionary selection, has presumably disfavored proteins with high tendencies to form such aggregates.

Accurate protein folding is required for cell viability. Aberrant protein folding can clearly cause serious illness. Given the extraordinary nature of the task and its failure, the process governing protein folding must be highly regulated. Figure **4.7.10.20** shows the steps determining intracellular concentrations and locations of normal and aberrant protein structures.

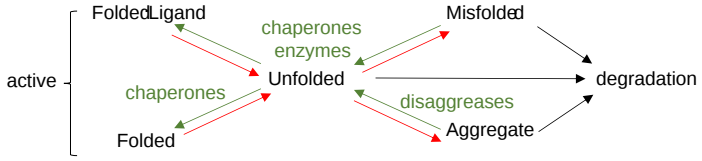


Figure **4.7.10.20**: Conversion of active folded proteins to misfolded and aggregated proteins (after Balch et al, Science, 319, 916, 2008)

Potential therapies for proteostasis diseases include replacing aberrant proteins, shifting the equilibrium toward active forms with small ligands, or modulating pathways with agents that influence processes such as signal transduction, transcription, translation, degradation, and translocation. This can be achieved using molecules such as siRNAs to modulate the levels of chaperones, disaggregases, and signaling pathways.

Recently updated: 10/3/25

4.7.10.6: Functions of Amyloid Proteins in Microbes

Are all amyloid proteins toxic? It depends on the cell and your reference frame. It turns out that both bacteria and fungi can form **biofilms**. These are beneficial to them, but not to infected humans, as these films are often resistant to drug treatment. Biofilms are aggregates, or perhaps better, communities of microorganisms that bind to a surface and encapsulate the enclosed cells within an extracellular matrix (ECM). The films can stick to almost any surface, including rocks, plant and animal tissue, pipes, teeth, and medical implants. If you ever walked on a dock near a lake and felt you were on ice, you were likely "skating" on an algal biofilm. Amyloid proteins from the cells in the biofilm are critical for biofilm formation and structure. They are involved in the initial attachment of cells and in cell:cell interactions. The cells in the biofilm are protected from the effects of environmental chemicals, including antibiotics and immune regulators. They are also resistant to dehydration and pH changes. As biofilms protect internal cells, the reversible conversion between dispersed ↔ enclosed cells is highly regulated. For example, in E. coli, the transcription factor CsgD leads to the synthesis of cellulose and a fibrous structure called the Curli fimbriae, which contain bacterial amyloid proteins. The amyloid protein forms aggregates enriched in β-sheets, which then form fibrils.

An example of a fungal amyloid protein is the **Heterokaryon incompatibility protein S (**HetS) from the filamentous fungus *Podospora anserina.* Fungi can reproduce sexually by releasing spores, which can grow filamentous hyphae as part of their life cycle, as illustrated in Figure 4.7.10.21. The figure shows two spores fusing to form a hyphal network, which, in this case, leads to the growth of a recognizable mushroom.

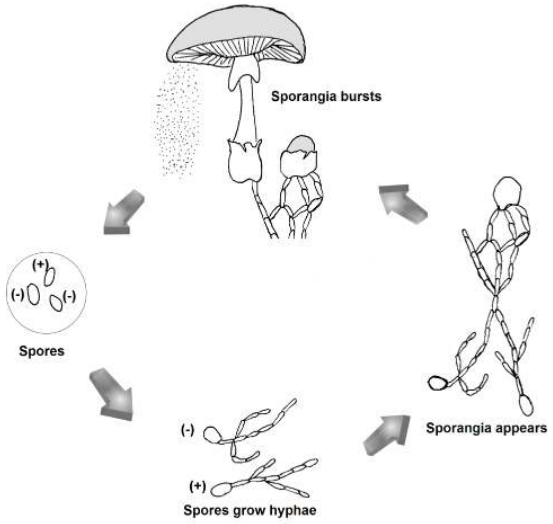
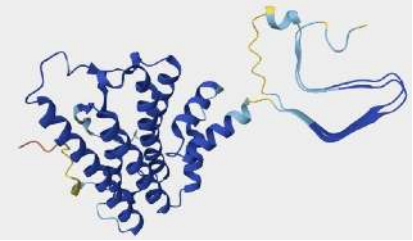
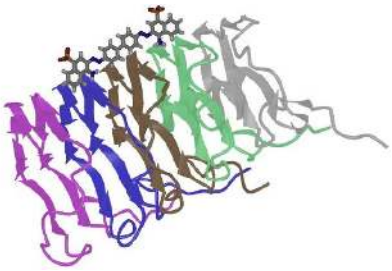


Figure 4.7.10.21:
Fungi reproduction. Fungi_sessuate_reproduction.png: M.violante. https://commons.wikimedia.org/wiki/F...production.png.

There are two forms of the HETS protein from two different alleles: Het-**s**, which can exist in both a soluble "normal" form and a prion amyloid version, and Het-**S,** a soluble form of the protein. Het-S has a few amino acid differences in its N-terminus that prevent it from adopting a prion form on its own.

What prevents two different strains from fusing? HetS is involved. If one strain expresses Het-s and the other Het-S, the strain with Het-s induces cell death in the Het-S strain. The het-s prion interacts with and induces a prion form of Het-S, which leads to conformational changes in Het-S to generate an amyloid β-solenoid form. In addition, the N-terminal of Het-S changes conformation to expose a hydrophobic region that interacts with membranes. At the membrane, Het-S forms amyloid oligomers that puncture it, leading to cell death.

Figure 4.7.10.22 shows interactive iCn3D models of the predicted AlphaFold structure of the soluble form of Het-s (**Q03689)** and the C-terminal amino acids 218-289 prion form (2LBU). The prion aggregate contains five monomers containing amino acids 218-289 of the protein, and is bound to Congo Red, a dye that binds to amyloid fibers.

 **Unavailable during US Government shutdown.** AlphaFold model of the HET-s soluble monomer (AF-Q03689-F1). (Copyright; author via source). Click the image for a popup or use this external link:	 NMR structure of the HET-s amyloid (multimer bound to Congo Red. (2LBU). (Copyright; author via source). Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...9cbf00612c5921

As mentioned earlier, biofilms are beneficial to microorganisms, but not humans. Recently, the mouth bacteria *viridans streptococci*, which are found in dental biofilms, were found in about 60% of coronary artery plaques in autopsies of people who died suddenly. Since the bacteria were in biofilms, they would be resistant to antibiotics and immune mediators. In heart attacks, the plaques rupture, likely releasing the bacteria, which are then recognized by the immune system. This leads to a burst of immune activity from the innate (macrophages) and adaptive (T-cell) immune systems. The sequestration of bacteria in coronary artery biofilms might explain the limited effect of antibiotics in treating cardiovascular disease. These findings link oral health and heart attacks, and suggest that antibiotics might be useful at the onset of a heart attack.

One last twist in the case of biofilms. **Helicobacter pylori** is a stomach bacterium found in a high percentage of people. It has been associated with ulcers and gastric cancer. The CagA protein (**Cytotoxicity-associated immunodominant antigen**) produced by the bacterium is a virulence factor. It's injected into epithelial cells that line the gut. Once in the cell, it alters morphology and signaling, leading to the loss of contact inhibition in epithelial cells. This can promote tumor formation, making the protein a bacterial oncoprotein. The N-terminal of the protein, **CagA^{N}**, is involved in amyloid protein binding.

However, the secreted form has several advantages from both the bacterial and human perspectives. It binds to both amyloid fibril proteins of bacteria (curli, CsgA, *E. coli*) and FapC (*Pseudomonas*) with high affinity and also human amyloid proteins (**Aβ42/Aβ40 peptides** and **tau** in Alzheimer's disease, and **α-synuclein** in Parkinson's disease), inhibiting their fibril formation. It can prevent biofilm formation in bacterial cells and destabilize existing ones. Binding again is high affinity. It affects different stages of amyloid fiber formation for different cell types, as illustrated in Figure 4.7.10.23 below.

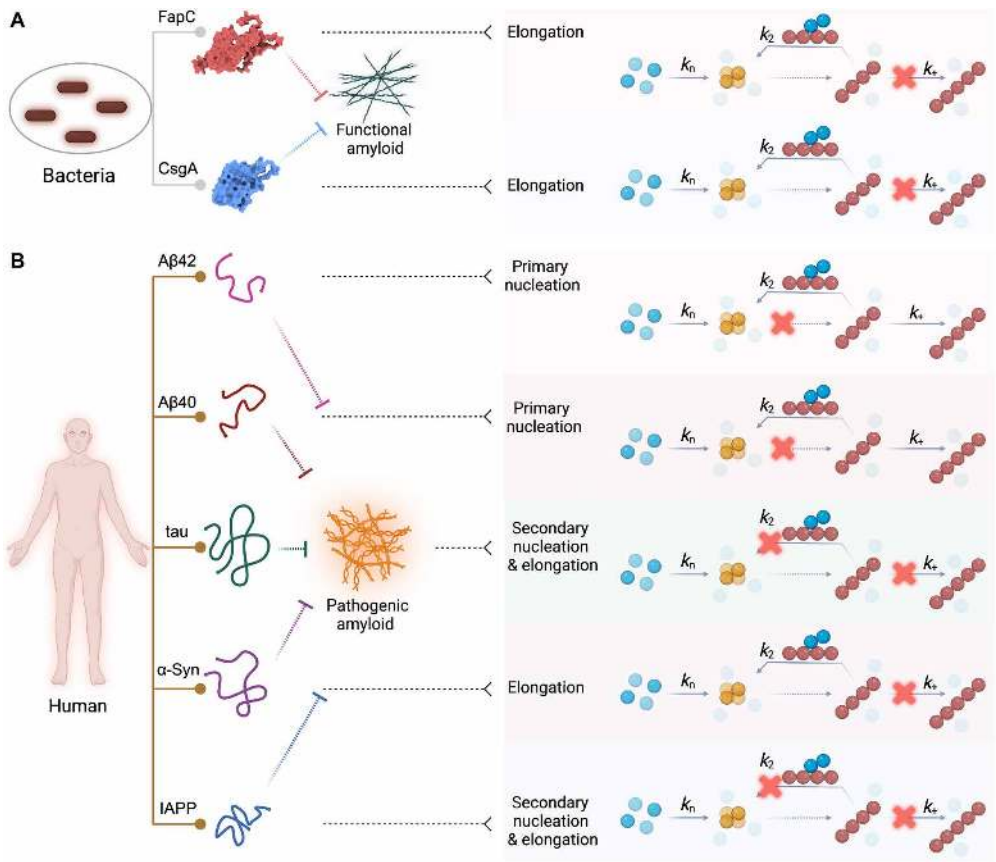


Figure **4.7.10.23**: CagA^N interferes with different microscopic events of the fibril formation process of various amyloid peptides or proteins.

Zhen Jin *et al. Helicobacter pylori* CagA protein is a potent and broad-spectrum amyloid inhibitor. *Sci. Adv.***11**,eads7525(2025).DOI:10.1126/sciadv.ads7525. https://creativecommons.org/licenses/by-nc/4.0/

The kinetics of amyloid aggregation involve distinct stages: In the primary nucleation phase, monomers come together to create a nucleus (kn), from which a fibril can start to elongate (k+); simultaneously, during secondary nucleation (k2), monomers adhere to the fibril's surface, catalyzing the development of a new nucleus and facilitating exponential fibril growth. Refer to the schematic representation in (A) and (B) for a visual depiction of this process. (A) CagA^N suppresses amyloid formation of different bacterial functional amyloid proteins, i.e., CsgA from E. coli and FapC from Pseudomonas. For these amyloidogenic proteins, CagA^N predominantly blocks elongation, as indicated by a red cross. (B) CagA^N impedes the formation of amyloid fibrils by various human pathogenic amyloid peptides or proteins through diverse mechanisms. In the case of AD-associated Aβ42 and Aβ40 peptides, primary nucleation is largely blocked, as CagA proteins likely stabilize oligomers. Conversely, for the AD-relevant protein tau, both secondary nucleation and elongation are affected by CagA^N. Regarding the PD-related protein α-synuclein, CagA inhibits elongation. The IAPP peptide, associated with T2D, is inhibited from fibrillizing by CagAN, affecting both secondary nucleation and elongation.

Figure **4.7.10.24** below shows an AlphaFold 3 model of the interaction of GagA^N and amyloid proteins. GagA^n has three domains, domain I (D1), domain II (D2), and domain III (D3).

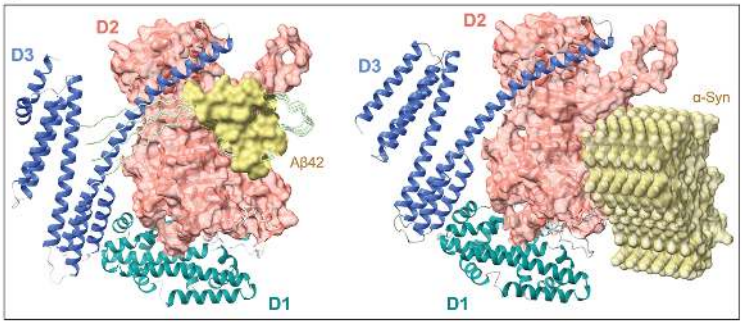


Figure 4.7.10.24: AlphaFold 3 models of the interaction of $GagA^N$ and amyloid proteins. Zhen Jin *et al.*, *ibid.*

AlphaFold 3 prediction of the interactions of $CagA^N$ and different amyloid peptides or proteins, with four Aβ42 subunits mimicking the aggregates (left) and six α-synuclein mimicking the fibrils.

4.7.10.7: Binding, Intracellular Granules and Droplets

The above structures are fascinating aggregates of specific proteins. The aggregates are quite large. In Alzheimer's Disease, they vary from around 150-500 μm^2, giving a length of 12-22 μm if they were squares. In comparison, intracellular "**granules**" are much smaller, with diameters of 200-500 nm (0.2-0.5 μm). The term "granule" describes particles in cells that are barely visible by light microscopy. Granules are found in many cells and mostly contain protein. Platelet granules contain many proteins involved in clotting. Pancreatic beta cell granules contain insulin for secretion. Other types of granules in germ-line cells are dense bodies, perinuclear P granules in Caenorhabditis elegans, germinal granules in Xenopus laevis, chromatoid bodies in mice, and polar granules in Drosophila. It contains both RNA and proteins. These are often referred to as ribonucleoprotein (RNP) granules. Plants and livers also contain starch granules (a carbohydrate). The granules don't appear to be surrounded by a membrane. Rather, they are just aggregates of proteins or of RNA and proteins. Chapters 10 and 11 show analogous particles for lipids, nonpolar "insoluble" molecules that self-aggregate into micelles and membrane **bi**layers. In contrast to the granules mentioned above, lipid droplets containing TAGs and cholesterol esters are surrounded by a phospholipid **mon**olayer with adsorbed protein. Understanding the structure and properties of phase-separated granules may provide insight into the aggregates formed in neurodegenerative diseases.

How do these granules form? What principle underlies the specificity of proteins and RNAs found in them? The aggregates are not toxic, unlike the beta-amyloid aggregates discussed above. Granule formation can be caused by a classic "phase transition/separation", not unlike gaseous water, which can self-associate through hydrogen bonds to form liquid drops, which can freeze with the formation of more hydrogen bonds to form solids. Soluble biomolecules in cells can **reversibly** aggregate through multiple, weak noncovalent interactions to form storage granules. This balance might be perturbed if storage granules aggregate further, potentially through a process that is irreversible and has health consequences, as observed in neurodegenerative diseases. Let's delve into new insights into the processes involved in droplet formation.

Imagine small amounts of a sparingly soluble oil added to an aqueous solution. Initially, it is in solution, but at higher concentration, dipole-dipole interactions, along with the “hydrophobic effect,” would drive the oil out of solution into liquid drops. This phase separation, also known as **liquid-liquid demixing,** occurs when two liquids (solubilized oil in water and separated oil drops) separate. This process has produced various types of non-membrane-bound droplets (distinct from membrane-bound vesicles) within the cell.

Figure 4.7.10.25 below shows the many types of condensates found in cells.

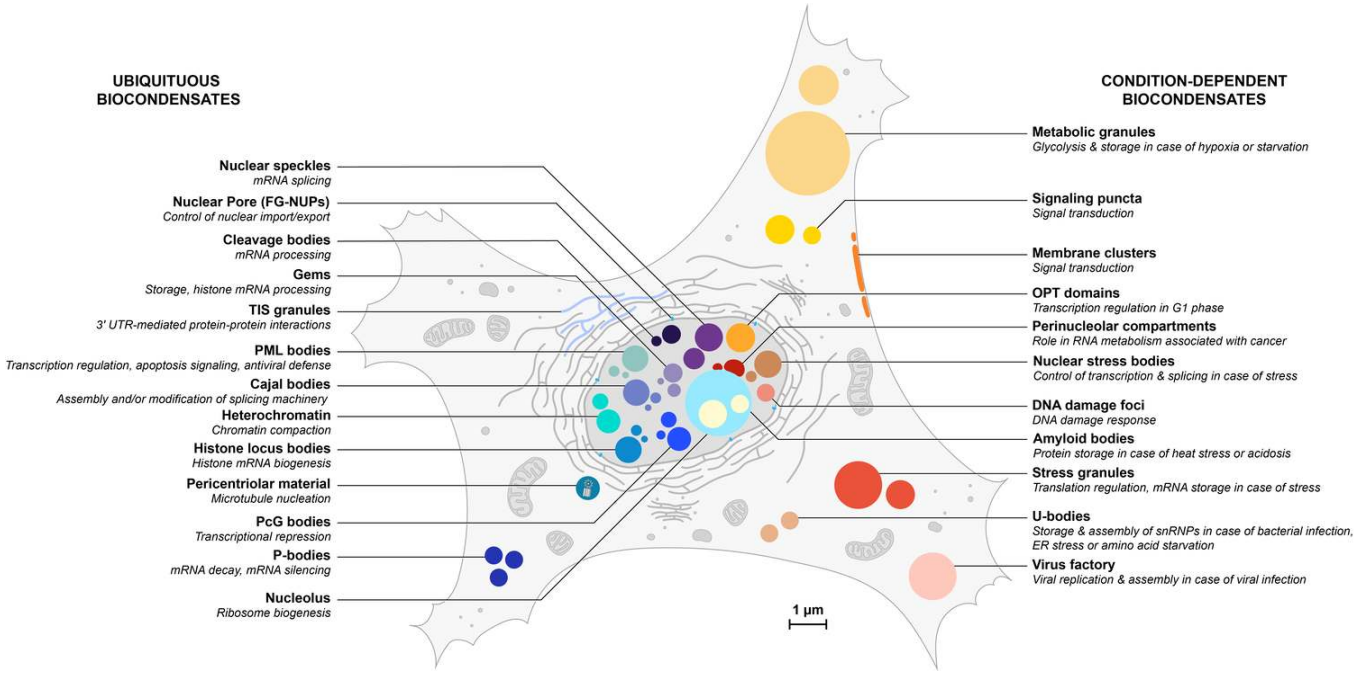


Figure 4.7.10.25: Condensates found in cells. PHASE SEPARATION 101. Animation Lab. Margot Riggi, Janet Iwasa, et al. Creative Commons Licensing CC BY 4.0 - https://creativecommons.org/licenses/by/4.0/

This phenomenon has also been seen with intrinsically disordered proteins and proteins with such domains. These are characterized by amorphous structures composed of repeated, often positively charged amino acids and contain a limited number of distinct amino acids. An example of a protein with a domain that has low sequence complexity is the SP1 transcription factor, a DNA-binding protein. One of its transactivation domains has almost 20% glutamines, with regions within it having even higher percent abundances. It has been estimated that up to 20% of eukaryotic proteins lack a stable shape, as they are, in part, intrinsically disordered and contain **low-complexity domains (LCDs)**. They are found in the N- and C-terminal ends of all mammalian intermediate filament proteins, as well as in almost all RNA-binding proteins, lining the nuclear pore, and on the cytoplasmic faces of mitochondrial, lysosomal, peroxisomal, and Golgi integral membrane proteins. They are the target of up to 3/4s of post-translational modifications. LCDs, therefore, appear to facilitate promiscuous binding of various proteins, especially those that add or remove covalent tags.

Under the right conditions, these can aggregate and “precipitate” from the solution. What is the nature of the precipitate? It may have properties more akin to distinct liquid droplets, so that this process could be referred to as **liquid-liquid demixing**.

Properties of demixed drops include reduced rates of material diffusion into and out of the drop, coupled material movements within the drop, and probable weak hydrophobic-dependent aggregation, which makes drops sensitive to agents such as detergents. Liquid-like diffusion inside the drop is observed, as evidenced by the rapid fluorescence recovery from partially photobleached internal components in the drop.

As with forming a crystalline solid from a liquid solution, the process must be initiated by seeding. This process can be “catalyzed” by poly (ADP-ribose), a nucleic acid-like polyanion that interacts with intrinsically disordered proteins. The negative charges would counterbalance the positive charges in the disordered protein domain, which, without neutralization, would interfere with protein-protein contacts necessary for aggregation/droplet formation and demixing. Aggregation in these cases may arise from hydrophobic interactions (even though hydrophobic side chains are underrepresented in the disordered domains).

- A database, PhaSePro, is dedicated to proteins that drive liquid-liquid phase separation (LLPS) in living cells.

The solubility of proteins in cells is a fascinating topic in its own right. A high ATP concentration (5 mM) in the cell facilitates protein solubilization. ATP is considered a **hydrotrope**. It’s a small molecule with a distinct polar part (polyphosphate and ribose) and a more nonpolar part (the adenosine ring). Hence, it acts like a mini-detergent (an amphiphile), but it doesn’t form micelles. It helps stabilize more nonpolar regions of proteins in solution and has been shown to inhibit aggregate formation and disaggregate

preexisting aggregates. Figure 4.7.10.26 shows a protonated form of energy-minimized ATP with its dipole moment as an arrow from the + to - ends. The dipole moment would only be larger if the ATP was deprotonated and had negative charges.

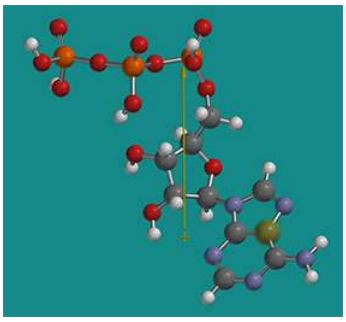


Figure 4.7.10.26: Dipole moment of ATP

Biochemists also use the term gel (examples include polyacrylamide gel or fibrin blood clots, which are chemically cross-linked) and a "gel" form of a bilayer (Chapter 10), when they wish to describe a structure that is neither solid nor liquid. Structures like the cytoskeleton or the actin-myosin network would be examples of the latter.

Noncovalent gels would be characterized by the regulatable dissociation of subunits and, hence, short half-lives. A gel (either covalent or noncovalent) with a high water content is called a hydrogel, as it contains hydrophilic components. An example is RNA- and protein-containing particles.

Here is a link to Phase Separation 101, a wonderful description of phase separations in cells by Janet Iwasa and the Animation Lab. It contains numerous animations illustrating phase separations in chemical and cellular systems, and it describes the underlying chemical and physical processes.

4.7.10.7.1: Granules

Granules that contain RNA and proteins are called ribonucleoprotein bodies (RNPs) or RNA granules. Specific examples include cytoplasmic processing bodies, neuronal and germ granules, and nuclear Cajal bodies, nucleoli, and nuclear dots/bodies. Some granules contain proteins, including inclusion bodies with misfolded and aggregated proteins, as well as those with active proteins involved in biosynthesis, such as purinosome (for purine biosynthesis) and cellulosomes (for cellulose degradation).

The presence of trinucleotide repeats in DNA and RNA can lead to the aggregation of RNAs and proteins encoded by genes containing the repeats. We saw above examples of proteins with extended stretches of glycine and glutamine that aggregate. What about expansions of a charged amino acid like arginine? It likely would not aggregate with itself given charge-charge repulsion. More likely, it would interact with a polyanion, like RNA or DNA.

Arginine (Arg, R) is encoded by six codons: CGU, CGC, CGA, CGG, AGA, and AGG. As with the cases of Gly and Gln, runs of Arg occur using the codons CGG, AGA, and AGG. Short stretches of Arg (5-20) can normally occur, and are used as signals to translocate a translated protein into the nucleus and also ot bind to RNA. These functional proteins do not aggregate abnormally. However, proteins produced by RAN synthesis from expanded repeats can form aberrant proteins that damage cells. One example is the RAN synthesis from the hexanucleotide repeats of GGGGCC (G4C2) found in the first intron the protein *C9orf72* gene, resulting in a protein hghly enriched in arginine.

This gene is the most common genetic cause of Lou Gehring's Disease or amyotrophic lateral sclerosis (ALS) and also of frontotemporal dementia (FTD). This gene lacks an AUG start codon and can be read to produce five different dipeptide repeat (DPR) proteins, the most common being poly-GR and poly-PR, both of which are highly toxic. Since they have a high positive charge, they are likely disordered. They bind readily to RNA (negatively charged) and nucleoli. Hence, they impact transcription, splicing, and the formation of the membraneless organelle, the nucleolus, which is the site of ribosome formation for normal protein synthesis. They can lead to phase separation of RNA and other RNA-binding proteins, forming stress granules. They are toxic even at low levels. In Fragile X syndrome, 230-4000 repeats of the **CGG** triplet occur in the noncoding parts of the genome, compared to less than 50 in the normal gene.

Trinucleotide expansion occurs in the non-protein-coding intronic DNA, which also leads to trinucleotide repeats in the RNA spliced out of the transcribed RNA. For example, a CTG DNA repeat in an intron leads to a poly (CUG) RNA, which, through unusual base pairing, can form RNA aggregates, as seen in myotonic dystrophy.

In vitro experiments show that small complexes are soluble. Still, as the size increases, liquid-liquid demixing (or a liquid-gel transition) can occur, forming spherical droplets of RNA particles. This would explain the observation that pathologies occur above a certain repeat length. If misfolded proteins are also present, these particles might combine to form larger gels.

In the control experiment, demixing and spherical particle formation were not observed when the repeats were scrambled. In an experiment similar to adding 1,6-hexanediol to intrinsically disordered proteins, small antisense trinucleotide repeats, such as (CTG)8, which could interfere with the weak H bonds between G and C in the aggregates, were added. The size of RNA drops (foci) was reduced. In vivo experiments showed characteristic drop-like structures only when the repeats were sufficiently large.

Researchers found that in vitro, RNA drop formation was inhibited by monovalent cations. In the presence of 0.1 M ammonium acetate, which permeates cells without affecting pH, CAG RNA droplets in vitro disappeared.

Aggregation of mRNA may be one way to regulate its translation and, in turn, indirectly regulate gene activity. There are advantages to regulating the translation of a protein from mRNA, especially if the mRNA's "activity" can be dynamically regulated. This would be useful if new protein synthesis were immediately required. Hence, reversible aggregation is one way to regulate mRNA activity (in addition to degradation).

4.7.10.7.2: Protein drops and granules

The cytoskeletal proteins actin and tubulin (a heterodimer of alpha and beta chains) can exist in either soluble or condensed, filamentous states (as actin filaments and microtubules, respectively). GTP hydrolysis is required for the formation of tubulin polymers. Actin binds ATP, which is necessary for filament formation, but ATP cleavage is required for depolymerization. Hence, nucleotide binding/hydrolysis regulates the filament equilibrium, unlike simple phase changes in water.

Since only certain proteins form granules, they must have similar structural features that facilitate reversible binding interactions. These proteins have multiple, weak-binding sites, but if they act collectively, they provide multivalent (multiple) binding interactions that allow robust but not irreversible granule formation. Here are some characteristics of proteins found in granules:

- the protein NCK has three repeated domains (SH3) that bind to proline-rich motifs (PRMs) in the protein NWASP. These proteins are involved in actin polymerization. In high concentrations, they precipitate from the solution and coalesce to form larger droplets;
- repeating interaction domains are widely found, especially among RNA-binding proteins;
- some proteins contain Phe-Gly (FG) repeats separated by hydrophilic amino acids in portions of the protein that are intrinsically disordered.
- a biotinylated derivative of 5-aryl-isoxazole-3-carboxyamide (Figure **4.7.10.27**) precipitates proteins, which are enriched in those that bind RNA (RBPs). The precipitated proteins were generally intrinsically disordered and characterized by low-complexity sequences (LCS). One such example contained 27 repeats of the tripeptide sequence (G/S)Y(G/S). The proteins could also form hydrogels (made of hydrophilic polymers and crosslinks) and transition between the soluble and gel phases, with extensive hydrogen-bond networks. The hydrogel gel phase exhibited an X-ray diffraction pattern similar to that of beta-structure-enriched amyloid proteins. Short-range, weak interactions between LCS may then drive reversible condensation into gel-like granule states, characterized by extensive hydrogen bonding (similar to that during ice formation). If this process goes awry, more continued and irreversible formation of a solid fibril (as seen in neurodegenerative diseases) might occur from the hydrogel state.

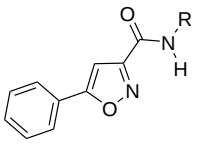


Figure **4.7.10.27**: Structure of 5-aryl-isoxazole-3-carboxyamide

- RNAs form granules when proteins bind to them via RNA-binding domains that interact via low-complexity sequences, leading to phase separation and the formation of hydrogel-like granules. Around 500 RNA-binding proteins have been found in the

human RNA interactome. They are enriched in LCSs and have more tyrosines than average proteins in the entire proteome, in which tyrosines are often found in the (G/S)Y(G/S) motif. Phosphorylation of tyrosines (Y) in LCS may decrease association and hydrogel stability.

Given that many neurodegenerative diseases are associated with unfolded/misfolded protein aggregates, the high protein concentrations in protein-containing liquid drops may pose cellular problems. If sufficiently high, the equilibrium might shift from a liquid drop to a solid precipitate, with severe cellular consequences. The progression to the solid state may irreversibly affect the cell.

4.7.10.8: Low complexity domains (LCD) and neurodegenerative disease

The aggregation of alternatively folded proteins is associated with neurodegenerative disease. Mutations that cause disease are associated with the formation of low-complexity domains and aggregates, a phenomenon increasingly described as phase separation. The demixed phases are stabilized by interchain backbone hydrogen bonds, as shown in the many beta-sheet aggregates described above. Evidence suggests that labile structures with the potential for interchain H-bonds and beta-strand formation lead to fibril formation. The nascent interactions would involve short stretches of interchain H bonds. If so, mutations that enrich such nascent structural interaction would promote fibril formation, while those that inhibit the nascent interactions would inhibit fibril formation. A study (Zhou et al, Science, 377, 2022. DOI: 10.1126/science.abn5582) verifies this.

The investigators created single amino acid variants of the low-complexity domains of an RNA-binding protein, TDP-43 RNA, that prevented the single amino acid within a region involved in interchain beta strand formation from forming a hydrogen bond via its amide hydrogen. They achieved this by methylating the single main-chain amide nitrogen, which prevents its participation in hydrogen bonding. The modification is shown in Figure **4.7.10.28**.

Figure **4.7.10.28**: Methylation of a single backbone nitrogen in a region involved in interchain hydrogen bond and beta sheet formation in the low complexity domain of proteins

Of the 23 variants they made, 9 were within a continuous stretch, inhibiting phase separation. As determined by cryo-EM, these nine were at the same sites as hydrogen bonds between adjacent chains of the TDP-43.

Next, they looked at other proteins with low-complexity domains that form aggregates/polymers. The proteins examined were the neurofilament light (NFL) chain protein, the microtubule-associated tau protein, and the heterogeneous nuclear RNA-binding protein A2 (hnRNPA2). They identified 10 mutations in LCDs that are known to be associated with neurological diseases. Indeed, these mutations allow for an extra hydrogen bond in low-complexity domain sequences and enhance aggregate formation, presumably through the extra interchain H bond. Specifically, the known mutations replace individual prolines. This cyclic amino acid lacks an amide hydrogen and, therefore, cannot donate a hydrogen bond to another amino acid, which allows for one additional hydrogen bond. Each known mutation was associated with neurological disease and increased stable aggregate/polymer formation. This increased aggregation/polymer (phase separation) was reversed in vitro by chemical methylation of the single amino acid change in the mutant, which prevented it from forming hydrogen bonds. Site-specific methylation was performed by linking synthetic peptides containing the single Nα-methyl amino acid to other synthesized peptides that comprise the protein. The semisynthetic NFL protein, for example, was incubated under conditions conducive to the assembly of mature intermediate filaments.

In vitro experiments were conducted using different synthetic head domains of the neurofilament light (NFL) chain protein, in which the P8 residue contained a different amino acid at that position. Figure **4.7.10.29** shows the variant amino acids.

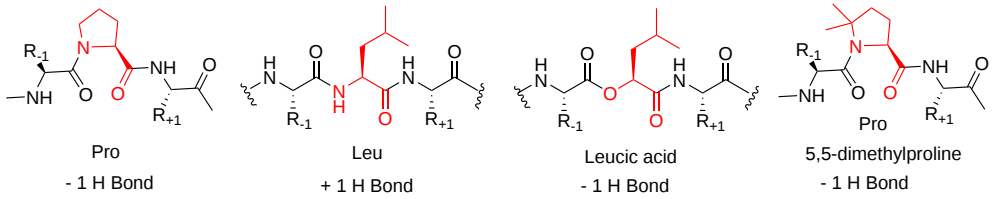


Figure 4.7.10.29: Variant amino acids used at position P8 in the low complexity head domain of the neurofilament light (NFL) chain protein (after Zhou et al, ibid)

Only variants containing Leu at position P8 could form filaments, as measured in vitro using fluorescent assays. Experiments like this are crucial in determining whether phase separation/aggregation causes are causative or merely correlated with the development of complex neurodegenerative diseases.

4.7.10.9: Summary

Protein aggregation is a critical phenomenon that occurs when certain proteins, rather than folding into a single, unique native conformation, adopt alternative, often misfolded structures that self-assemble into larger, insoluble aggregates. Although many proteins naturally form functional oligomers and filaments through specific, reversible interactions, a subset of proteins—including prions and amyloid-forming proteins—can undergo a drastic conformational change that transforms predominantly α-helical or random coil structures into β-sheet–rich forms. This conversion exposes new hydrogen-bonding surfaces that drive the templated, self-propagating formation of fibrils.

Such amyloid fibrils are characterized by a common cross-β architecture, in which β-strands from individual monomers align perpendicularly to the fibril axis and are stabilized by extensive intermolecular hydrogen bonding. Although the monomeric forms of these proteins may appear benign, the aggregated state is associated with several devastating neurodegenerative diseases such as Alzheimer's, Parkinson's, and prion diseases. In these conditions, the aggregates can interfere with normal cellular functions, ultimately leading to cell death.

Research has revealed that the aggregation process is not random; it can be influenced by specific amino acid sequences, point mutations, and post-translational modifications that either promote or inhibit the conversion to the aggregated state. Moreover, while intracellular protein quality control mechanisms—such as molecular chaperones and the unfolded protein response—normally prevent harmful aggregation by assisting correct folding and targeting misfolded proteins for degradation, these systems can become overwhelmed under stress conditions or due to genetic mutations.

Recent structural studies using techniques such as cryo-electron microscopy and NMR have provided detailed models of amyloid fibrils from various proteins, revealing surprising similarities in their cross-β structures despite differences in the native states of the precursor proteins. In addition, emerging evidence on phase separation reveals that some cellular condensates and RNA granules form through reversible, weak multivalent interactions—a process that, when dysregulated, may also contribute to pathological aggregate formation.

Overall, this chapter emphasizes the delicate balance between protein folding, aggregation, and cellular proteostasis. It highlights how understanding the molecular mechanisms of aggregation—from initial misfolding events to the formation of toxic fibrils—can inform therapeutic strategies to prevent or reverse protein aggregation in neurodegenerative diseases.

This page titled 4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.11: Biomolecular Visualization - Conceptions and Misconceptions

Learning Goals (ChatGPT o1, 1/27/25)

Below is a series of learning goals designed for junior and senior biochemistry majors based on the provided information:

1. **Relate Structure to Function and Reactivity:**
   - Explain how the three-dimensional structure of a molecule dictates its physical properties (melting point, boiling point, solubility) and chemical reactivity (acid–base behavior, redox properties, nucleophilicity/electrophilicity).
   - Connect classical representations (e.g., Lewis structures) with modern molecular orbital theory to describe bonding and reactivity.
2. **Differentiate Visualization Representations:**
   - Identify various molecular renderings (line, stick, ball-and-stick, spacefill, surface, and cartoon) and articulate the type of structural and functional information each provides.
   - Evaluate the strengths and limitations of each representation in conveying details such as atom connectivity, bond angles, steric hindrance, and molecular surface features.
3. **Interpret Structural Data from Experimental Methods:**
   - Describe how techniques like X-ray crystallography, NMR spectroscopy, and cryo-EM are used to determine biomolecular structures and discuss the inherent limitations (e.g., static snapshots, missing hydrogen atoms).
   - Explain how computational visualization tools are essential for interpreting and “fleshing out” structural data, especially for large and complex biomolecules.
4. **Apply Concepts from Small Molecules to Macromolecules:**
   - Use the example of oleic acid to illustrate how small molecule renderings can inform our understanding of molecular shape, polarity, and function.
   - Extend these ideas to proteins (e.g., titin, superoxide dismutase) by discussing how different renderings help reveal features like hydrophobic cores, active sites, and metal ion binding.
5. **Analyze Functional Implications of Surface Properties:**
   - Interpret electrostatic potential surfaces to understand how charge distribution contributes to enzyme function, such as the attraction of a negatively charged substrate (e.g., superoxide in superoxide dismutase).
   - Discuss how the visualization of metal ions and co-factors (e.g., Cu and Zn in superoxide dismutase) is critical for deducing catalytic mechanisms.
6. **Integrate Static and Dynamic Views of Protein Structure:**
   - Appreciate that while crystallographic structures offer static images, protein dynamics (bond stretching, angle bending, torsional rotations) are vital for understanding real-life function.
   - Recognize the role of computational methods (like molecular dynamics) in bridging the gap between static structures and dynamic behavior.
7. **Critically Evaluate Structural Models and Data:**
   - Assess potential sources of error in structural models (e.g., missing hydrogen atoms, steric clashes) and consider strategies to resolve these issues in both experimental and computational settings.
   - Discuss how multiple visualizations provide complementary insights and can prevent misconceptions about molecular structure and function.

By mastering these learning goals, students will be able to interpret various molecular renderings critically, understand the experimental and computational methods used to derive them, and apply these insights to explain how molecular structure underpins biological function.

Structure determines everything in biology and chemistry. Since you learned to represent molecules with Lewis structures, it's been drilled into you that the structure of a molecule determines its physical and chemical properties. Physical properties would include melting points, boiling points, and solubility. In contrast, chemical properties include acid/base, redox, precipitation, and general chemical reactivity determined by the presence of Lewis/Brønsted acids/bases and nucleophiles/electrophiles. More modern analyses of reactivity would include molecular orbital theory descriptions of bonding.

We can't see molecules but make inferences from data (x-ray crystallography, NMR spectroscopy, and cryo-EM) about the structure of a molecule (atom type, atom/bond connectivity, and geometry). As molecules get larger (consider the muscle protein titin, also called connectin, with a molecular weight of around 3.8 million), we must use computer visualization to understand its structure and infer its resulting function and activity. As with small molecules, we can render the molecule in different ways to better understand the attributes that confer function and activity. We ask students to view a biomolecule and infer its properties from the rendering without giving them devoted attention and instruction on how to do that.

4.7.11.1: Small molecules

Let's start with a small molecule like oleic acid, a long chain carboxylic acid with 18 carbon atoms and one cis (Z) double bond between carbons 9 and 10, as shown in Figure **4.7.11.1**.

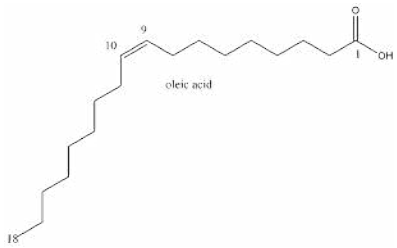


Figure **4.7.11.1**: Line drawing of oleic acid

Table **4.7.11.1** below shows multiple ways to render the molecule. Each rendering offers insight into the molecule's function/activity, but might, at the same time, leave students struggling to interpret it and reinforce or instill misconceptions. Each representation below shows the very same molecule. The top row shows representations without H atoms; the bottom row shows those with H atoms.

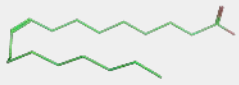
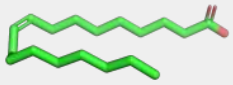
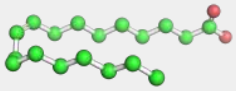
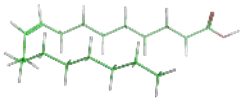
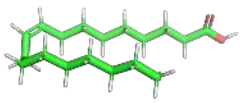
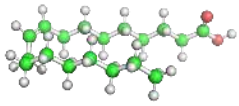
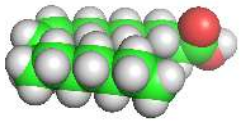
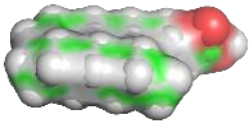
line	stick	ball sphere	sphere	surface
				
				

Table **4.7.11.1**:Different renderings of oleic acid

Here are some important things to remember about biomolecular structures, including both small and large molecules:

- Structures obtained using X-ray crystallography are constructed from relative electron densities calculated from diffraction patterns. Computer programs calculate the structure based on these electron density maps and known bond lengths, bond angles, and atom types. Most of the structures used in this book are available in the Protein Data Bank. The structures seen in computer models are visualized data and may contain errors (missing atoms, steric clashes, incorrect atom assignments), although structural refinement techniques minimize such problems.
- Structures derived from X-ray and cryomicroscopy analyses are static and represent only one conformation from a large ensemble. As you learned from studying simple molecules in organic chemistry, bond length and angle changes continually occur within molecules. Bonds connecting two atoms can stretch, angles connecting three atoms can bend, and the torsional

angle around the center bond in a four-atom, three-bond system can rotate to form eclipsed and staggered (gauche and anti) conformers.

- PDB structures obtained by X-ray crystallography contain no H atoms as they are too small and contain too few electrons to diffract/scatter X-rays. So get used to adding them in your mind when you see a structure. Programs are used to calculate and show H bonds between a slightly positive H atom on an O or N atom in a protein and another slightly negative Os or Ns on the same or different molecule. The H bonds are often shown between N and O atoms. You should look at the atoms involved and the distance between them, and visualize a hydrogen atom connected to one of them. Figure 4.7.11.2 shows an example of an H bond between two base pairs in a DNA molecule.

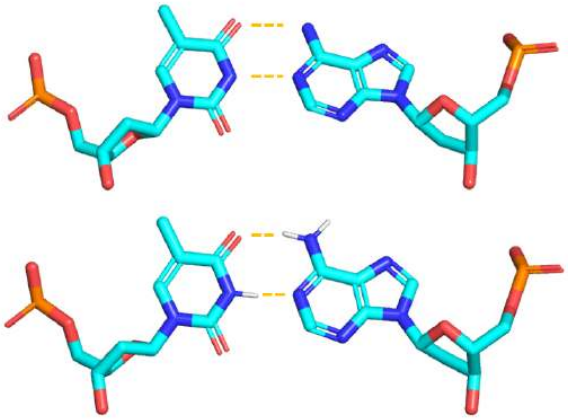


Figure 4.7.11.2: Hydrogen bonds between two base pairs of DNA, with one showing actual hydrogen atoms

- Double bonds or likely resonance structures are not typically shown in PDB structures.
- Line, ball-and-stick, and stick renderings are useful for showing atom connectivity and bond angles. However, they are not particularly useful for showing how atom size might affect a molecule's structure and properties. This type of information is better demonstrated when spacefill renderings that show the sizes of the atoms (based on their Van der Waals radii) are used or when the surface of the molecule, calculated from the contact surface created between the van der Waals surface of the atoms and a rolling probe (often an O atom mimicking water) is displayed.

4.7.11.2: Large molecules

As molecules get larger, line, stick, and ball-and-stick representations become increasingly useless. New ways of visualizing the molecular structure are needed. The importance of multiple renderings to clarify structure/function relationships becomes apparent when you wish to understand protein structure. Various renderings of the protein superoxide dismutase (2sod) are shown in Table 4.7.11.2 below.

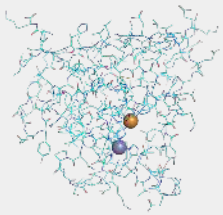
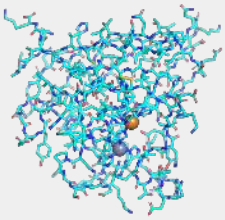
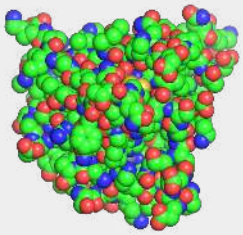
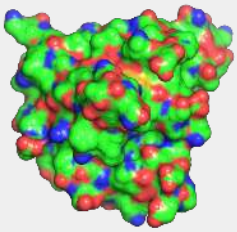
line	stick	cartoon	sphere	surface
				

Table 4.7.11.2: Multiple renderings of the protein superoxide dismutase

- The same features and limitations described above for small molecules apply to large ones. It is best to omit most atoms and use mixed renderings within a single display to reveal the biomolecule's crucial structural features. The cartoon representation of superoxide dismutase shows a single tiny alpha helix (red) and many beta strands (yellow). All side chain and backbone atoms

have been removed. The green line traces the backbone of amino acids not involved in secondary structure. The Cu and Zn ions are shown as spheres.

Another type of surface rendering, the electrostatic potential surface, is beneficial. Figure 4.7.11.3 shows the electrostatic potential surface of superoxide dismutase taken from two different angles after simply rotating the protein.

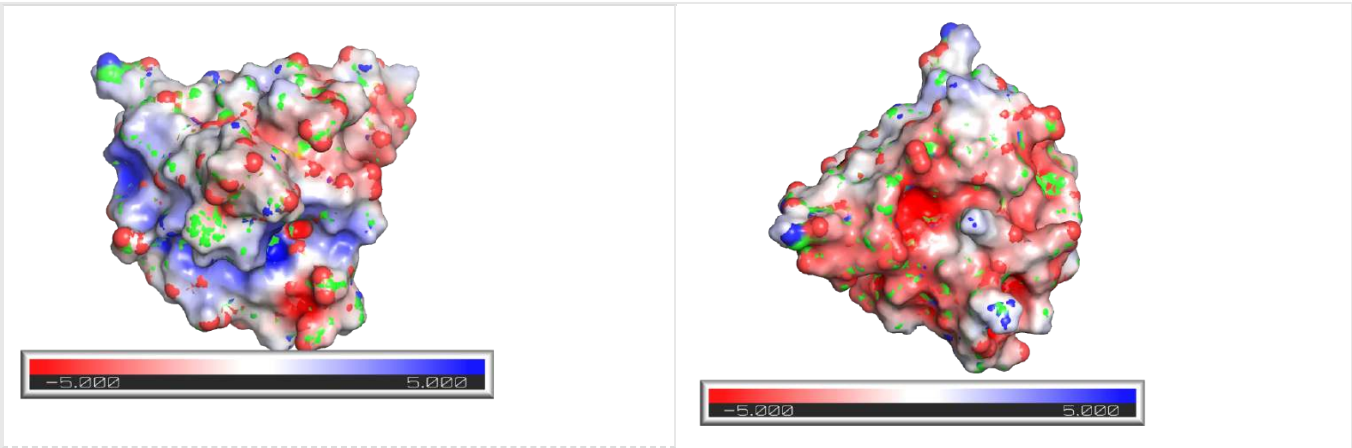


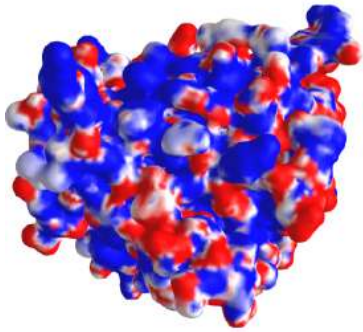
Figure 4.7.11.3: Electrostatic potential surface of superoxidase dismutase from two perspectives

The red represents minimal (most negative) potential. This part of the structure would be enriched in slightly negative Os and Ns or wholly negative Os (i.e., have the highest electron density). Blue represents the positive potential centered in regions containing a slight or full positive charge (I, the lowest electron density). This enzyme binds superoxide, O_2^-, a toxic free radical reduction product of dioxygen, O_2. It catalyzes this reaction:

$$2\,O_2^- + 2\,H^+ \longrightarrow O_2 + H_2O.$$

The enzyme can effectively scavenge superoxide in its vicinity as the negative superoxide is drawn into the active site with the Cu and Zn atoms by the positive potential surrounding the active site, enhancing the typical diffusion encounter rate of the reactant with the enzyme pocket.

Figure 4.7.11.4 shows an interactive iCn3D model of the electrostatic potential of superoxidase dismutase (2sod).



NCBI iCn3D Figure 4.7.11.4: Electrostatic potential of superoxidase dismutase (2sod) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...QQy2hAXehWLBW8

4.7.11.3: Summary

In this chapter, we explore the foundational concept that “structure determines function”—a principle that spans from small molecules to large proteins. We begin with the idea that a molecule’s Lewis structure dictates its physical properties (melting/boiling points, solubility) and its chemical reactivity (acid/base, redox, precipitation). Modern approaches, including molecular orbital theory, further refine our understanding of chemical bonding.

Because we cannot directly see molecules, scientists rely on experimental techniques such as X-ray crystallography, NMR spectroscopy, and cryo-EM to infer molecular structures. These methods provide detailed information about atom types, connectivity, and geometry. As the complexity of molecules increases—as in the case of gigantic proteins like titin (with a molecular weight ≈ 3.8 million)—computer visualization becomes essential for understanding how structure influences function.

For small molecules like oleic acid, various rendering techniques (line drawings, stick models, ball-and-stick, spacefill, and surface representations) highlight different aspects of molecular structure. For example, line models emphasize connectivity and bond angles, while spacefill models reveal the size and shape of atoms as defined by their van der Waals radii. These diverse visualizations help us infer how the structure might influence the molecule's function and reactivity.

When we turn to larger biomolecules, such as the enzyme superoxide dismutase (SOD), additional rendering techniques become useful. Cartoon renderings, which simplify the protein by showing its secondary structure elements (like alpha helices and beta strands), help reveal the overall folding pattern. In contrast, electrostatic potential surfaces illustrate the distribution of charge over the protein's surface, thereby shedding light on how SOD attracts and binds its substrate—superoxide—via complementary charge interactions.

Overall, this chapter underscores that while experimental structures are invaluable, they are snapshots of dynamic systems. Each rendering provides distinct insights into the molecular determinants of biological function, underscoring that the interplay between structure and dynamics is central to understanding both small-molecule and macromolecular behavior in biochemistry.

This page titled 4.7.11: Biomolecular Visualization - Conceptions and Misconceptions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.11: Biomolecular Visualization - Conceptions and Misconceptions** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation

Learning Goals (ChatGPT o1, 1/27/25)

1. **Comprehend Experimental Methods:**
   - Explain how various techniques (UV absorbance, fluorescence, circular dichroism, and viscosity measurements) are used to monitor protein denaturation as a function of temperature or denaturant concentration.
   - Describe what each observable (e.g., A_{280}, fluorescence intensity) indicates about the protein's conformational state.
2. **Quantitative Analysis of Denaturation Curves:**
   - Derive and interpret the equation $y=f_N y_N+f_D y_D$ and show how to calculate the fraction denatured (f_D) using the equation $f_D=y-y_N y_D-y_N$.
   - Relate the fractional denaturation to the equilibrium constant for the unfolding reaction: $K_{eq}=f_D/(1-f_D)$
3. **Thermodynamic Calculations of Protein Stability:**
   - Use the relationship $\Delta G^0=-RTlnK_{eq}$ = $-RTlnK[f_D/(1-f_D]$ to calculate the standard free energy of unfolding from experimental data.
   - Explain how ΔG^o changes with temperature and denaturant concentration, and describe what conditions yield $\Delta G^o = 0$ (the melting temperature, T_m).
4. **Understanding Denaturant Effects:**
   - Interpret the linear relationship $\Delta G^0=\Delta G_w^0-m[urea]$ and explain how the extrapolation to [urea] = 0 gives the free energy of unfolding in water (ΔG_w^0).
5. **Thermal Denaturation and van't Hoff Analysis:**
   - Analyze thermal denaturation data to construct a van 't Hoff plot (lnK_{eq} versus 1/T) and extract ΔH^o and ΔS^o from the slope and intercept, respectively.
   - Compare van 't Hoff enthalpy (ΔH_{vHoff}) with calorimetric enthalpy (ΔH_{cal}) obtained from differential scanning calorimetry.
6. **Critical Evaluation of Data and Models:**
   - Evaluate the assumptions behind a two-state model for protein unfolding and discuss the potential presence of intermediates.
   - Understand the limitations of extrapolating data (e.g., long extrapolation of ΔG^o to zero denaturant and infinite temperature considerations in van 't Hoff analysis).
7. **Application to Protein Mutagenesis:**
   - Explain how changes in the calculated ΔG^o (or $\Delta\Delta G$) for protein mutants can be used to assess alterations in protein stability under physiological conditions.

These learning goals aim to equip you with the conceptual and quantitative tools needed to analyze protein stability through experimental denaturation data and thermodynamic principles.

4.7.12.1: Introduction

Multiple methods can be used to investigate protein denaturation. These include UV, fluorescence, CD, and viscosity measurements. In all these methods, the dependent variable (y) is measured as a function of the independent variable, which is often temperature (for thermal denaturation curves) or denaturant concentration (such as urea or guanidine hydrochloride). From these curves, we want to calculate the standard free energy of unfolding (ΔG^O) for the protein (for the reaction N ↔ D). It is relatively easy to calculate if the denaturation curves show a sigmoidal, cooperative transition from the native to the denatured

state, indicating a two-state transition. The dependent variable can also be normalized to show fractional denaturation (fD). An idealized denaturation curve is shown below in Figure 4.7.12.1.

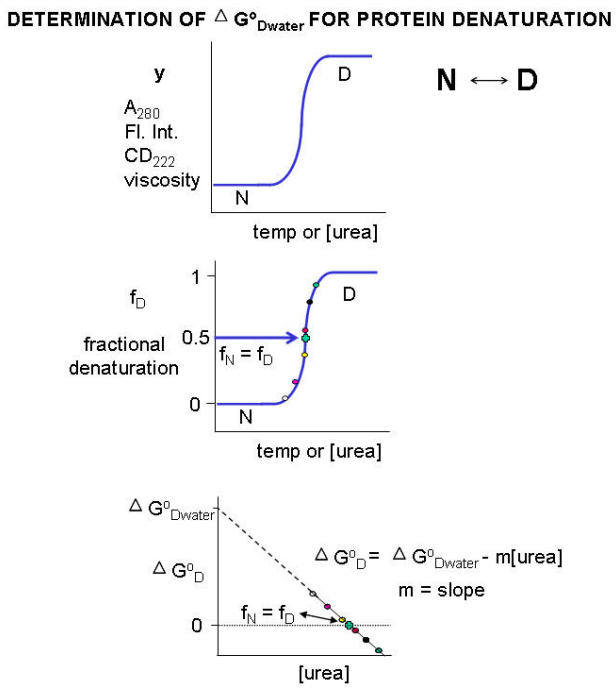


Figure 4.7.12.1: Graphical analysis of protein denaturation

A more realistic denaturation curve might show a small linear change in the dependent variable (fluorescence intensity, for example) with temperature or denaturant concentration, well before the major unfolding transition and above the concentration at which it is unfolded. In these cases, the mathematical analyses presented at the end are required.

4.7.12.2: Denaturation with urea or guanidine hydrochloride

For each curve, the value of y (either A_{280}, fluorescence intensity, viscosity, etc.) can be considered the sum contributed by the native state and the denatured states, which are both present in different fractional concentrations from 0 - 1. Hence, the following equation should be reasonably intuitive.

$$y = (f_N y_N) + (f_D y_D) \tag{4.7.12.1}$$

where f_N is the fraction native and y_N is the contribution to the dependent variable y from the native state, and f_D is the fraction denatured. y_D is the contribution to the dependent variable y from the denatured state. Conservation gives the following equation.

$$1 = f_N + f_D \text{ or } f_N = 1 - f_D \tag{4.7.12.2}$$

Substituting (2) into (1) gives

$$y = (1 - f_D)\, y_N + (f_D y_D) = y_N - f_D y_N + (f_D y_D) \tag{4.7.12.3}$$

Rearranging this equation gives

$$\mathrm{f_D} = \frac{\mathrm{y} - \mathrm{y_N}}{\mathrm{y_D} - \mathrm{y_N}} \tag{4.7.12.4}$$

Notice the right-hand side of the equations contains variables that are easily measured.

By substituting (2) and (3) into the expression for the equilibrium constant for the reaction N ↔ D we get:

$$K_{eq} = \frac{[D]_{eq}}{[N]_{eq}} = \frac{f_D}{f_N} = \frac{f_D}{1 - f_D} \tag{4.7.12.5}$$

From this, we can calculate ΔG^0.

$$\Delta G^0 = -R\,T\ln\ K_{eq} = -R\,T\ln\left[\frac{f_D}{1-f_D}\right] \tag{4.7.12.6}$$

Remember that ΔG^O (and hence K_{eq}) depends only on the intrinsic stability of the native vs denatured state for a given set of conditions. They vary as a function of temperature and solvent conditions. At low temperatures and low urea/guanidine HCl concentration, the native state is favored, and for the N ↔ D transition, $\Delta G^O > 0$ (i.e., denaturation is NOT favored). The denatured state is favored at high temperatures and urea/guanidine-HCl concentrations and $\Delta G^O < 0$.

At some temperature or urea/guanidine concentration value, both the native and denatured states would be equally favored. At this value, $K_{eq} = 1$ and $\Delta G^O = 0$.

The temperature at the maximal C_p value is called the protein's melting temperature (Tm). It is analogous to the T_m in the heat capacity vs temperature graphs for protein denaturation, as we saw in Chapter 4.9. This is illustrated in Figure **4.7.12.2**.

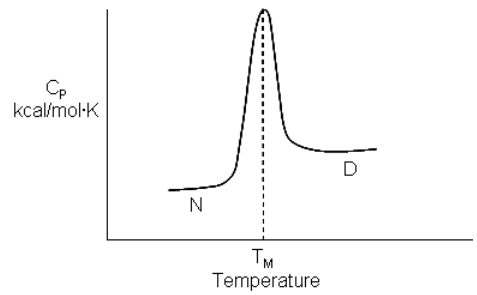


Figure **4.7.12.2**: T_m for thermal denaturation of a protein

Ordinarily, at a temperature much below the T_m for the protein or at a low urea concentration, so little of the protein would be in the D state that it would be extremely difficult to determine the protein concentration in the D state. Hence, it would be difficult to determine the K_{eq} or ΔG^o for the reaction N ↔ D. However, in the range where the protein denatures (either with urea or increasing temperature), it is possible to measure f_D/f_N, and hence ΔG^o at each urea or temperature.

Now, we can calculate the $\Delta \mathbf{G}^o{}_w$ for N ↔ D in water without urea. For a simple two-state N ↔ D, a plot of ΔG^o vs [urea] is linear and given by the following equation, which should be evident from (\PageIndex{1}\).

$$\Delta G^0 = \Delta G^0_w - m[\text{ urea }] \tag{4.7.12.7}$$

It is important to know the K_{eq} and ΔG^0 for the N ↔ D transition in the absence of urea and under "physiological conditions." A comparison of the calculated $\Delta G^0{}_w$ values in the absence of urea for a series of similar proteins (such as differing by a single amino acid, prepared by site-specific mutagenesis of the normal or wild-type gene) would indicate to what extent the mutants were stabilized or destabilized compared to the wild-type protein. $\Delta G^0{}_w$ for the N ↔ D transition of the protein in the **absence** of denaturant (i.e., in water) can be determined by extrapolating the straight line to [urea] = 0. Admittedly, this is a long extrapolation, but reasonable values can be obtained with high-quality data and a high correlation coefficient for the linear regression analysis of the best-fit line.

4.7.12.3: Denaturation with heat

Calculation of $\Delta \mathbf{H}^o$ and $\Delta \mathbf{S}^o$ for N <=>D at room temperature

K_{eq} values can be calculated from thermal denaturation curves, as described above, using urea as the denaturant by monitoring the change in an observable (e.g., spectral signal) vs. temperature. Knowing K_{eq} and ΔH^0, DS^0 can be calculated from the equation below since a semi-log plot of $\ln K_{eq}$ vs 1/T is a straight line with a slope of - $\Delta H^0 R$ and a y-intercept of + $\Delta S^0/R$, where R is the ideal gas constant.

$$\begin{aligned} \Delta G^0 = \Delta H^0 - T\Delta S^0 = -RT\ln K_{eq} \\ \ln K_{eq} = -\frac{\Delta H^0 - T\Delta S^0}{RT} \\ \ln K_{eq} = -\frac{\Delta H^0}{RT} + \frac{\Delta S^0}{R} \end{aligned} \tag{4.7.12.8}$$

From these equations, it should be evident that all the major thermodynamic constants (ΔG^0, ΔH^0, and ΔS^0) for the N ↔ D transition can be calculated from thermal denaturation curves.

Equation (8) below shows that the derivative of equation (7) with respect to 1/T (i.e. the slope of equation 7 plotted as $\ln K_{eq}$ vs 1/T) is indeed $-\Delta H^0/R$. Equation (8) is the van't Hoff equation, and the calculated enthalpy change is termed the van't Hoff enthalpy, $\Delta H^0 v_{Hoff}$.

$$\frac{d \ln K_{eq}}{d(1/T)} = -\frac{\Delta H^0}{R} = -\frac{\Delta H^0_{vHoff}}{R} \tag{4.7.12.9}$$

It is helpful to compare the van 't Hoff enthalpy, $\Delta H^0 v_{Hoff}$, with the enthalpy change determined directly using differential scanning calorimetry by analyzing a plot of Cp vs T. (Note that the area under the Cp vs T curve, as the protein transitions to the unfolded state, has units of kcal or kJ. The ΔH^0 for the unfolding is inversely proportional to the width of the curve.)

In contrast to the long extrapolation of the ΔG^0 vs [urea] to [urea] = 0 to get ΔG^0 (the y-intercept) in the absence of urea, which has some physical meaning, extrapolation of the straight line from the van 't Hoff plot from equation 8 to get $\Delta S^0/R$, the y-intercept, has little meaning since the 1/T value at the y-intercept is 0, which occurs when T approaches infinity. ΔS^0 can be calculated at any reasonable temperature from the calculated value of ΔG^0 at that temperature and the calculated ΔH^0_{vHoff}.

4.7.12.4: Summary

This chapter delves into how we quantify and understand protein stability through thermodynamic principles and experimental measurements. We begin by recognizing that protein stability is a delicate balance between favorable interactions (such as hydrophobic effects, hydrogen bonds, and van der Waals forces) that stabilize the native state and the inherent conformational entropy of the polypeptide chain that favors the denatured state.

To investigate these factors, various experimental methods—including UV absorbance, fluorescence, circular dichroism (CD), and viscosity—are used to monitor protein unfolding. By analyzing denaturation curves (plotted as a function of temperature or denaturant concentration), we can determine the fraction of protein in the native versus the denatured state. The relationship

$y=f_N y_N+f_D y_D$ and $f_D=y-y_N y_D-y_N$

allows us to quantify the fraction denatured (fDf_D) based on an observable yy (e.g., absorbance or fluorescence intensity). From there, the equilibrium constant for unfolding, $K_{eq}=f_D/(1-f_D)$, leads directly to calculating the standard free energy change for unfolding ($\Delta G^0=-RT\ln K_{eq}$.

In thermal denaturation, the melting temperature (Tm) is defined as the temperature at which half of the protein population is unfolded (i.e., K_{eq} = 1 and ΔG^0 = 0). By constructing van ’t Hoff plots ($\ln K_{eq}$ vs. $1/T_1$), we can extract the enthalpy (ΔH^0) and entropy (ΔS^0) changes associated with protein unfolding. These thermodynamic parameters are crucial for understanding how proteins respond to environmental changes and how mutations might alter protein stability.

Finally, the chapter emphasizes the importance of rigorous data analysis and careful extrapolation of denaturation curves to obtain accurate estimates of protein stability under physiological conditions. Such quantitative assessments provide a foundation for comparing the stability of wild-type proteins with mutant forms and for understanding the molecular forces that drive protein folding and function.

This integrated thermodynamic framework equips you with the conceptual and analytical tools necessary to interpret experimental data and to appreciate the fine balance that maintains protein structure and function.

This page titled 4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools

Search Fundamentals of Biochemistry

Recent Updates: New Chapter Section 11/1/23

Written by Logan Hallee and Henry Jakubowski

Learning Goals (ChatGPT o1, 1/27/25

Learning Goals: Protein Language Models and Their Applications in Biochemistry

1. **Understand the Motivation and Context for Protein Language Models (pLMs):**
   - Explain why traditional protein structure determination methods (X-ray, NMR, cryo-EM) are limited by cost and time.
   - Describe how massive sequence databases and computational models are used to predict protein structure and function from sequence alone.
   - Discuss the importance of integrating computational predictions with experimental biochemical assays.
2. **Comprehend the Fundamentals of Transformer-Based Neural Networks:**
   - Define what a transformer is and explain its role as an attention-based neural network in natural language processing.
   - Differentiate between various transformer architectures (e.g., encoder, decoder) and their applications (BERT vs. GPT).
   - Summarize the concept of "Attention Is All You Need" and how attention mechanisms enable efficient handling of sequential data.
3. **Learn the Core Components of Protein Language Models:**
   - **Tokenization and Token Embedding:**
     - Explain the process of converting protein sequences into tokens (using amino acids as tokens) and mapping them to numerical representations.
     - Understand the role of token embedding matrices and how they store semantic information.
   - **Attention and Self-Attention:**
     - Describe the attention mechanism, including the roles of query, key, and value matrices.
     - Explain how self-attention enables each amino acid token to dynamically interact with every other token in the sequence.
     - Understand the mathematical formulation of attention, including the use of softmax and normalization by the key dimension.
   - **Multi-Head Attention and Transformer Stacks:**
     - Explain how multi-head attention allows the model to learn different aspects of sequence relationships in parallel.
     - Describe how stacking transformer layers enhances the model's capacity to capture complex sequence semantics.
4. **Explore Training Strategies for Protein Language Models:**
   - Describe how large corpora of protein sequences are used to train pLMs.
   - Understand training objectives such as masked token prediction (denoising) and next token prediction.
   - Recognize the role of special tokens (e.g., MASK, CLS, SEP) in guiding the model during training and task-specific applications.
5. **Examine Applications of Protein Language Models in Biochemistry:**
   - **3D Structure Prediction:**
     - Explain how models like AlphaFold integrate multiple sequence alignments (MSA) with transformer-based latent spaces and structure modules to predict protein 3D coordinates.
     - Discuss the advantages and limitations of MSA-based methods and the developments in pLMs for "protein orphan" sequences.

- **Protein Function Prediction:**
  - Understand how protein embeddings derived from pLMs can be used to predict enzyme activity (EC numbers), Gene Ontology (GO) annotations, and other functional properties.
- **Protein Sequence Generation and Design:**
  - Describe how generative models (e.g., ProtGPT2, ProGen, ANKH) can design new protein sequences with desired structural or functional properties.
  - Discuss the potential impact of protein sequence generation in therapeutics, synthetic biology, and industrial catalysis.

6. **Integrate Interdisciplinary Perspectives:**

- Appreciate how protein language modeling bridges biochemistry, bioinformatics, and machine learning.
- Discuss how advances in natural language processing (NLP) are transforming protein modeling and accelerating biological discovery.
- Recognize the evolving nature of computational methods in biochemistry and the importance of staying updated with emerging technologies.

These learning goals aim to equip you with a comprehensive understanding of how modern protein language models work, their underlying computational principles, and their transformative applications in predicting protein structure and function.

4.7.13.1: Introduction

So far in this chapter, you have learned about protein structure and its determination in the laboratory. After decades of work in modeling protein structure and properties, the life science community has built massive databases organizing this information. While sequencing DNA and discovering protein sequences has become relatively cheap, the actual characterization of protein structure and function is still time and cost-intensive. Instead, researchers look to model and predict protein properties from their amino acid sequence alone to speed up the work necessary in the lab. The most recent and effective tool in this quest is the **protein language model (pLM)**, which models proteins as a biological language of amino acids. At their core, pLMs are **transformers** with a protein vocabulary.

The transformer, an attention-based neural network, emerged as a game-changer for the scientific community with the iconic 2017 paper “Attention Is All You Need.” The crux of this work is the revolutionary idea that by strategically organizing simple neural networks, performance can be enhanced beyond merely scaling up a singular neural network. A neural network is a type of AI/machine learning process, often described as deep learning, which is patterned after the brain with nodes (neurons) interconnected by lines (axon/dendritic connections) as in the brain.

Transformers have become the bedrock of modern **natural language processing (NLP)** and are especially adept at processing sequential data like time series or sentences. This technology is grounded in understanding their trained vocabulary's semantic and contextual intricacies.

4.7.13.2: The Essence of Protein Language Models

Tokenization & Token Embedding

The fundamental problem in NLP is encoding text, a string datatype, into a meaningful numerical representation; it is extremely challenging to perform mathematical operations on words. One approach to the problem is to give every sub-part of the vocabulary, say a word, a unique integer. That way, any sequence of words can be turned into a vector of integers, and we can easily do the math on vectors; think back to physics or math doing the dot product on collections of numbers that have *direction*. This process creates a look-up table called a **tokenizer** because it translates tokens (strings of words, letters, or characters) to integers.

? Vectors - A Simplified Review

To understand how AI/Machine Learning can be used to predict structure and function, we need to know about vectors and their use in physics and mathematics, particularly in matrices. Most students likely need a refresher. Click the link below for a guided view that will allow you better to understand the rest of the material in this section.

The Review!

Vectors – A Simplified Review

Most biochemistry students have taken physics in high school and college. In those courses, you were introduced to scalar and vector quantities. Scalar quantities like distance and work have no direction. Vector quantities like displacement and force have both a magnitude and a direction. Vectors are shown as arrows, with the length representing the magnitude and the direction by an arrow at the end of the vector.

Let's review a simple concept from elementary physics work. Work is a scalar quantity, and you probably remember that mechanical work is done on an object when an external force moves an object a given distance. Consider a force **F** (bold represents a vector) applied to a block, which causes it to move a distance along the surface. The distance and direction together are described as **displacement d**, as shown in Figure 4.7.13.1 below.

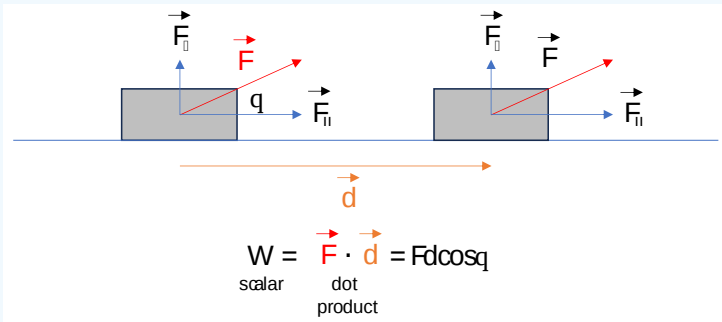


Figure 4.7.13.1: Forces on a block moving along a surface

The block would not move along the surface if the force were applied vertically, so no work is done on the box. If the force is applied at some angle, only the horizontal component of the force would cause the block to slide.

The horizontal component of the force is described as **F**cosθ. (When $\theta = 90^0$, $\cos\theta = 0$, so no work is done.) Hence, the work W = Fcosqd, the "dot product" of the vectors **F** and **d**.

Vectors are also used in math and can be considered directed line segments. Let's consider the equations for circles and spheres. These equations are based on the Pythagorean Theorem. As you learned in high school geometry, the Cartesian equation of a circle is:

$x^2 + y^2 = r^2$

To generate a circle, set the r-value to a fixed number (such as 1 for a "unit circle"), and for a multiple number of x values, calculate y values (where $-1 \leq x,y \leq +1$) from the equation. Then plot the x and y coordinate pairs, and presto, you have a circle, as shown in Figure 4.7.13.2 below.

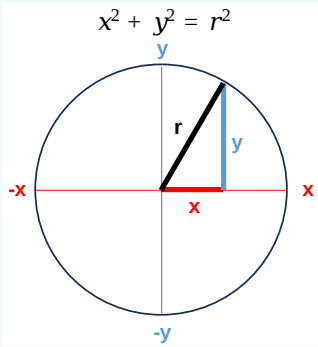


Figure 4.7.13.2: Cartesian graph of a circle

Each of the (x,y) pairs can be considered a 2D vector with the origin at 0,0, a magnitude of 1 (the fixed radius in this example), and a direction described by the specific x,y points that fall on the circle. Two simple (x,y) pairs are (1,0) and (0,1) for the unit circle. Another x,y pair that satisfies the Pythagorean theorem is (0.5, 0,866)

The Pythagorean Theorem can be extended to three dimensions to give the Cartesian equation for a sphere:

$x^2 + y^2 + z^2 = r^2$

To generate a sphere, solve for z for many x and y values and a fixed r value. Then plot the x, y, and z values, and presto, you get a sphere, as shown in Figure **4.7.13.3** below for a "unit" sphere of radius 1.

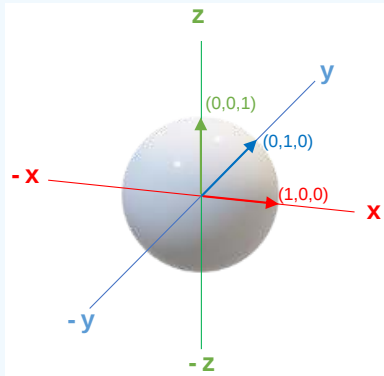


Figure **4.7.13.3**: Cartesian graph of a sphere

The sets of x, y, and z points that land on the surface are vectors (directed line segments).

Now vectors are also used to describe matrices. Matrices are two-dimensional arrays of numbers. A matrix with just one row or one column is called a row vector or column vector, respectively, as shown in Figure **4.7.13.4** below.

$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$	Row vector	$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	Column vector
$\begin{bmatrix} 1 \end{bmatrix}$	1 row, 1 column, scalar	$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$	Square matrix = # columns/rows

Figure **4.7.13.4**:

You should now see that all vectors defining a sphere can be written as a large matrix. Figure **4.7.13.5** below shows a x3 square matrix representing the unit vectors along the x, y, and z axes.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Figure **4.7.13.5**:

In this example, any vector will lie on the unit sphere's surface if the three components have the relationship of the Cartesian equation above.

These examples showcase how values in vectors can represent a position in space, but they can also be measurements of an object. For example, a car with four tires, four cylinders, and 180 horsepower could be represented as (4, 4, 180), where a more detailed description will lead to a longer vector with more components.

Another use of vectors and matrices comes from more complicated mathematical graphs. Graphs are arbitrary objects with nodes and edges, where edges connect nodes. A popular example of a mathematical graph is the social network Facebook. You can represent the entirety of the Facebook network with each member as a node and an edge between each node when the nodes are friends on the site. Figure **4.7.13.6** below shows an example of a social networking graph.

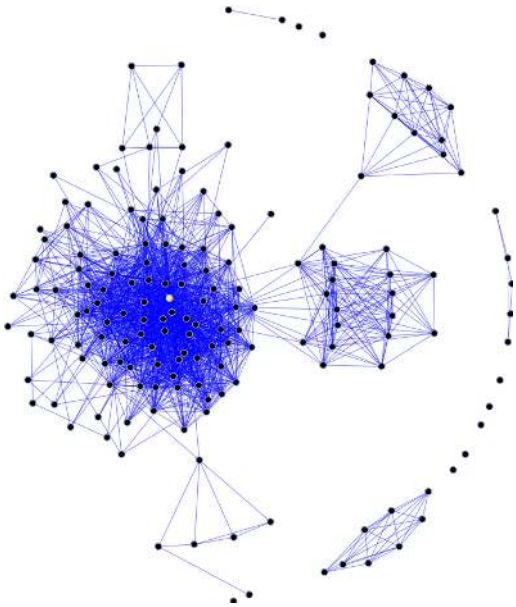


Figure 4.7.13.6: Graph showing social relationships using graph theory. Darwin Peacock. CC BY 3.0, https://commons.wikimedia.org/w/inde...?curid=6057981

This example is an unweighted, undirected graph because the edges do not have a specific value or direction. We could make a weighted graph representing Facebook using weighted edges, perhaps the number of mutual friends between members. However, this is still an undirected example.

Graphs are important in computational biochemistry because molecules can be represented as graphs, with atoms as nodes and bonds as edges. This is illustrated in Figure 4.7.13.7 for a multidentate adsorbate complex.

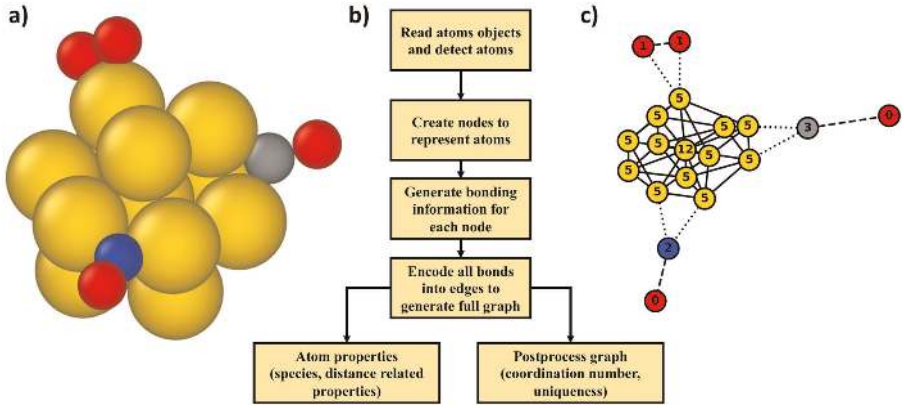


Figure 4.7.13.7: Graph theory-based algorithm to generate graphs for a given atomic model. Deshpande, S., Maxson, T. & Greeley, J. Graph theory approach to determine configurations of multidentate and high coverage adsorbates for heterogeneous catalysis. npj Comput Mater 6, 79 (2020). https://doi.org/10.1038/s41524-020-0345-2. http://creativecommons.org/licenses/by/4.0/. Creative Commons Attribution 4.0 International License.

Panel a shows an atomic model for a simple nanoparticle with adsorbates. Panel b is an algorithm to generate graph-based representations. Panel C shows the generated graph mode.

Figure 4.7.13.8 below shows a protein structure graph (right) for a short stretch of an alpha helix (left).

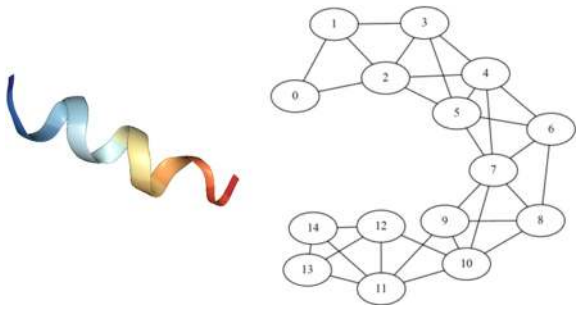


Figure 4.7.13.8: Protein and protein graph. Baranwal, M., Magner, A., Saldinger, J. et al. Struct2Graph: a graph attention network for structure-based predictions of protein–protein interactions. BMC Bioinformatics 23, 370 (2022). https://doi.org/10.1186/s12859-022-04910-9. http://creativecommons.org/licenses/by/4.0/

Figure 4.7.13.9 shows a graph not of a single protein but of a small protein:protein interaction network.

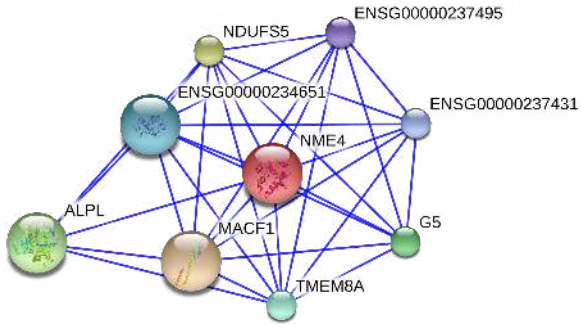


Figure 4.7.13.9: protein interactions of TMEM8A in humans. https://commons.wikimedia.org/wiki/F...for_TMEM8A.png

You can even produce graphs that represent entire networks of molecules and their relationships. A directed molecular graph might showcase proteins and their substrate. Having a direction in an edge is important in this distinction because a protein may use a substrate for a chemical reaction, but a substrate might not act on a protein on its own. Many molecular relationships are weighted and directed. A weight in the protein substrate case might be the relative affinity of binding between the molecules.

These graphs contain three types of edges: undirected, directed, and weighted, as illustrated in Figure 4.7.13.10 below.

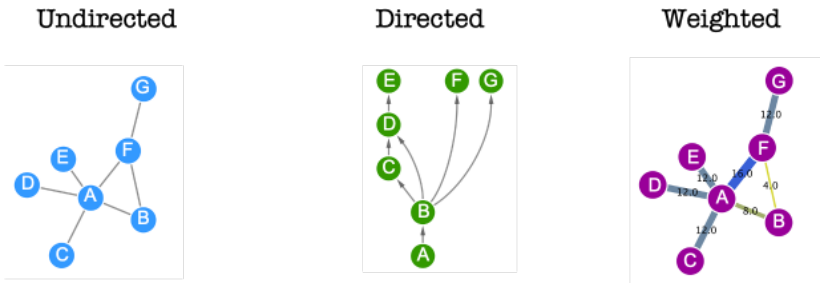


Figure 4.7.13.10: The main types of edges found in a network. https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/introduction-to-graph-theory/graph-theory-graph-types-and-edge-properties/ . Attribution 4.0 International (CC BY 4.0) license

- Undirected edges: Connections in protein-protein interactions, as shown in Figure 9 above, are examples. The proteins are connected through binding but without implied flow between them;
- Directed edges: These are found in metabolic and signaling pathways when arrows indicate the flow of reactants/products in a pathway. These can be arranged in complex hierarchies, as those familiar with metabolic and signaling pathways know.

- Weighted edges: Either undirected or directed edges can have a quantitative weight associated with them. These many reflect affinities, similarities between genes, fold effects, etc.

These examples are great for showcasing the versatility of mathematical graphs, but how can we use matrices to represent them? Enter adjacency matrices.

Adjacency matrices state which notes are connected. Each node can have multiple features. Figure 4.7.13.11 below shows a network of 5 nodes, the adjacency matrix, and a features matrix with each node having features. For example, if an atom is a node, the features could be electronegativity, partial charge, size, etc.

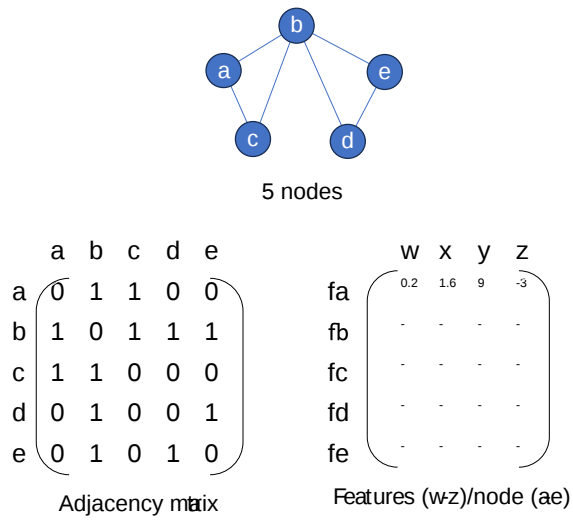


Figure 4.7.13.11: Properties of a 5 node network

An adjacency matrix stacks n vectors together for a graph with n nodes. The vectors are also n long, so the resultant matrix is n by n. At the ith jth index of the matrix is a number dictating how many edges the node shares. So if the 1st node has an edge to the 2nd node, the 1st row and 2nd column of the adjacency matrix will have a 1. These tend to be symmetric; in this example, there would also be a 1 at the 2nd row and 1st column. Figure 4.7.13.12 below does a great job at explaining:

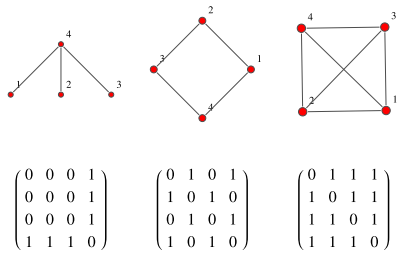


Figure 4.7.13.12: https://mathworld.wolfram.com/AdjacencyMatrix.html

If there is a weight on an edge, the number in the adjacency matrix can be used to store the weight instead of the count of edges between nodes. This can also hold a direction by allowing for positive and negative entries.

Any complex network can be described mathematically as an adjacency matrix with rows and columns indicating nodes and an edge as a number. Unweighted and undirected edges lead to symmetric matrices with just 0 and 1. Directed and weighted can be more complicated with different numbers used to show relationships, such as affinity. +/- values can be used where + is an activation and – is an inhibition. These matrices can be manipulated using linear algebra. Examples of adjacency matrices for Undirected, Directed, and Weighted networks are shown in Figure 4.7.13.13 below.

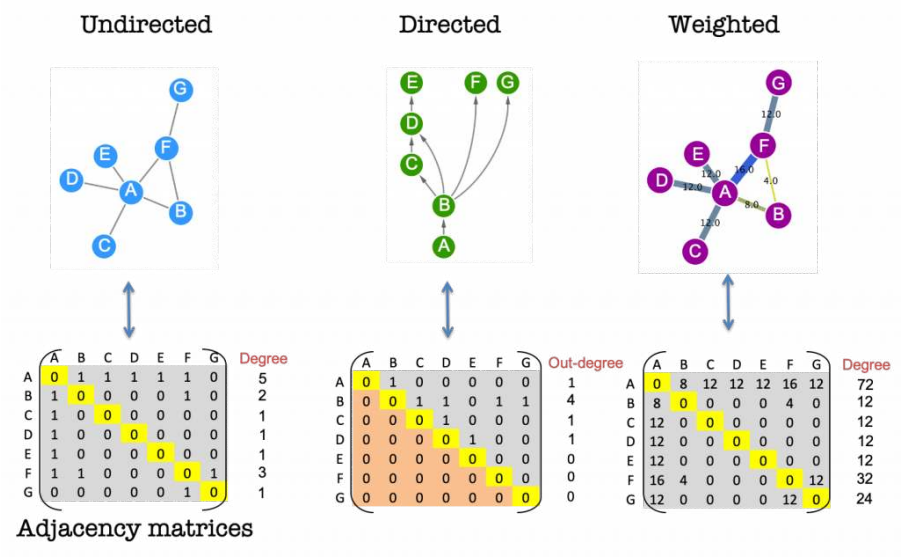


Figure **4.7.13.13**: Adjacency matrices from undirected, directed, and weighted networks. https://www.ebi.ac.uk/training/online/courses/network-analysis-of-protein-interaction-data-an-introduction/introduction-to-graph-theory/graph-theory-adjacency-matrices/ . Attribution 4.0 International (CC BY 4.0) license

Hopefully, after this brief introduction to vectors and matrices, you can understand how amino acid positions and their semantic properties (polarity, charge, size, etc.) can be described as large matrices and their associated vectors.

4.7.13.3: Back to Tokens

Here are a few types of tokenizers commonly used in NLP:

1. Word Tokenizers: Word tokenizers split text into individual words based on spaces or punctuation marks. This approach assumes that words are the primary units of meaning in a language. For example, given the sentence "The cat is sleeping," a word tokenizer would split it into tokens: ["The", "cat", "is", "sleeping"].
2. Subword Tokenizers: Subword tokenizers split text into subword units that capture partial linguistic information. This approach is useful for handling out-of-vocabulary words like abbreviations or reducing vocabulary size, which saves computational resources. Popular subword tokenization algorithms include Byte Pair Encoding (BPE), Unigram Language Model, and SentencePiece. Subword tokenizers can be complicated, but one possible subword tokenization for our example above would be [“The#”, “cat#”, “is#”, “sleep”, “ing#”] where # has been added to showcase the ending of a word.
3. Character Tokenizers: Character tokenizers treat each character as a separate token. This approach is beneficial when dealing with languages without explicit word boundaries or for character-level modeling tasks. Our sleepy cat is now [“T”, “h”, “e”, “\s”, “c”, “a”, “t”, “\s”] and so on. Here, we need to add a space token so the model can tell where the words start and finish.

For pLMs, researchers typically treat amino acids as tokens or “words” and protein sequences as “sentences.” For example, a protein sequence like "MVKLTA" would be tokenized into individual amino acids: “M,” “V,” “K,” “L,” “T,” and “A.”

The main problem with tokenizing sentences or sequences is that this numerical space has no *semantic* meaning. The grammar and word meanings are not encoded here. We will define another look-up table to work on this semantic information storage. However, this time, we will make it a large matrix with some arbitrary dimension, say, 768. We assume we want to model approximately 50,000 English words, so our matrix is 50,000 x 768. Now, we will simply connect our tokenizer and our matrix. If the word “dog” corresponds to the token 5, the *vector* representing the dog will be the 5th row of our matrix. Maybe “protein” is token 800. The *vector* that represents the protein is the 800th row in our matrix.

These vectors representing each word are where we will store the semantic information behind each word, the meaning, and related grammar within the vocabulary. Unlike an object like a car, with easy numerical features to pick out (4 doors, 4 wheels, 6 cylinders, etc.), we need to learn the features of a word. And so, within the transformer neural network, we have a vector representing each word containing learnable weights. Through the process of gradient descent, these weights will be adjusted from a random starting point to contain the necessary information for language modeling. One neat fact is that within a well-trained English language

model, the vectors for words King - Man + Woman roughly equal Queen. How amazing that concepts behind gender and royalty can be encoded in a meaningful numerical space!

Of course, many words have different meanings in different contexts. This is still the case in proteins, where specific residues may be important because of their charge or maybe just because of the space they occupy. That is why we need a portion of our transformer that handles the contextual understanding. This is where **attention** comes in.

4.7.13.4: Attention

Attention is used to signify the importance of a word or its part. For example, if high attention is given to a word, less overall information is needed to predict the word. Attention can be conferred by weighting the importance of the word/token. In real language, we attend to certain words in sentences and weigh them more to provide context and make predictions. Consider this sentence referring to a TV show or movie. Predict which show/movie it is from.

..and Spock said that the needs of the many outweigh the needs of the few.

The answer comes quickly (if you are a Star Trek fan) by attending to the word Spock. You could maybe predict the entire sentence using databases by perhaps attending to another word, such as outweigh.

Using attention enables the transformers we discussed above to access long-term memory and concentrate (attend) to previously generated tokens.

The attention mechanism identifies dependencies and relationships between tokens within an input sequence. What makes attention especially compelling is its ability to dive deep into a relational space and seamlessly bridge back into the token embedding space we previously discussed.

Another feature is self-attention, which focuses on the relationships within a given sentence.

Self-attention, where every token in a sequence attends to every other token, allows for a dynamic weighting of significance.

Three key inputs are required. The Query is what you are asking (such as the input in a web search box), the Key is the search results, and the Values are the returned content in each search result. In a web search, the program has to find the best matches between the query and the keys.

Now, consider both the query and key to be vectors. The similarity between them can be determined using a cosine similarity function, which is a bit similar to the dot product of the vectors = A·K/|A||K|= Akcosθ/|A||K|. The denominator is the product of the vector lengths. The cosine function makes sense since it varies between +1 (when the vectors are in the same direction, q = 1) and -1 (when they point in opposite directions, q = -1). This gives the degree of similarity.

Mathematically, the relationships among query, key, and values matrices are formulated as

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4.7.13.1}$$

Here, the Q, K, and V stand for query, key, and value matrices, respectively, and d_k is the dimension of the queries and keys. These are obtained directly by multiplying learned weights with our token embeddings.

$$Q = W_e \times W_Q, \quad K = W_e \times W_K, \quad V = W_e \times W_V, \tag{4.7.13.2}$$

The softmax function turns a vector of K real values into a vector of K real values that sum to 1, effectively converting them into probabilities (that sum to 1).

We symbolize the token embedding matrix, or, at least, the portion extracted from our tokenizer and token embedding process. d_K denotes the dimension of the key matrix, which is used to normalize the values.

During the attention operation, there is a length-wise representation and an embedding-wise representation. The length-wise representation is typically called an “attention matrix” and is often analyzed by researchers to determine how a model makes decisions. Coming from the Softmax(QK^T/d_K) during the attention operation is this attention matrix: $A_{LxL,}$ where L symbolizes the sequence length, and values are scaled from 0 to 1. When $A_{\{i,j\}}$ is close to 1, it signifies that the ith amino acid bears contextual

relevance to the jth one. In pLMs, such associations often resonate with spatial or chemical bonds—perhaps the ith amino acid is nestled close to the jth one in the protein's 3D structure. Notably, $A_{\{i,j\}}$ scores are perpetually high, emphasizing an amino acid's inherent significance to itself. When A is multiplied further by V, we go from this relational space back to the embedding space, where the matrix is L x embedding dimension. In our previous example, this was 768, but this number is completely arbitrary. A smaller embedding dimension essentially leads to a smaller model with a worse theoretical top-end performance, with the advantage that it is much cheaper to use and train. A larger embedding dimension and model will have a theoretically high top-end performance, require more data and computational resources to train, and be more expensive.

Diving deeper, multi-head attention further refines this process. It transforms the input sequence into multiple smaller queries, keys, and values, each an independent attention head brandishing unique weights. Each head, in isolation, computes its own version of the attention matrix on a section of the input. These matrices are then woven together, under an additional learned linear transformation, into a final cohesive representation. This multi-headed approach has proven important for pLMs, where an increased headcount is correlated with better performance.

Bringing it all together ...

Lastly, bridging back from the realm of relational depths to the embeddings, transformers don't solely rely on attention. They also encompass feed-forward layers, which further improve the generalized modeling capabilities. Together, these components are the essence of the transformer: a multi-headed attention layer and a linear layer. These transformer layers are simply stacked on top of each other, so the input of one is the output of another. These transformer stacks are referred to as large language models (LLMs). LLMs have a couple of main forms and some other important parts that we did not discuss. For example, there are transformer encoders (BERT) and decoders (GPT) that have different organizations of layers. There are also some additional embeddings called positional embeddings, which help transformers learn the intended order of input sequences because the native mathematical operations are position invariant. Also, there are different types of normalizations and skip connections that are also extremely important for protein modeling. So while this introduction to transformers is enough to get your feet wet, we have included several other resources to refine your knowledge further if desired.

4.7.13.5: Training transformers

At a high level, transformers are trained just like any other neural network. For a large corpus of training data, there is an input and a known desired output called the ground truth. The input is fed to the model, and the output is compared against the ground truth with a loss function. This is all set up so that the loss function measures some error between the output and the ground truth, so that the loss function and corresponding error are minimized with optimization techniques. Regardless of the technical details, it is essentially a strategic trial and error, where the model is rewarded for producing outputs close to or equal to the ground truth. This way, through optimization, the error is reduced over time, and the model learns how to do the specified task effectively.

However, because transformers have such a large parameter count, they require a large amount of data to adjust for a specific task successfully. Thus, we need a way to produce lots of labeled data without wasting too much human time on annotating data. Luckily, sequences of strings are perfect for this. If we feed a transformer a sequence, we can simply hide a token and ask the model to predict what word went there.

Now, in the input, let’s replace some words (or tokens) with **Mask** tokens randomly. Mask tokens are used to hide tokens from the model so that we can train the model to recover the missing words, where the desired output is the original sequence. Such a task forces the model to learn what words are around it, hence building the model's semantic and contextual understanding. This is called denoising because we artificially injected noise into our input, and our model used the surrounding context to eliminate the noise. Another popular training objective is next token prediction, where part of a sequence is input, and the model predicts what goes next. Different organizations of transformer layers are better or worse at these various tasks.

The subject of a MASK token brings up the broader topic of special tokens. Special tokens are tokens that are added to a specific vocabulary that serve a specific purpose. As we discussed, a model can learn to replace a MASK token with a correct token that belongs in the sequence.

Some other popular tokens are **CLS** and **SEP**, which stand for **classification** and **separator,** respectively.

CLS tokens are typically prepended to the beginning of sequences so that models can learn to summarize the entire input in a single vector, which is useful for classification tasks. Separator tokens are typically placed at the end of sequences or between sequences

if more than one was fed to the model at a time. This way, the model can treat separate sequences as individual entities even if they are input simultaneously. Researchers will often create new special tokens to use alongside model training for specific tasks. For example, a pLM called ProstT5 has a special token that indicates a translation from structure to amino acid sequence and an additional token that does the reverse. The respective token was prepended to the necessary inputs during training so that the model could intuit which task it was supposed to be performing. This is particularly useful if you are designing a transformer model that needs to do multiple distinct tasks.

4.7.13.6: Applications of protein language models

Now that we have learned substantially about how modern language modeling works, we can explore how researchers utilize these techniques to advance our understanding of biochemistry further.

Prediction 3D Structure

By and large, the most famous pLM is AlphaFold: The deep learning model that first successfully mapped amino acid sequences to protein structure at a large scale. Since then, other pLMs have also been able to learn from a large corpus of sequences and structure data to perform well on unseen sequences. The fundamental organization of sequence-to-structure models is made up of

- a transformer (pLM) that builds a semantic and contextual understanding of sequences;
- a structure module that maps the latent sequence representation to 3D coordinates.

AlphaFold is special in a few ways. Firstly, the transformer used for AlphaFold is two transformers in one! Let us call the first transformer T1. T1 is an MSA transformer that works on several sequences at once instead of a single sequence. This comes from the concept of multiple sequence alignment (MSA), a common bioinformatics search method that compares amino acid strings based on an evolutionary-informed algorithm. The output of an MSA is a list of strings aligned based on similarity and substitution probability. Figure 4.7.13.13 below …

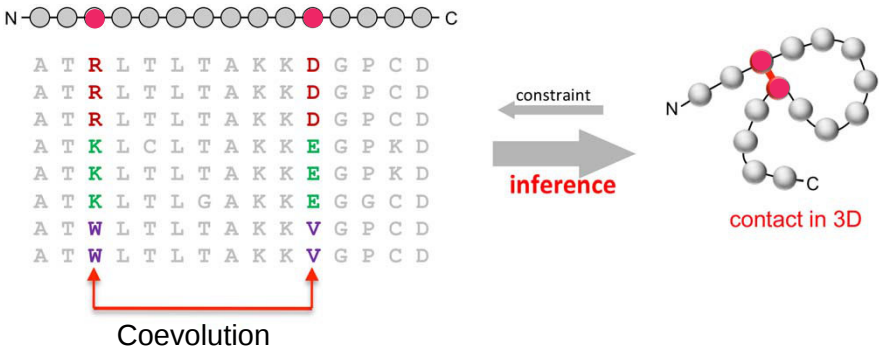


Figure 4.7.13.13: Correlated mutations carry information about distance relationships in protein structure.https://www.blopig.com/blog/2021/07/alphafold-2-is-here-whats-behind-the-structure-prediction-miracle/

The sequence of the protein for which the 3D structure is to be predicted (each circle is an amino acid residue, typical sequence length is 50–250 residues) is part of an evolutionarily related family of sequences (amino acid residue types in standard one-letter code) that are presumed to have essentially the same fold (iso-structural family). Evolutionary variation in the sequences is constrained by many requirements, including the maintenance of favorable interactions and indirect residue-residue contacts (red line, right). The inverse problem of protein folding prediction from sequence addressed here exploits pair correlations in the multiple sequence alignment (left) to deduce which residue pairs are likely to be close to each other in the three-dimensional structure (right). A subset of the predicted residue contact pairs is subsequently used to fold up any protein in the family into an approximate predicted 3D shape ('fold'), which is then refined using standard molecular physics techniques, yielding a predicted all-atom 3D structure of the protein of interest.

Including this information in the model allows the model to learn more about the protein than it could from a single sequence because of the concept of coevolution. Coevolution is a simple process that necessitates meaningful substitutions in amino acid sequences through various mutations if an organism is going to stay fit.

However, as we have discussed, transformers are incredibly adept at processing and understanding single sequences; how would one process many sequences simultaneously? The MSA input to the pLM is many sequences stacked on top of each other, so the attention mechanism needs to be modified. The MSA goes through row-wise attention, which picks out the important residues, and column-wise attention, which picks out the important sequences. This creates a protein latent space built from an MSA instead of a single sequence.

The other transformer, T2, also has modified attention (triangular self-attention). T2 is for computing over a pairwise representation of the single input sequence. It builds a representation similar to a distogram, a matrix representation of the distance between every residue and every other residue. T1 and T2 make up the section of AlphaFold known as the EvoFormer, and there are 48 of these EvoFormer layers in total. The latent representation of the MSA input from T1 and the pair-wise representation from T2 are both input structure modules.

The structure module also has some fancy attention (invariant point attention) that merges the information from T1 and T2. A simple computer-vision-inspired architecture called a ResNet used the attention output to predict side chain and backbone torsion angles at the atomic level. From these outputs, the atom coordinates for the entire protein are calculated, and the structure is relaxed with Amber, which removes any structural violations based on the charges and locations of the atoms.

This entire process is repeated three times, where the structure information and MSA information can inform each other through skip connections and linear transformations. This recycling greatly improves the final 3D structure. Figure **4.7.13.14** below.

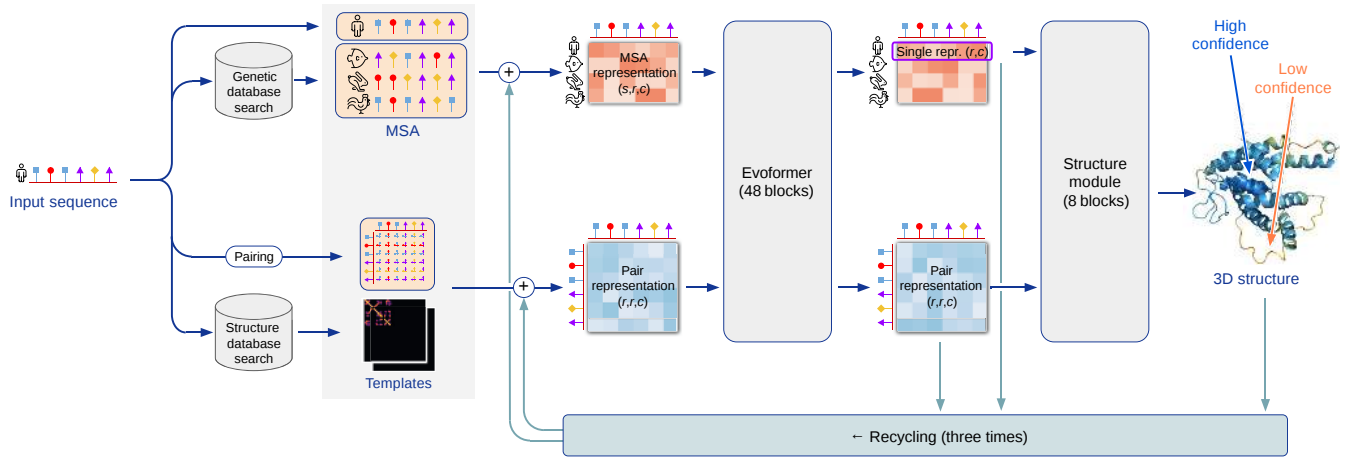


Figure **4.7.13.14**: Model architecture. Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Arrows show the information flow among the various components. Array shapes are shown in parentheses, with s indicating the number of sequences, r indicating the number of residues, and c indicating the number of channels.

In summary, AlphaFold combines three main networks that talk to each other and work together. The input sequence is searched with MSA through a large database that pulls out similar sequences. These similar sequences are input to T1 which builds a semantic and contextual understanding of the amino acids called the latent space. This latent space informs a pair-wise distogram that tracks the distances between amino acid residues in the original input sequence. All of this information is utilized by the structure module, which does some fancy math to calculate the 3D coordinates of each atom. The entire process is repeated so that the structural information can inform the MSA, and the MSA can further inform the structure. Throughout the process, a ResNet and Amber are used to prevent any weird side chains or backbone angles that cannot exist in nature.

Other projects have also taken the MSA approach to structure prediction. RoseTTAFold and xtrimoGPLM utilize an MSA with different networks and attention variants, achieving similar accuracy to AlphaFold. However, MSA can be quite computationally expensive and requires a database with similar sequences. These methodologies fall short for so-called protein orphans, who have little to no sequence homology in recorded repositories.

To address the computational and protein orphan concern, various projects like ESM, OmegaFold, and Ember have utilized a standard pLM trained through mask denoising and a structure module, also to obtain high accuracy on structure prediction, outperforming MSA-based methods on protein orphans. Breakthroughs from AlphaFold, like recycling, are standard practice throughout these different approaches.

Some researchers utilize these various projects in parallel to get multiple structure predictions for the same sequence. This ensemble approach allows accelerated progress by leveraging the advantages of each model and averaging out the disadvantages. Low confidence regions or disagreements in models may also indicate intrinsically disordered regions of protein structure, which are incredibly important in biological function.

Overall, sequence-to-structure mapping has been effectively correlated with modern computational methods. The backbone of pLMs and structure modules has enabled the large-scale annotation of protein sequences with high-quality structures, which the scientific community can utilize for accelerated breakthroughs. However, structure prediction is not the only thing you can do with pLMs.

Protein function prediction

The latent space learned from vast mask denoising on corpora of amino acid sequences correlates extremely highly with protein structure, which is vital for the study and annotation of proteins (link to other parts of the textbook). Interestingly, this latent space also correlates highly with other useful annotation types like function. By averaging across the length of the last hidden state output of pLMs, one can build an effective vector representation of a protein, called a vector embedding. This way, every protein has the same size numerical representation and can easily be fed to machine learning classifiers like support vector machines, k-nearest neighbors, random forests, and more. The pLM can also be fine-tuned as a classifier, given enough annotated data. Precomputed protein embeddings from popular models like ProtT5 can be downloaded from UniProt, as well as large amounts of annotated data if you want to try this for yourself.

Researchers are typically interested in annotations involving details like EC and GO classes. EC stands for Enzyme Commission, and EC numbers break down protein functionalities into a hierarchical organization scheme delimited by the type of reaction the proteins catalyze.

For example, the hierarchy of an EC number can be illustrated as:

- 1st digit: Represents one of the six primary classes of enzymes, e.g., '1' stands for oxidoreductases.
- 2nd digit: Describes a subclass within the primary class. If an enzyme is a '1.1', it specifically acts on the CH-OH group of donors.
- 3rd digit: Categorizes the enzyme even further by specifying the acceptor. For instance, '1.1.1' would mean that the enzyme acts on the CH-OH group with NAD+ or NADP+ as the acceptor.
- 4th digit: Provides a unique identifier for each enzyme within its specific class, subclass, and sub-subclass. So, '1.1.1.1' is the EC number for alcohol dehydrogenase.

In total, there are currently over 8000 unique EC numbers! pLMs have shown remarkable competency in predicting them from sequence alone, often achieving between 80-90+% accuracy on unseen data.

Another type of popular annotation is Gene Ontology (GO), which labels genes based on what their protein products do in a biological context. GO is split into three main categories

1. Biological Process (BP): Describes a series of events accomplished by one or more ordered assemblies of molecular functions. For instance:

   GO:0006955 - Immune response

   GO:0006958 - Complement activation, classical pathway

   GO:0045087 - Innate immune response

   … and so on

2. Cellular Component (CC): Describes parts of a cell or environment a protein product likely localizes to. For example:

   GO:0005634 - Nucleus

GO:0005654 - Nucleoplasm

GO:0005694 - Chromosome

… and others

3. Molecular Function (MF): Describes catalytic activities, such as binding or catalysis, that occur at the molecular level. This subcategory is very similar to EC numbers. For instance:

   GO:0003824 - Catalytic activity

   GO:0016491 - Oxidoreductase activity

   GO:0016614 - Oxidoreductase activity, acting on CH-OH group of donors

   … and further subcategories

The main difference in organization is that GO terms have parent-child relationships that are not unique. In simpler terms, a GO term can have multiple parent terms, while an EC number may only have one. Regardless, pLMs also show a wide breadth of impressive performances in predicting GO terms from sequence alone.

Additionally, researchers are interested in the complex interplay of proteins in the cell. How do proteins modify each other and their surrounding cellular components? Why does a specific gene expression cause a disease state? Which chaperones or post-translational modifications can contribute to homeostasis, and which ones are detrimental? All of these questions can be answered by building an understanding of protein-protein interactions and networks.

Protein-protein interactions (PPIs) can be defined in various ways, but the literature typically focuses on some mediated chemical or conformational changes when one protein comes in contact with another. Some other terms are often added to the definition, requiring an interaction to have a nonredundant function in some sense. Regardless, researchers typically drastically simplify the problem by treating it as a binary classification: Proteins either interact or do not. In a biological context, it is much more complicated than this, but good PPI classifiers are still informative towards the questions we mentioned above.

Recently, pLMs have received increasing attention for their ability to compare protein sequences and guess about interactions in biological contexts. Vast databases of positive interactors enable this type of analysis. However, confirming that two proteins for certain never interact is a much harder problem. Clever data science towards training PPI classifiers with massive inherent class imbalance is challenging, but there are many promising modern approaches. Hence, PPI classification is another way to partially uncover protein function computationally via pLMs.

Protein sequence generation

Lastly, we discuss one more general application of pLMs in protein sequence design and generation. As we discussed above, transformer decoders (or GPT models) are often trained to predict the next token, given some other tokens for context. The popularized ChatGPT does this incredibly well for the English language. Generative pLMs perform the identical task on amino acid tokens, generating sequences from scratch or completing sequence prompts. Many pLM projects are notable in this space.

- ProtGPT2: Stacked transformer decoders that generate viable nature-like sequences from scratch.
- ProGen: Stacked transformer decoders that generate plausible sequences given control tags for context, thus being able to generate sequences of a particular family or ontology.
- ANKH: A general-purpose encoder-decoder pLM that can generate proteins of a specific superfamily or plausible variants with possible increased functionality.
- ProtDT: A pLM fine-tuned by contrasting vector embeddings with an English language model to enable protein generation based on English natural language input.
- xtrimoGPLM: A massive general-purpose pLM that has an extremely capable sequence design. It can even generate sequences with nearly identical structures with almost no sequence similarity.
- PostT5: A fine-tuned version of ProtT5, which is also capable of sequence generation. This encoder-decoder architecture can generate a sequence given a structure input that approximates said structure. This is a bilingual model with amino acid and structure-based tokens.
- SAProt: An encoder-only system with a similar bilingual vocabulary as ProstT5 that can also translate between sequence and structure, enabling sequence generation based on a structure.

4.7.13.7: Concluding remarks

Protein language modeling is an interdisciplinary science at the intersection of bioinformatics, biochemistry, and computational sciences. Such modeling techniques are becoming an integral part of biochemical research through the fast-paced progress of NLP and computational hardware. It is easy to recognize the potential of protein language modeling in general life sciences: The generation of novel sequences for therapeutics, industrial catalysts, and synthetic biology, all the while annotating newly sequenced and generated proteins alike.

Importantly, amino acid-based vocabularies are not the only biochemically relevant uses of NLP models. DNA, codon, and even atom-wise vocabularies are being explored in many applications, including genomics, phylogenetics, and small-molecule-to-protein interactions. There are many avenues ready to explore in the biological NLP field.

As someone learning competency in biochemistry, it is important to remember the capabilities of protein language modeling while recognizing it is a relatively new and rapidly evolving field. Computational protein modeling may look vastly different in two years and will likely be completely different a decade from now. Regardless, computational tools serve as a helper to biochemists, not replacements. Standard biochemical assays to determine the structure and function of proteins will always be necessary to confirm and further inform the findings of a computational domain. This computational domain can simply weed out the plausible options for a given problem away from the implausible ones, enabling the impossible-seeming search of biochemical space.

4.7.13.8: Summary

Chapter Summary: Protein Language Models—Transforming Protein Structure and Function Prediction

In this chapter, we explore how modern computational tools, particularly protein language models (pLMs), are revolutionizing the prediction of protein structure and function from sequence data. Traditionally, protein structure determination via experimental methods (X-ray crystallography, NMR, cryo-EM) is both time-consuming and costly, even though sequencing has become routine. In response, the scientific community has developed massive sequence databases and powerful computational approaches to infer protein properties without requiring full experimental characterization.

At the heart of these advances are pLMs, which treat protein sequences as a biological language. These models are based on transformer architectures—attention-based neural networks that excel at processing sequential data. By tokenizing amino acid sequences (assigning each residue a unique numerical representation) and embedding them into high-dimensional vectors, pLMs capture both the semantic and contextual relationships among amino acids. The transformer's attention mechanism, particularly self-attention and multi-head attention, allows the model to weigh the relevance of each residue with respect to all others, which is crucial for understanding the complex, long-range interactions that dictate protein folding.

Training these models involves objectives such as masked token prediction (denoising) and next-token prediction, enabling the pLM to learn the "grammar" of protein sequences from vast corpora of data. This learned latent space not only correlates with protein structure but also with function, allowing the models to predict enzyme classifications (EC numbers), Gene Ontology (GO) annotations, and even protein–protein interactions.

A flagship application of pLMs is in 3D structure prediction, exemplified by AlphaFold. AlphaFold uniquely combines multiple sequence alignment (MSA) data with transformer-derived latent representations to build distograms and ultimately predict atomic-level protein structures. This breakthrough, along with complementary models like RoseTTAFold, has enabled high-accuracy structural predictions, even for proteins with little sequence homology ("protein orphans").

Beyond structure prediction, pLMs are applied to protein function prediction and even de novo protein design, generating novel sequences with potential applications in therapeutics, industrial catalysis, and synthetic biology. As interdisciplinary tools at the nexus of biochemistry, bioinformatics, and machine learning, pLMs illustrate how computational methods can drastically accelerate our understanding of protein biology.

While the rapid progress in protein language modeling offers unprecedented capabilities, the chapter also emphasizes that computational predictions must be validated with experimental biochemical assays. Ultimately, pLMs serve as powerful, complementary tools that help narrow down the vast biochemical search space, guiding experimentalists toward the most promising candidates for further study.

4.7.13.9: References

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. arXiv December 5, 2017. https://doi.org/10.48550/arXiv.1706.03762.

EMBL-EBI. Introduction to graph theory | Network analysis of protein interaction data. https://www.ebi.ac.uk/training/onlin...-graph-theory/ (accessed 2023-10-28).

Geetansh Kalra. Attention Networks: A simple way to understand Self Attention. https://medium.com/@geetkal67/attent...n-f5fb363c736d

DeepFindr. Understanding Graph Attention Networks. https://www.youtube.com/watch?v=A-yKQamf2Fc

Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. Highly Accurate Protein Structure Prediction with AlphaFold. Nature 2021, 596 (7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2.

Lin, Z.; Akin, H.; Rao, R.; Hie, B.; Zhu, Z.; Lu, W.; Smetanin, N.; Verkuil, R.; Kabeli, O.; Shmueli, Y.; dos Santos Costa, A.; Fazel-Zarandi, M.; Sercu, T.; Candido, S.; Rives, A. Evolutionary-Scale Prediction of Atomic-Level Protein Structure with a Language Model. Science 2023, 379 (6637), 1123–1130. https://doi.org/10.1126/science.ade2574.

Chen, B.; Cheng, X.; Geng, Y.; Li, S.; Zeng, X.; Wang, B.; Gong, J.; Liu, C.; Zeng, A.; Dong, Y.; Tang, J.; Song, L. xTrimoPGLM: Unified 100B-Scale Pre-Trained Transformer for Deciphering the Language of Protein. bioRxiv July 6, 2023, p 2023.07.05.547496. https://doi.org/10.1101/2023.07.05.547496.

Elnaggar, A.; Heinzinger, M.; Dallago, C.; Rehawi, G.; Wang, Y.; Jones, L.; Gibbs, T.; Feher, T.; Angerer, C.; Steinegger, M.; Bhowmik, D.; Rost, B. ProtTrans: Toward Understanding the Language of Life Through Self-Supervised Learning. IEEE Trans Pattern Anal Mach Intell 2022, 44 (10), 7112–7127. https://doi.org/10.1109/TPAMI.2021.3095381.

Elnaggar, A.; Essam, H.; Salah-Eldin, W.; Moustafa, W.; Elkerdawy, M.; Rochereau, C.; Rost, B. Ankh: Optimized Protein Language Model Unlocks General-Purpose Modelling. arXiv January 16, 2023. https://doi.org/10.48550/arXiv.2301.06568.

Su, J.; Han, C.; Zhou, Y.; Shan, J.; Zhou, X.; Yuan, F. SaProt: Protein Language Modeling with Structure-Aware Vocabulary. bioRxiv October 2, 2023, p 2023.10.01.560349. https://doi.org/10.1101/2023.10.01.560349.

Heinzinger, M.; Weissenow, K.; Sanchez, J. G.; Henkel, A.; Steinegger, M.; Rost, B. ProstT5: Bilingual Language Model for Protein Sequence and Structure. bioRxiv July 25, 2023, p 2023.07.23.550085. https://doi.org/10.1101/2023.07.23.550085.

Hallee, L.; Rafailidis, N.; Gleghorn, J. P. cdsBERT - Extending Protein Language Models with Codon Awareness. bioRxiv September 17, 2023, p 2023.09.15.558027. https://doi.org/10.1101/2023.09.15.558027.

Hallee, L.; Gleghorn, J. P. Protein-Protein Interaction Prediction Is Achievable with Large Language Models. bioRxiv June 9, 2023, p 2023.06.07.544109. https://doi.org/10.1101/2023.06.07.544109.

Ferruz, N.; Schmidt, S.; Höcker, B. ProtGPT2 Is a Deep Unsupervised Language Model for Protein Design. Nat Commun 2022, 13 (1), 4348. https://doi.org/10.1038/s41467-022-32007-7.

Liu, S.; Zhu, Y.; Lu, J.; Xu, Z.; Nie, W.; Gitter, A.; Xiao, C.; Tang, J.; Guo, H.; Anandkumar, A. A Text-Guided Protein Design Framework. arXiv February 9, 2023. http://arxiv.org/abs/2302.04611 (accessed 2023-02-14).

4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools** has no license indicated.

4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24)

Learning Goals (ChatGPT o1, 1/27/25

Below is a list of learning goals designed for junior and senior biochemistry majors to help them master the key concepts in predicting protein structure from sequence and the reverse protein folding problem:

1. **Understand the Evolution of Protein Structure Determination:**
   - Explain the shift from experimental techniques (NMR, X-ray crystallography, cryo-EM) to computational structure prediction.
   - Discuss how large databases (e.g., PDB) and evolutionary information enable structure prediction from protein sequences.
2. **Grasp the Fundamentals of the Protein Folding Problem:**
   - Describe the “protein folding problem” and why predicting 3D structure from a linear sequence is both challenging and crucial.
   - Compare traditional homology modeling with modern machine learning approaches (e.g., AlphaFold, RoseTTAFold).
3. **Evaluate Computational Structure Prediction Methods:**
   - Interpret and compare metrics (RMSD, TM-score, ipTM, pTM) used to assess the accuracy of predicted protein structures and complexes.
   - Explain how AI models leverage hidden evolutionary signals in sequences to generate accurate 3D models.
4. **Explore Protein-Protein Interaction Prediction:**
   - Understand how tools like AlphaFold3 are used to predict the quaternary structures of protein complexes.
   - Evaluate the statistical comparisons between experimentally determined and predicted complex structures.
5. **Analyze the Reverse Protein Folding Problem:**
   - Describe methods for deducing protein function from 3D structure when sequence homology is limited.
   - Explain how structural comparison tools (e.g., FoldSeek) utilize a 3D interaction alphabet to compare protein structures.
6. **Understand De Novo Protein Design:**
   - Outline the steps involved in designing a protein from scratch (e.g., using RFDiffusion and ProteinMPNN).
   - Describe how conditioning on functional motifs or binding sites can lead to the creation of novel protein binders, catalysts, or symmetric assemblies.
7. **Assess Applications and Case Studies:**
   - Review examples such as de novo designed IL-6 mimetics, influenza HA binders, pentameric helical bundles, and symmetric oligomers.
   - Discuss how these designed proteins have potential applications in therapeutics, vaccines, biosensors, and industrial catalysts.
8. **Critically Evaluate the Integration of Computational and Experimental Methods:**
   - Recognize the importance of experimental validation (e.g., X-ray crystallography, cryo-EM, NMR) for confirming computational predictions.
   - Reflect on how computational tools are used to narrow down plausible candidates for experimental testing, not replace bench work.
9. **Consider Future Challenges and Directions:**
   - Identify remaining challenges in designing effective binders, catalysts, and proteins with dynamic conformational flexibility.

- Discuss the potential impact of continued advancements in AI and machine learning on the field of protein engineering and functional annotation.

These learning goals will guide you through understanding how modern computational techniques—especially protein language models and transformer architectures—are transforming our ability to predict, analyze, and design protein structures and functions.

Recent Updates: November 2024

4.7.14.1: The Protein Folding Problem: Sequence to 3D Structure

In Chapter 3.4: Analyses of Protein Structure, we discussed how protein structure can be experimentally analyzed and the 3D structure of a protein determined using NMR, X-ray crystallography, and cryo-EM. Now, using sequence and structural databases (like the PDB with over 227,000 structures), we can often predict the 3D structure of a protein just from its linear sequence by comparing the sequence of a protein of unknown tertiary structure to homologous proteins (by sequence) whose 3D structures are known. Machine learning and artificial intelligence have extended earlier and simpler "homology modeling" attempts to allow structural predictions for millions of protein sequences using programs such as RoseTTAFold and AlphaFold. RoseTTAFold and AlphaFold produce high-quality structure predictions when trained using the vast sequence information in the Protein Data Bank. Embedded in those linear sequences is a large amount of hidden (to the human eye) evolutionary information that machine learning and AI can harness to predict 3D structures. They work less well when very limited sequence comparisons are available.

In general (for smaller proteins), the protein folding problem, the prediction of structure from sequence, appears to have been "solved". The Nobel Prize in Chemistry in 2024 was awarded to Demis Hassabis and John M. Jumper from Google DeepMind for developing AlphaFold and David Baker for developing RoseTTaFold and other powerful techniques described below. (Check out the previous section, Chapter 4.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools, if you are interested in a "deep dive" into how these programs work!)

A comparison of protein structures obtained using these programs with known 3D structures obtained through X-ray crystallography or other techniques shows them almost identical. Different metrics can be used to compare predicted structures to the actual ones. The root mean squared deviation (RMSD) is a common one. RoseTTAFold uses a **TM-score to assess** the topological similarity of protein structures. Compared to RMDS, the TM-score weights smaller distance errors higher than larger ones, making it sensitive to the global fold, not local structural differences. TM values range from 0-100 (100 is a perfect match). Scores below 17 indicate no topology match, while those above 50 suggest a common fold.

AlphaFold uses a "neural network, meaning it simultaneously considers patterns in protein sequences, how a protein's amino acids interact, and its possible three-dimensional structure. In this architecture, one-, two-, and three-dimensional information flows back and forth, allowing the network to collectively reason about the relationship between a protein's chemical parts and its folded structure". Programs of this type might allow the generation of proteins with new therapeutic or commercial potential based on sequences. These include vaccines, sensors, specific immune system suppressors or activators, and antivirals. AlphaFold has now been used to predict the structure of 214 million proteins from more than one million species — essentially all known protein-coding sequences. We have included many AlphaFold iCn3D models throughout this book.

- AlphaFold Database - Protein Structure Database

Figure **4.7.14.1** shows the backbone tube cartoon of the x-ray pdb structure of the small protein (1xww, cyan) and the structure predicted by both RoseTTAFold program and AlphaFold (magenta) just from its primary sequence. Sulfate, a competitive inhibitor, is shown (spacefill) bound in the active site. The alignment is spectacular, except for the N-terminal 5 amino acids at the bottom of

the figure (6 o'clock). This stretch has more disorder even in the x-ray structure as the amino acids have high B-factors, indicating more conformational flexibility.

Figure **4.7.14.1**: Comparison of the x-ray and computationally predicted structures of human low molecule weight protein tyrosine phosphatase

Left panel: X-ray structure (cyan) of low molecular weight protein tyrosine phosphatase with bound SO_4^{2-} (1xww) and corresponding structure predicted by the RoseTTAFold (magenta). Right panel: Same structures using AlphaFold for the structural prediction.

4.7.14.2: Prediction of Protein-Protein Interactions

AlphaFold3 has also been used to predict the structure of protein complexes in which multiples of the same or different proteins combine to form a larger, quaternary structure. Figure **4.7.14.2** shows an interactive iCn3D model of a recent stunning example of a predicted AlphaFold complex required to bind a human sperm and egg. The complex consists of three transmembrane human sperm proteins and a human egg protein attached to the egg membrane with a posttranslational lipid anchor (not shown).

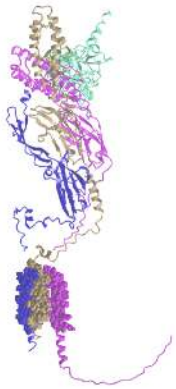


Figure \(\PageIndex{2}\): Human sperm proteins and egg protein complex predicted by AlphaFold. (Copyright; author via source). The spacefill section of the three sperm proteins attached the proteins to the sperm membrane. The egg protein is JUNO (cyan) and the three sperm proteins are 1ZUMO1 (magenta), SPACA6 (brown/gold), and TMEM81 (blue)

Reference for PDB file: Deneke, V. et al. A conserved fertilization complex bridges sperm and egg in vertebrates. Cell. October, 2024. https://doi.org/10.1016/j.cell.2024.09.035. Creative Commons Attribution (CC BY 4.0)

You can download this iCn3D file and load it in iCn3D using these commands to see the structure as rendered in the image above: **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

- Open iCn3D
- **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

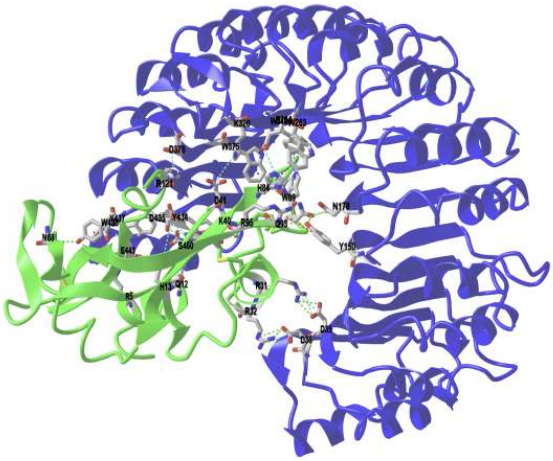
Here is the link to the AlphaFold Server 3. AlphaFold3 is based on a machine-learning process called diffusion, which is explained in more detail below.

The following biological species can be modeled in AlphaFold3:

- macromolecules, including proteins, DNA and RNA
- common ligands including ATP, ADP, AMP, GTP, GDP, FAD, NADP, NADPH, NDP, heme, heme C, myristic acid, oleic acid, palmitic acid, citric acid, chlorophylls A and B, bacteriochlorophylls A and B
- common ions such as Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{3+}, K^{+}, Mg^{2+}, Mn^{2+}, Na^{+}, Zn^{2+}, and Cl^{-}
- common post-translational modifications (PTMs) of amino acid residues such as phosphorylation of serine, threonine, tyrosine, and histidine, acetylation of lysine residues, methylation of lysine and arginine, malonylation of cysteine, hydroxylation of proline, lysine, and asparagine, palmitoylation of cysteine, succinylation of asparagine, S-nitrosylation, formylation of tryptophan, crotonylation of lysine, citrullination of lysine and arginine
- glycan chains (including branched chains) composed of some sugars, including alpha/beta-D-glucose, alpha/beta-D-mannose, alpha-L-fucose, beta-D-galactose, N-acetyl-beta-D-glucosamine
- common chemical modifications of the DNA (including methylation of cytosine, guanine, and adenine, carboxylation of cytosine, oxidation of guanine, formylation of cytosine) and RNA (including isomerization of uridine into pseudouridine, formylation of cytosine, and methylation of cytosine, guanine, adenine, and uracil
- structures composed of multiple proteins, nucleic acids, ligands, ions, and chemically modified derivatives.

Note that simple drugs are not on the list since those applications are proprietary (at present). AlphaFold 3 can now be downloaded for academic (non-commercial) applications and likely includes drugs. A more limited web version is available here.

It is important to statistically compare the PDB experimentally determined and AF3-predicted structures of complexes. One such comparison involves the wild type and their mutated forms, which should involve just subtle conformational changes. For example, consider human angiogenin and placental ribonuclease inhibitors. Figure 4.7.14.3 shows an interactive iCn3D model of the experimental human angiogenin and placental ribonuclease inhibitor complex (1A4Y).



NCBI iCn3D Figure 4.7.14.3: Human angiogenin-placental ribonuclease inhibitor complex (1A4Y). (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...DaUAJzhh8H6LU8

There are 27 mutant variants of the complex whose experimental structures are known. They are included in the SKEMPI database, which contains thermodynamic and kinetic data for wild-type and mutant complexes with known PDB structures. One widely used thermodynamic parameter we have seen before is the change in thermodynamic stability (i.e. $\Delta\Delta G^0 = \Delta G^0_{mutant} - \Delta G^0_{wildtype}$), in this case for the complex, when key residues lining the binding pocket are mutated. The entire database contains over 317 protein-protein complexes and 8338 mutations. How well does AF3 predict the structure of mutant complexes?

Three statistical values are used to compare the experimental and AF-predicted complex structures:

- RMSD (Root-Mean-Square Deviation) measures the average distance between equivalent atoms in the complex subunits. Lower values show great similarity between the experiential and AlphaFold 3 models.
- ipTM (Interface Predicted Template Model) measures changes in the interface of the subunits between the experimental and AF3 structures. Higher values indicate a closer match.

- pTM (Predicted Template Model), as with simple AF prediction, measures the overall accuracy of the predicted structure (based on both backbone and sidechain orientations). Higher pTM values indicated a more accurate prediction.

Figure 4.7.14.4 shows the structure of the wild-type RNase inhibitor-Angiogenin complex with 27 red dots indicating mutants with experimental structures (Panel A) and a comparison of the wild-type and a mutant structure (Panel B). Statistical results comparing experimental and AF structures for all 317 complexes in the SKEMPI database are shown in Panels C and D.

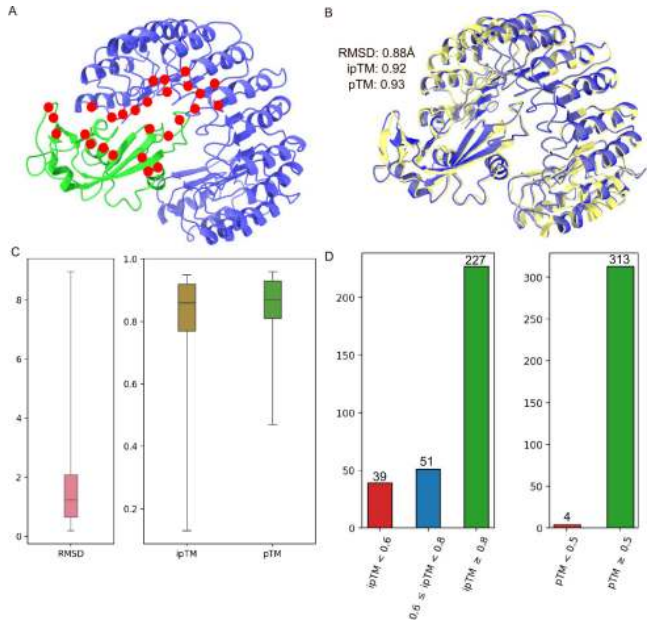


Figure 4.7.14.4: Wildtype and AF predicted composite structure of RNase inhibitor-Angiogenin complexes (Panel A and B) and statistical comparison of all structures in the SKEMPI database (panels C and D). JunJie Wee and Guo-Wei Wei. J. Chem. Inf. Model. 2024, 64, 16, 6676–6683. https://doi.org/10.1021/acs.jcim.4c00976. Published August 8, 2024. CC-BY 4.0 .

Panel A: The cartoon representation of ribonuclease inhibitor-angiogenin complex (PDB ID: 1A4Y). The ribonuclease inhibitor is shown in blue, and the angiogenin is shown in green. 27 mutation spots of 1A4Y in the S8338 data set are indicated in red.

Panel B: The structural alignment of 1A4Y with its AF3 predicted complex.

Panels C and D below are based on the complexes studied, not just the RNase Inhib-Angiogenin complex.

Panel C: The boxplot for RMSD, ipTM and pTM distributions of 317 predicted AF3 protein–protein complexes. RMSDs refer to the overall RMSD calculated by structurally aligning an AF3 complex with its original PDB complex.

Panel D: The breakdown of AF3 protein-protein complexes based on their ipTM and pTM scoring criteria.

The results for average statistical values for the wildtype and mutant structures are 1.61 Å (RMSD), 0.803 (ipTM), and 0.847 (pTM), respectively. The graphs clearly show that most predictions (72%) had high ipTM scores ($\geq$ 0.8), while 99% had reasonably high pTM scores ($\geq$0.5). However, they were outliers, as shown in Panel C, with RMSD values> 4 Å, which indicates poorer performance with AF3. Most complexes with high prediction values also had low RMSD values. Experiments like these will be used to continually refine programs such as AlphaFold 3.

4.7.14.3: Reverse Protein Folding Problem: 3D Structure to Function

So now we have structures of 200 million plus proteins. Pick one, Protein X, of unknown function. What might its function be? Surely, its sequence could be compared to the entire database to find homologous proteins that might give a clue to the function of Protein X. But what if the sequence of Protein X is very divergent from potential sequence homologs, since they are very distant from each other evolutionarily? Also, what if comparison proteins in the database are underrepresented? For example, our knowledge of the sequences and structures of proteins from pathogens (viruses and bacteria) is incomplete, especially since we have studied just a small fraction of the virus and bacterial world.

To circumvent this problem, the actual predicted or determined 3D structure of Protein X (not its sequence) could be compared to the 3D structures from the databases. This should be difficult since it would be a 3D comparison, not a 1D comparison of linear sequences. A program called **FoldSeek** allows 3D structural comparisons to be made computationally more easily.

Instead of using an alphabet of actual sequences (such as the single letter code for the 20 naturally occurring amino acids - ACDEFGHIKLMNPQRSTWYV), a new "structural alphabet" based on the conformations of short stretches of 3-5 alpha C (C_α) atoms in the protein backbone has been used, but this doesn't explicitly contain tertiary interactions found in proteins. Instead, FoldSeek uses a 3D interaction alphabet (3Di) with 20 states (one for each amino acid), each with 10 interaction "features." A conformational state for residue X is defined for its closest spatial residue, y. The state description is less dependent on the next amino acid in the linear sequence for a given amino acid. The defined state has more information when x is in a conserved and packed protein core than in a nonconserved, more flexible loop. In contrast, there would be less information if just the backbone structural alphabet were used. Figure **4.7.14.5** below gives a pictorial view of how the 3Di state for a single amino acid Val at a specific position in a 3D structure is defined. Note that the state has 10 3D features, more than just the conformation of a backbone of 3 amino acids or the next amino acid in the linear sequence.

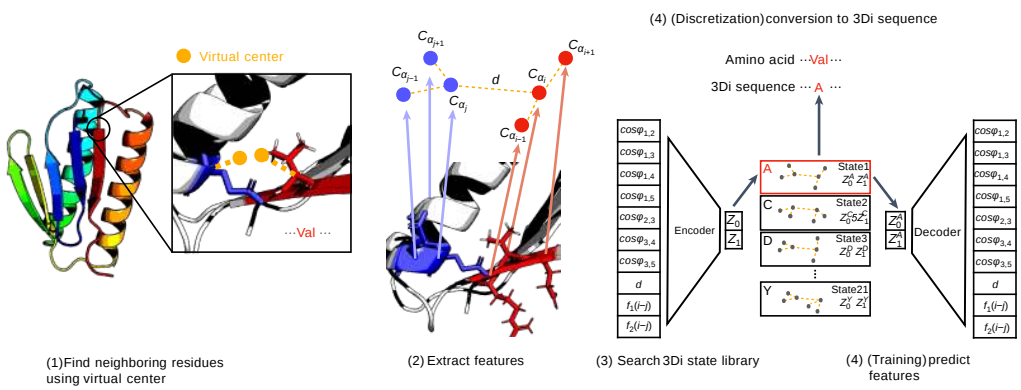


Figure **4.7.14.5**: Learning the 3Di alphabet. van Kempen, M., Kim, S.S., Tumescheit, C. *et al.* Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* **42**, 243–246 (2024). https://doi.org/10.1038/s41587-023-01773-0. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

(1) 3Di states describe the tertiary interactions between a residue *i* and its nearest neighbor *j*. Nearest neighbors have the closest virtual center distance (yellow). Virtual center positions were optimized for maximum search sensitivity. (2) To describe the interaction geometry of residues *i* and *j*, we extract seven angles, the Euclidean C_α distance, and two sequence distance features from the six C_α coordinates of the two backbone fragments (blue and red). (3) These 10 features are used to define 20 3Di states by training a VQ-VAE modified to learn states that are maximally evolutionarily conserved. The encoder predicts the best-matching 3Di state for structure searches for each residue.

Recently, FoldSeek has been used to find structural and functional similarities of over 67,000 newly predicted viral proteins (underrepresented in the PDB) with other proteins of known structure. Of these:

- 62% had distinct structures and were not homologous to proteins in the AlphaFold database (as we indicated above).
- Many of the 38% left were structurally homologous to nonviral proteins, suggesting a similarity in viral protein function to host analogs.

Similar 3D structures imply similar functions, so probable functions could be described for some novel proteins. Some were involved in the viral escape from the host's innate immune system. We'll explore that in Chapter 5.4: Complementary Interactions between Proteins and Ligands - The Immune System.

The FoldSeek server allows multi-database searches, including AlphaFoldDB (version 4: Proteomes and Swiss-Prot), AlphaFoldDB (version 4) and CATH[25] clustered at 50% sequence identity, ESM Atlas-HQ and Protein Data Bank (PDB). These Google Colab sites are also available:

- Foldseek Viral Structure Alignment

- Foldseek Viral Structure Visualization

In summary, FoldSeek is useful in several circumstances:

- you have a protein sequence, but a comparison to other sequences doesn't give you enough information. A structure-based search would then be helpful.
- you want to design a brand new protein (de novo protein synthesis), and you would want to know that its structure is not similar to other proteins
- you want to design a protein with a particular function, and want to compare its structure to other proteins with similar function

4.7.14.4: Reverse Protein Folding Problem: 3D Shape to Linear Sequence - Designing Proteins From Scratch

In yet another use of machine learning and artificial intelligence, programs can start with a desired 3D shape (protein backbone, for example) and determine the amino acid sequence necessary to get it. Two programs, ProteinMPNN and RoseTTAFold Diffusion (RFDiffusion), developed by David Baker (who also won the Nobel Prize in Chemistry in 2024) et al, have enabled these predictions. It allows protein structure design, not structure prediction.

Yet another dream that seemed so distant not so long ago was to design a protein from scratch with no linear sequence (hence little alignment information) but with a final desired structure or function in mind. Here are some possible "de novo" design examples of novel proteins that...

- are soluble versions of a known membrane protein, which could advance drug design;
- bind with high affinity to a desired small molecule (much like an antibody), enabling the creation of sensors and protective agents;
- bind target molecules and catalyze their chemical conversion to products, allowing the development of new and nontoxic catalysts;
- bind to another target protein and modulate its function by activating or inhibiting it;
- have novel, unrepresented folds that could further elucidate key principles of protein folding and stability while creating new functionalities.

RFDiffusion

This dream has also been accomplished in large measure. David Baker is a pioneer in *de novo* protein structure design and prediction. His group has developed and used several programs, including RoseTTAFold Diffusion (RFDifffusion), which uses AI to design new proteins with novel structures and functions. RFDiffusion is freely available to anyone for use in Google Collaboratory. It creates new structures by combining structure prediction from RoseTTAFold with an AI "Diffusion" model.

To understand the term diffusion in structure prediction, let's first explore AI/machine learning models for generative image creation. Instead of starting with no previous information, start with a clear image, add random (Gaussian) noise to it (noising), and then try to recreate the original image by a "denoising" process. Some previous information and additional programs would help to constrain the denoising process for generative image creation for a requested image. This process is illustrated in Figure 4.7.14.6 below. Note that the arrows are reversible.

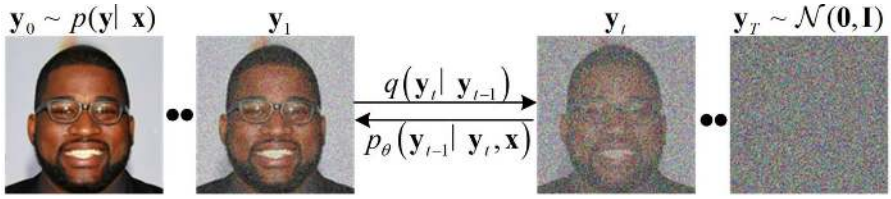


Figure 4.7.14.6: Xiao, H., Wang, X., Wang, J. *et al*. Single image super-resolution with denoising diffusion GANS. *Sci Rep* **14**, 4272 (2024). https://doi.org/10.1038/s41598-024-52370-3. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Simplistically, this is similar to solving X-ray structures of proteins. A given protein in a crystal produces an X-ray diffraction pattern specific to the atoms and their arrangement in the crystal lattice. In the reverse process, the X-ray diffraction

pattern can be computationally analyzed to produce the arrangement of atoms (from an initial electron density map) in the lattice that would generate the given diffraction pattern.

In a diffusion model for generative protein structure creation using RFDiffusion, randomly disordered small chemical fragments diffuse together to form a more ordered and realistic protein structure. Information from the database of known protein structures is used to constrain the generative processes through a deep learning–based protein sequence design method called **ProteinMPNN (Protein Message-Passing Neural Network)**. It differs from Rosetta, a physically-based method that maximizes sidechain packing to produce the lowest energy state. Designing a sequence that produces the lowest energy state is more computationally challenging than finding it for a given sequence. Calculating energies of unwanted nonproductive oligomeric and aggregated states makes this approach intractable.

What is more doable is to carry out these two steps in succession:

- first, search for the lowest-energy sequence for a given backbone structure;
- then search the "universe" of possible structures for the sequence created in the first step to determine if it is indeed the lowest energy conformation.

Methods like Rosetta use physical "rules" to minimize undesired results. For example, restrictions are used in placing hydrophobic side chains on the surface of a protein, as these might promote unwanted aggregation states. ProteinMPNN overcomes these issues since it uses data from all solved structures to find the most probable amino acid at a given position. It requires less human theoretical knowledge as it extracts an energy-minimized folded state from an immense amount of structural data. It's a bit like deriving Newton's Laws of Motion from data without the underlying theory, even though the data was acquired from systems whose motions and positions are well described by Newton's Laws.

Figure 4.7.14.7 below shows the noising (right to left) and denoising (left to right) processes that can generate a protein structure in a diffusion model. It parallels the image deconstruction and reconstruction shown in Figure 4.7.14.6 above.

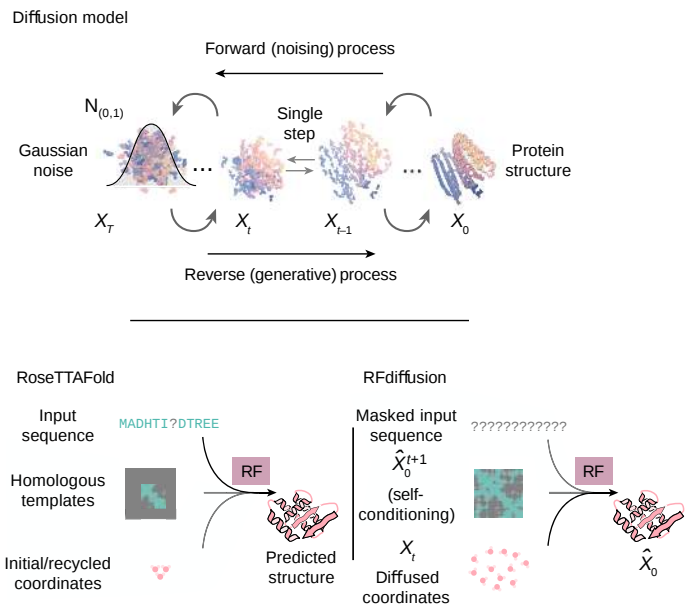


Figure 4.7.14.7: **Protein design using RFdiffusion.** Diffusion models for proteins are trained to recover corrupted (noised) protein structures and to generate new structures by reversing the corruption process through iterative denoising of initially random noise X_T into a realistic structure X_0 (top panel). Watson, J.L., Juergens, D., Bennett, N.R. *et al*. De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089–1100 (2023). https://doi.org/10.1038/s41586-023-06415-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

RFDiffusion can create a protein sequence without introducing conditions to the final structure (unconditional process). Conditions placed on the denoising process can lead to conditioned structures. Conditions such as symmetric noise, a binding target, a functional motif such as pre-positioned amino acids in an active site, and a symmetric motif can lead to synthetic oligomers, a

binder protein that interacts with a target protein, an active site with the correct 3D disposition of catalytic residues, and symmetrical scaffolds, respectively. These examples are illustrated in Figure 4.7.14.8 below.

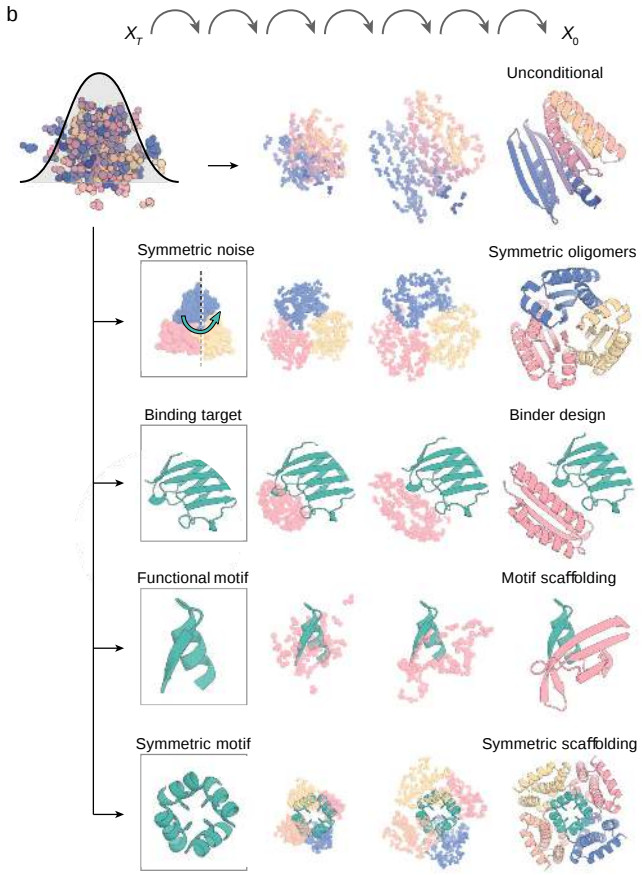


Figure 4.7.14.8: **b**, RFdiffusion is broadly applicable for protein design. RFdiffusion generates protein structures without further input (top row) or by conditioning on (top to bottom): symmetry specifications; binding targets; protein functional motifs, or symmetric functional motifs. In each case, random noise and conditioning information are input to RFdiffusion, which iteratively refines that noise until a final protein structure is designed. Watson, J.L. et al., ibid.

Figure 4.7.14.9 below shows that the final structure predicted for a 300 amino acid protein sequence by AlphaFold (bottom row) is almost identical to the final structure produced by RFDiffusion (top row).

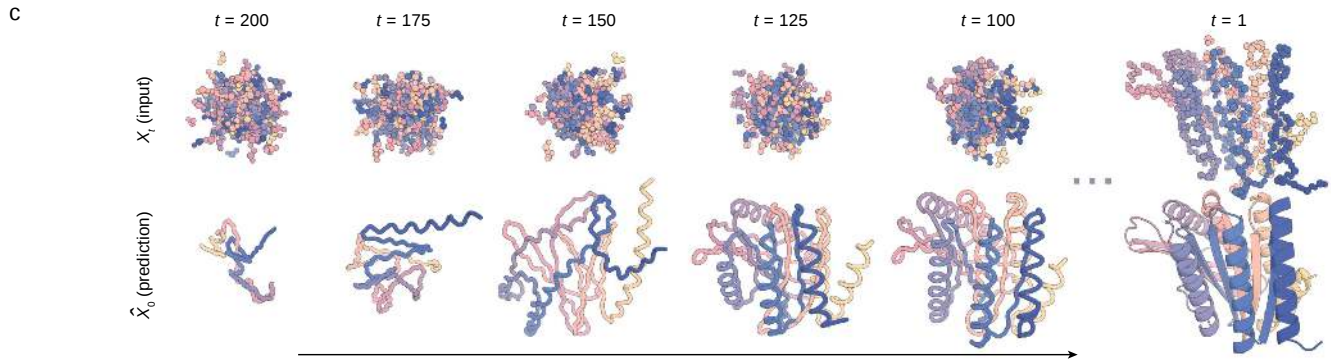


Figure 4.7.14.9: An example of an unconditional design trajectory for a 300-residue chain, depicting the input to the model (X_t) and the corresponding X^0 prediction. At early timesteps (high t), X^0 bears little resemblance to a protein but is gradually refined into a realistic protein structure. Watson, J.L. et al., ibid.

Unconditional RFDiffusion models for protein sequences up to 600 amino acids are essentially the same as AlphaFold's.

Figure 4.7.14.10 below shows how hot spots (key binding residues) in a target protein are used as a condition in an RFDiffusion model in the *de novo* synthesis of a mini-binder for a target protein.

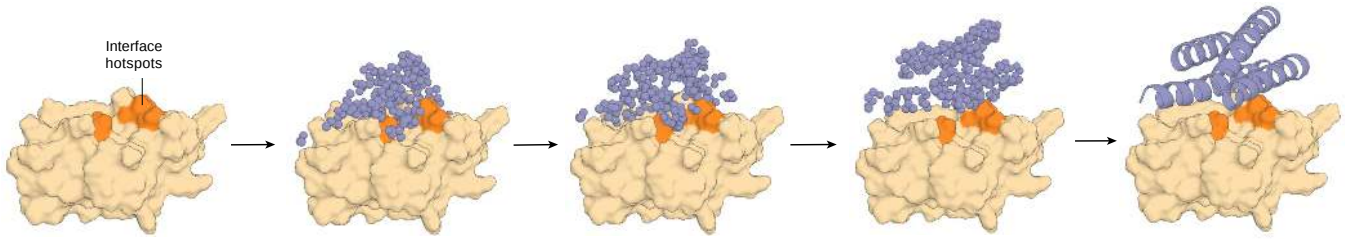


Figure 4.7.14.10: RFdiffusion generates protein binders given a target and specification of interface hotspot residues. Watson, J.L. et al., ibid.

The video below from the Baker lab (obtained from YouTube at https://youtu.be/geqlzPsigQo) shows how a protein structure can be created that binds to a predefined structure, in this case, the insulin receptor, using conditional RFDiffusion.

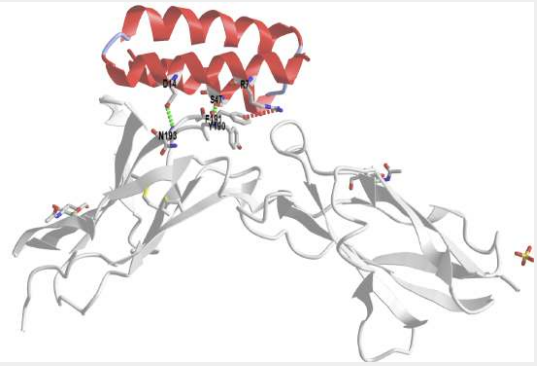
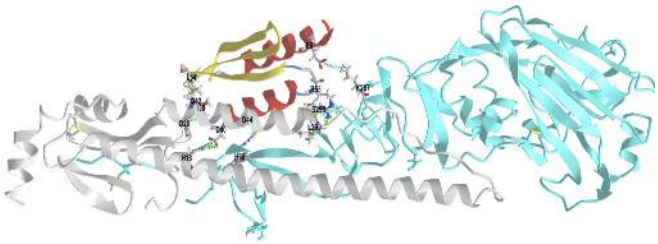
The same video is found at the Baker site at: https://www.bakerlab.org/2023/07/11/...rotein-design/

The structures produced by RFDiffusion and ProteinMPNN for any given sequence can be verified by making the protein and analyzing its structure using X-ray crystallography, NMR, or cryoEM. Less precise methods, such as CD-spectroscopy, are also used to get a simpler measure of the overall secondary structure of the synthesized protein.

Examples

The following iCn3D models of crystal structures using these methods illustrate the power of RFDiffusion methods in creating new

proteins of defined structure and function.

Examples of *de novo* protein design	Interactive iCn3D model with links
GP130 (IL6 coreceptor) in complex with a *de novo* designed IL-6 mimetic (8UPA) Cytokine storms (cytokine release syndrome) are often deadly inflammatory responses accompanying bacteria or viral infections (such as COVID-19 in those with severe disease). The storm is associated with the overexpression and release of two proinflammatory cytokines, interleukin 1 (IL-1) and interleukin 6 (IL-6), by activated immune cells such as macrophages. IL-1 and IL-6 bind to their receptors (IL-1R and IL-2R) with high affinity. IL-6 also binds to a coreceptor (GP130) needed for cytokine release. Inhibitors used to interfere with the cytokines:receptor complex can persist too long and have deleterious effects, since an appropriately amplified immune response is needed against bacteria and viral infections. *A* small protein antagonist (a minibinder or MB) with a high affinity (pM to nM dissociation constant) for the receptor and the IL-6 coreceptor was made through *de novo* design and proved protective against a cytokine storm in animal models. The structure of the IL-6 mimetic with its coreceptor, GP130, is shown in the iCn3D model to the right. The computationally designed structure was a close match to the X-ray structure. Reference: Huang, B., Coventry, B., Borowska, M.T. *et al.* De novo design of miniprotein antagonists of cytokine storm inducers. *Nat Commun* **15**, 7064 (2024). https://doi.org/10.1038/s41467-024-50919-4	 NCBI iCn3D Figure **4.7.14.11**: GP130 (IL6 coreceptor, gray) in complex with a de novo designed IL-6 mimetic (8UPA, red). (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...rAQXeMxWaq74Y6 Here is another link to see a surface representation of the two interacting proteins: https://structure.ncbi.nlm.nih.gov/i...3bbFWhy2eHNGAA
Designed Influenza HA binder, HA_20, bound to Influenza HA (8SK7) Hemagglutinin (HA) from the influenza virus is a trimeric membrane protein. Each "monomer" is a heterodimer consisting of two different chains, HA1 and HA2. The HA1 domain binds to a particular sugar, sialic acid, found on many human cells, but most importantly in the respiratory tract. The HA2 subunit is transmembrane. We receive a vaccine each year that recognizes the HA protein since the globular head part of HA that interacts with the human cell surface mutates so quickly from year to year. Large shifts in the structure of the influenza HA protein lead to pandemics. Parts of the HA molecule are more conserved and are somewhat sequestered from the human immune system. Targeting them could lead to a more permanent and universal vaccine. A small influenza binder was synthesized de novo and tightly bound (nanomolar dissociation constant). The de novo-designed protein had essentially the same structure as the AlphaFold computational model. An iCn3D model showing the interaction of the HA binding with one HA1:HA2 heterodimer is shown to the right. Reference: Nature 620, 1089–1100 (2023). https://doi.org/10.1038/s41586-023-06415-8. Creative Commons	 NCBI iCn3D Figure **4.7.14.12**: **Designed Influenza HA binder, HA_20, bound to Influenza HA** (8SK7). (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...xaAUkDPBBUeob7 The gray is HA2, the cyan is the HA1, and the red/yellow-coded secondary structure is the designed HA minibinder. The biological HA complex contains three copies of the heterodimeric structure shown above Here is another link to see a surface representation of the three interacting proteins: https://structure.ncbi.nlm.nih.gov/i...a9LQq3SjGXp9JA

Attribution 4.0. International License. http://creativecommons.org/licenses/by/4.0/.

Pentameric helical bundle protein (8U5W)

De Novo protein synthesis was used to create a single protein chain (i.e., a monomeric protein with a single C5 rotational symmetry axis. Open the iCn3D model to the right. It contains a single rotational axis (red line). Rotation around the axis by $360^0/5$ reproduces the identical structure.

The designed protein also displays near-infrared fluorescence when it binds to a synthetic dye, merocyanine. The protein forms a covalent Schiff base with the dye. If the Schiff base is protonated, the fluorescence spectra show a large red shift in both the excitation and emission wavelengths. The protein/dye complex can be used for tissue imaging at a greater depth than other visible light fluorophores.

Reference: https://www.researchsquare.com/article/rs-4652998/v1

NCBI iCn3D Figure 4.7.14.13: **Pentameric helical bundle protein (8U5W)**. (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...Xm7aBQhm2GEKWA

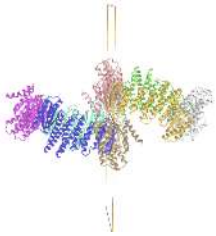
Symmetric Oligomers

In contrast to the previous example of a symmetric monomer, the *de novo* protein models to the right contain multiple subunits in oligomers that display different types of cyclic symmetry.

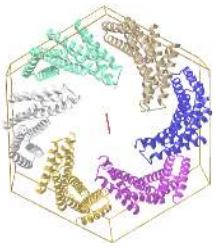
Figure 4.7.14.14 to the right displays C2 symmetry, with a rotation of $360^0/2$ around the axis resulting in an identical structure. The dimer also displays allostery - it changes its shape globally with the addition of effector molecules. Allostery will be explained in Chapter 5.

Figure 4.7.14.15 to the right is a homo 6-mer of identical subunits and C6 symmetry. Rotation of $360^0/6$ around the axis results in an identical structure.

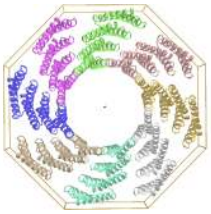
Figure 4.7.14.16 to the right is a homo 8-mer of identical subunits and C8 symmetry. Rotation of $360^0/8$ around the axis results in an identical structure.



NCBI iCn3D Figure 4.7.14.14:Allosterically Switchable De Novo Protein sr322, In Closed State (8UTM). (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...HSZ7Ddt8v9R3J9



NCBI iCn3D Figure 4.7.14.15:Designed modular protein oligomer C6-79 (8f6r). (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...SDwnTmNhovAny9



NCBI iCn3D Figure 4.7.14.16:Designed modular protein oligomer C8-71 (8f6q).. (Copyright; author via source). Click the image for

a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...ydRkeJAaGGag38

A protein with an active site

A protein was designed using RFDiffusion that recreates an active site containing three key catalytic residues from a native enzyme, cytotoxic ribonuclease alpha-sarcin (1DE3).

The left model in Figure 4.7.14.17 below and the iCn3D model in Figure 4.7.14.18 in the adjacent right column show the three active site residues used for conditional de novo protein modeling. The middle two images below show the isolated input catalytic "triad" and the structure created by RFDiffusion. The right image below is a zoomed image of the active site in the designed protein.

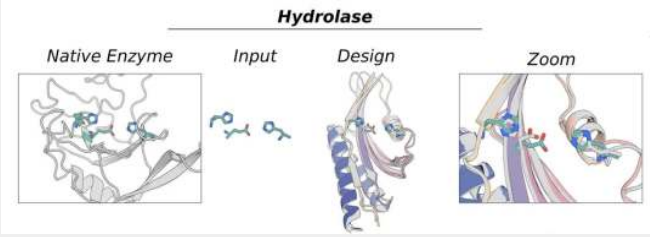


Figure 4.7.14.17: Comparison of native ribonuclease sarcin and RFDiffusion designed protein. Supplemental Figure, Watson, J.L. et al., ibid.

This iCn3D model is for cytotoxic ribonuclease alpha-sarcin (1DE3). Three active site amino acids, H50, E96, and H137, were conditionally used to create the *de novo* protein with the same active site residues (shown to the left).

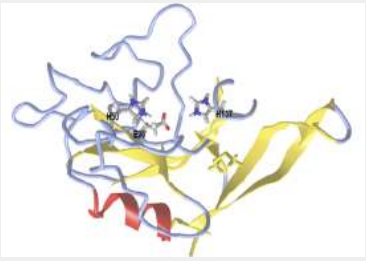


Figure 4.7.14.18: Cytotoxic ribonuclease alpha-sarcin (1DE3). (Copyright; author via source). Click the image for a popup or use this link: https://structure.ncbi.nlm.nih.gov/i...WR1oo3fru633d6

In Chapter 11.1, we will explore how RFDiffusion can create novel membrane proteins and soluble versions.

One final comment: Structures predicted by these AI programs must be subjected to experimental validation of structure and function. Since creating new structures with designed functions is so easy, we must be careful not to blindly accept the results without supporting experimental validation.

AlphaProteo

AlphaProteo from Google DeepMind is also used to design protein binders for target sites on proteins. Download this file for a video of a synthesized protein binder designed for the SARS-CoV-2 spike receptor-binding domain (reference). This program is not yet freely available (as of 11/11/24) for use. The machine learning methods used in AlphaProteo were not reported in the preprint reference because of "biosecurity and commercial considerations," so we can't explain the basis of the program as we did above for RFDiffusion. Figure 4.7.14.19 below shows, in general, the steps involved in developing binders that interact with "hotspots" sites on target proteins.

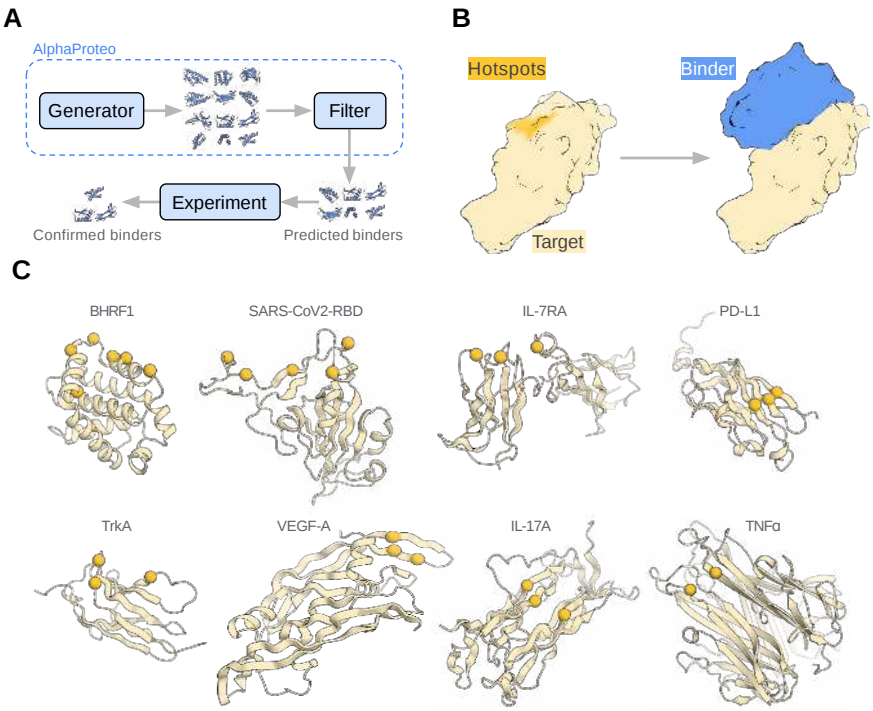


Figure 4.7.14.19: Overview and experimental performance of AlphaProteo. Vinicius Zambaldi et al. De novo design of high-affinity protein binders with AlphaProteo. Submitted 9/12/24. https://arxiv.org/abs/2409.08022. https://creativecommons.org/licenses/by-nc-sa/4.0/

Panel (A) Schematic of the design system. The generative model outputs designed structures and sequences of binder candidates, and the filter is a model or procedure that predicts whether a design will bind.

Panel (B) Schematic of target-structure-conditioned binder design as performed by the generative model.

Panel (C) Crystal structures (light yellow) and hotspot residues (dark yellow spheres) of seven target proteins for binder design experiments in this work. VEGF-A and IL-17A are both disulfide-linked homodimers. See Table S1 for PDB IDs and hotspot residue numbers.

Figure 4.7.14.20 below shows the interactions of the de novo synthesized binder with seven target proteins.

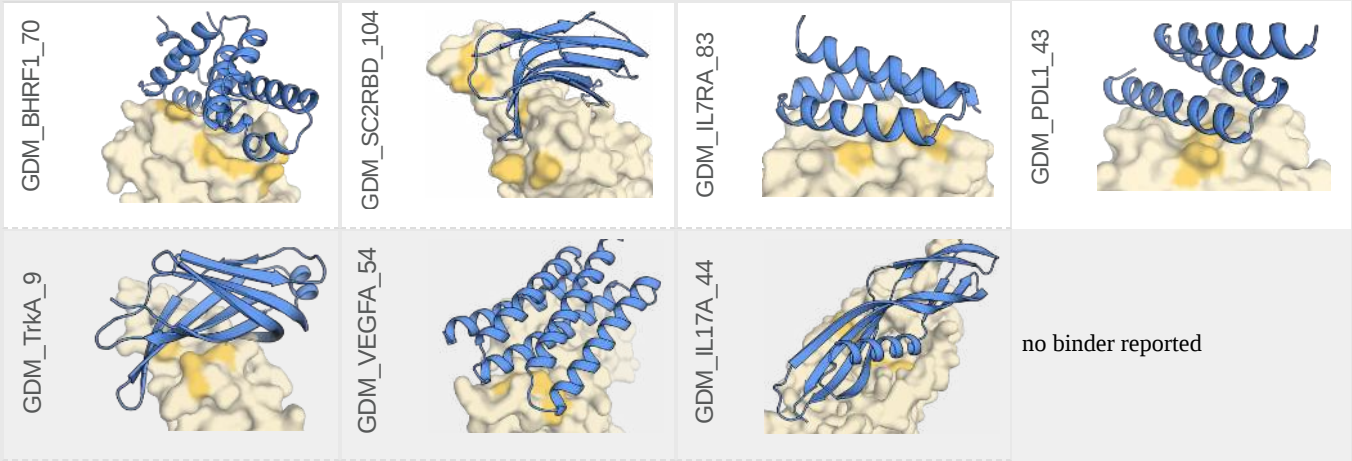


Figure 4.7.14.20: Biochemical characterization of representative binders for each target-design model.

The binders all interacted tightly with their target protein.

4.7.14.5: Challenges that remain

Here are some examples that pose challenges

- Binders that affect the function of a protein: These include both small-molecule binders (i.e., drugs) that target orthosteric or allosteric sites. Essentially, this is the task of the drug and pharmaceutical industry. Designing binders is especially hard for membrane proteins. Also, binders that mimic small drugs are difficult since the databases are more limited and often proprietary, so the training set is smaller. In addition, the differences between binders that activate or inhibit a target protein can be subtle.
- de novo synthesis of protein catalyst: Much of a protein structure is used to bring key groups into a stable configuration for catalysis. Synthetic chemists try to make small transition metal catalysts that mimic the function of proteins with a catalytic site that often contains a metal ion. This suggests that natural proteins might not be the most efficient mimics to produce novel protein catalysts. Also, proteins that differ in 3D structure can carry out similar reactions.
- Conformational flexibility in proteins: Unless we look at the dynamic structures of proteins, our minds can be trapped into creating just the most stable, low-energy protein structure. Yet flexibility and conformational changes are key to protein function and regulation. Programming conformational change into the algorithms for de novo synthesis is another complicated task.
- Creating proteins and protein complexes with functions other than catalysis: Many macromolecular assemblies (inflammasomes, proteasomes, regulated membrane pores, mobility proteins, etc) provide critical cellular functions. Creating new ones could offer novel ways to modulate cell function. One example would be to create nanoparticles that can deliver "cargo" (such as vaccines) into cells or potentially sequester and eliminate deleterious intracellular components (like misfolded and aggregated proteins).

4.7.14.6: Summary

In this chapter, we explore how the tremendous progress in computational biology has transformed our approach to predicting protein structures from their amino acid sequences—a challenge once known as the "protein folding problem." Traditionally, protein structures were determined through experimental methods such as NMR spectroscopy, X-ray crystallography, and cryo-EM, but these techniques are time-consuming and costly. With the advent of massive structural databases like the Protein Data Bank (PDB) and the explosion of sequence data, researchers have shifted toward computational methods to predict 3D structures from primary sequences.

Modern machine learning approaches—most notably AlphaFold and RoseTTAFold—have revolutionized the field. These methods leverage hidden evolutionary signals embedded within protein sequences to infer spatial contacts and predict highly accurate 3D models, often matching experimental structures with remarkable fidelity. Key metrics, such as RMSD and TM-scores, are used to evaluate the accuracy of these predictions by comparing them to known structures.

Beyond predicting the 3D structure of individual proteins, recent advances like AlphaFold3 extend these capabilities to protein complexes, offering insights into protein–protein interactions that underpin many cellular processes. Meanwhile, structure-based methods like FoldSeek utilize a “structural alphabet” (3Di) to compare and infer protein function from 3D structures, especially useful when sequence homology is limited.

The chapter further addresses the reverse protein folding problem—determining function from structure—and delves into the emerging field of de novo protein design. Techniques such as RFDiffusion and ProteinMPNN allow scientists to generate novel protein sequences with desired structural features and functions, paving the way for innovative applications in therapeutics, enzyme engineering, and synthetic biology.

Overall, this chapter highlights how computational protein modeling, powered by advanced neural network architectures (transformers and protein language models), is bridging the gap between sequence and structure. It emphasizes that while these computational methods greatly accelerate research, experimental validation remains essential to confirm predicted structures and functions, ensuring that these digital tools serve as powerful complements to traditional biochemical approaches.

This page titled 4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **4.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24)** by Henry Jakubowski and Patricia Flatt has no license indicated.

4.7.15: Chapter 4 Questions

Section 1 Questions:

Q1) Newman projection

Q2) Explain what conformation peptide bonds predominately fall into, cis or trans. Can you think of a situation in which the nonpredominate form could be found in a biological system?
A2) Most peptide bonds are in the trans conformation which helps to reduce steric hindrance R-groups of the amino acids but also decreases the strain on the peptide linkage. The cis form can be found in peptides, but at a low percentage. As the cis form causes strain on the peptide backbone. Enzymes can utilize this bond strain to overcome the activation energy for catalysis.

Q3) Ramachandran Plot example (define the regions)

Section 2 Questions:

Q1) Which amino acid(s) is(are) considered to be alpha helix breaker(s)? Explain why this occurs.
A1) Proline is consider to be the amino acids that "breaks" the secondary structure of an alpha helix. This occurs due to the amine ring in the proline R-group. This R-group is unable to participate in H-bonding needed for helix stabilization, and also is sterically locked in one conformation. Glycine is also an alpha helix breaker. Glycine has only a hydrogen as its R-group and therefore has no constrains for its Φ and Ψ angles. This makes glycine too flexible to hold the structure needed for am alpha helix.

Q2) Helical wheel question

Q3) Beta sheet/alpha helix counting from PDB file

Q4)

4.7.15: Chapter 4 Questions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.15: Chapter 4 Questions** has no license indicated.

4.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs

The exercises below use simple examples to demonstrate the power of machine learning/AI algorithms to predict protein structure and function and to design proteins with specific functions. You will then model the results using iCn3D. Many of the programs are quite new and likely will be changed when new and better software becomes available. This field is moving extremely fast. Some of the programs are commercial with limited free use sufficient for these exercises. Remember these exercises provide computational results. Experimental methods should be used to give these proposed structures additional support. The background for many of the methods used below is found in Chapter 4.14: Predicting Structure from Sequence and Sequence from Structure/Function

- Glossary of AlphaFold terms

4.8.1: Predicting structure from Sequence with AlphaFold

Exercise 4.8.1

1. Use AlphaFold3 from this link to determine a structure from this protein sequence:

MPGAISQLVSYGAQDVYLTGNPQITFFKAVYRRYTNFAMESIQQTFDGTTDFGKFPTVTISRNGDLAGPIWIEVNLPSLL
GYNITPTPAEGNTSNIAAISTVFKDDYNNYWWTYNPGTTPQYSNLIAAFSNVDYKYYANAVTSTYPPTALSNVVYSWP
YMITGNTGTRSTVAIPTANLRYVNGIGLALFNSIELELGGQRIDKHYSEWWDIWTELTETAEKIQGYNTMVGRYDPAV
YNAGWNISQAQGGTYYVPLKFCYNRNPGLYMPLVALSYHQMKLNFNINNYLNCVKCNYPVTALTSKNGANPLSITN
MKLYTDFVFLDAPERIRMSEIQHEYLVTQLQWQGSEPVTAPGDPNGSTNRKITLNFNHPVRELVFVYQAASNYDVDA
VTGNNIFDYEIPANPTATPPYAGGGEVFTEVKLIINGSDRFSGRPGAYFRLVQPYEHHVRVPSKSVYVYSFALEDADSRQ
PNGSANFTRYDSVQLQLTLNENLASGRVQIYAPNFNILRIAAGMGGLAFAN

   a. Select these in the input window and then paste in the above sequence.

   b. Select Continue and Preview Job
   c. Name the job: FOB_Unknown_Protein1
   d. Select Confirm and Submit Job
   e. When completed, check the job box and open the results
   f. Take a screen capture of the results window
   g. Select the download icon in the top menu bar to download a compressed zip file of the results
   h. Unzip the folder/file

Use iCn3D to open the first structure file: fold_fob_unknown_protein1_model_0.cif (it is a .cif file, not a pdb file) and render as follows:

- **File, Open File, mmCIF**
- **Analysis, Seq. & Annotations**
- In the Seq & Annotations window, choose the Details tab

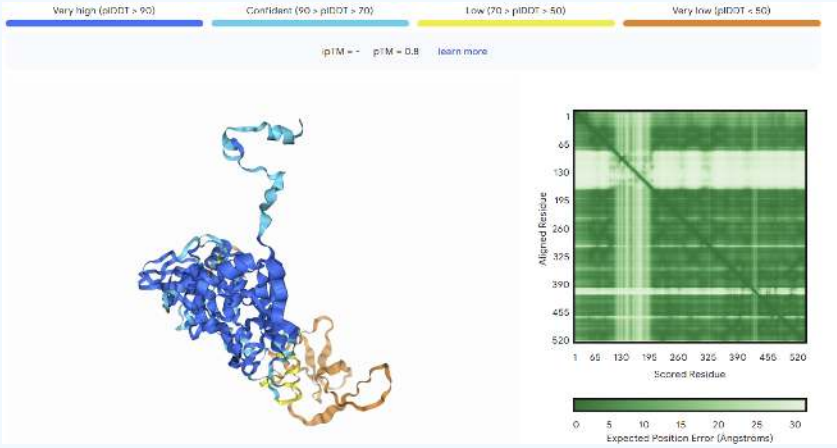
iCn3D shows 2 domains, the Capsid_N and Capsid_NCLDV domains (NCLDV stands for nucleocytoplasmic large DNA viruses).

- Click the blue Capsid_N domain name to highlight it
- **Color, Unicolor, Cyan, Cyan**
- **Select, Save Selection,** and name it Capsid_N domain
- Analysis, Label, Per Selection, then name it Capsid_N, and change the size to 10
- Click the blue Capsid_NCLDV domain name to highlight it
- **Color, Unicolor, Magenta, Magenta**
- **Select, Save Selection,** and name it Capsid_NCLDV domain

- Analysis, Label, Per Selection, then name it Capsid_NCLDV, and change the size to 10
- **Clear Selection**
- **Style, Background, Transparent**
- **File, Save File, iCn3D PNG image, Original Size and HTML**
- name it: fold_fob_unknown_protein1_model_0

Answer

f. Here is the screen capture



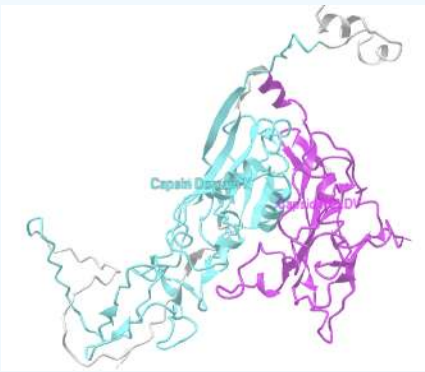
iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

- Open iCn3D
- **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Try this: https://www.ncbi.nlm.nih.gov/Structu...?origin=mt-web

https://www.ncbi.nlm.nih.gov/Structu...d/pdb/3GVU.png

Here is a static image:

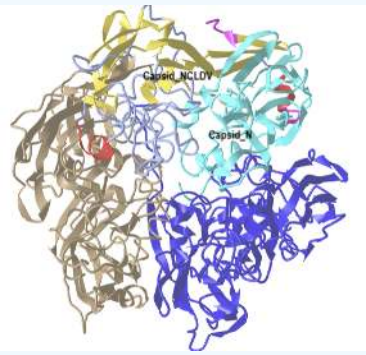


Domain information from iCn3D:

- Capsid_N: This is the N-terminal domain of the major capsid protein in several dsDNA viruses.(open details view...)

- Capsid_NCLDV: This family includes the major capsid protein of iridoviruses, chlorella virus, and Spodoptera ascovirus, which are all dsDNA viruses with no RNA stage. This is the most abundant structural protein and can account for up to 45% of virion protein. In Chlorella virus PBCV-1 the major capsid protein is a glycoprotein. The four families of large eukaryotic DNA viruses, Poxviridae, Asfarviridae, Iridoviridae, and Phycodnaviridae, are referred to collectively as nucleocytoplasmic large DNA viruses or NCLDV. The virions of different NCLDV have dramatically different structures. The major capsid proteins of iridoviruses and phycodnaviruses, both of which have icosahedral capsids surrounding an inner lipid membrane, showed a high level of sequence conservation. A more limited but statistically significant sequence similarity was observed between these proteins and the major capsid protein (p72) of ASFV, which also has an icosahedral capsid. It was surprising, however, to find that all of these proteins shared a conserved domain with the poxvirus protein D13L, which is an integral virion component thought to form a scaffold for the formation of viral crescents and immature virions. (open details view...)

The figure below shows an interactive iCn3D model of a capsid trimer from Paramecium bursaria Chlorella virus 1 (1M4X).



NCBI iCn3D Figure: Capsid trimer from Paramecium bursaria Chlorella virus 1 (1M4X). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7MCUVpKLytn9bA

The monomer in this trimeric capsid protein is a different capsid protein with a different sequence than the one you modeled. Two monomers are shown in brown and purple. The other monomer is colored to show the two Capsid domains (same as in the one you modeled). One is in cyan, and the other is colored by secondary structure. Each of these two domains contains a jellyroll domain with eight beta-strands that form two 4-stranded sheets. Thousands of these individual trimers form the capsid shell of the virus, as shown in the figure below.

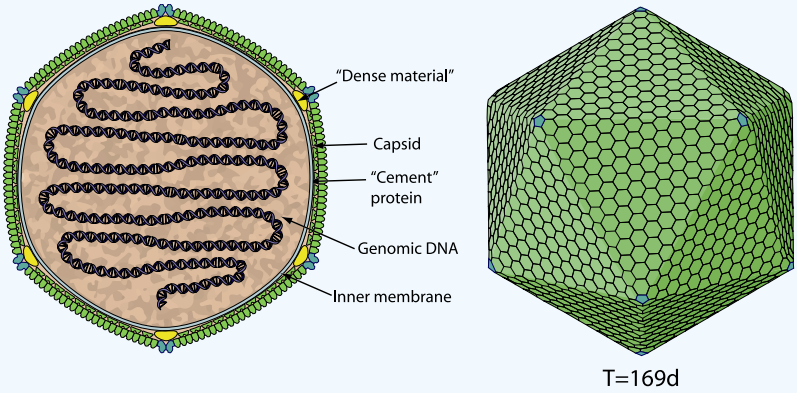


Figure: Structure of a Phycodnaviridae. https://viralzone.expasy.org/145

In this virus, there are 5040 copies of the major capsid protein (hexamers), 60 copies of the penton protein (pentamers) and 1800 minor capsid proteins of different types

In this example, you took linear information (the amino acid sequence) and turned it into 3D information, the tertiary structure of an unknown protein. Subsequent analyses within iCn3D show that this is a capsid protein from the dsDNA Paramecium bursaria

Chlorella virus 1. The protein you modeled (same sequence) is listed in Uniprot (O41104) and the PDB (8H2I) as an averaged PBCV-1 capsid**,** but no 3D structures of this protein are available in these databases.

4.8.2: Predicting Structure from Sequence with ESMFold

This program predicts structures from the sequence in the ESM Metagenomic Atlas. In contrast to genomics, which involves the DNA sequence analysis of a single organism, metagenomics involves DNA sequencing of a complex sample (such as soil, the gut microbiome, etc) with a large number of different organisms and viruses. The Metagenomic Atlas contains information on a vast number of proteins. Protein structures in this database are determined using another AI program, **ESMFold**. These structures can be viewed in iCn3D as shown in the figure below.

The Atlas contains predicted protein structures for over 700 million proteins. A million of these can be previewed on their Explore page. ESMFold uses statistical scoring metrics similar to AlphaFold. They include pTM, which is the predicted TM-score, giving the reliability of the global fold, and the mean predicted LDDT (pLDDT) score, which gives the reliability of the local structure. Over 225 million structures, including the sampling of those on the explore page, have high confidence (both pTM and pLDDT >0.7).

? Exercise 4.8.2

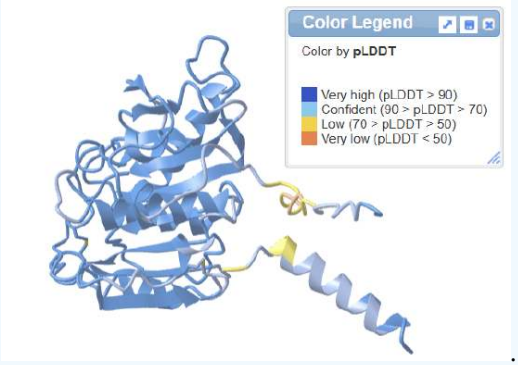
a. Go to iCn3D
b. File, Predict by Sequence, EMS Fold.
c. Input this sequence:
MIKNFFQNTEKSEFFKNSLVANFIIILTMVPQFLLVPIILNFWGEAQYSSYVVFLSVLNIAIQINGAMQNGYINR
LFQSGKISVAELSSMFLLNLGLLSFMLILLFVFYVSGYGAAFFSHITLIILSLPFLFALSSLKIIYQLGHKVTTPL
d. then select EMSFold
e. Render the structure in iCn3D as you wish.
f. File, Save File, iCn3D PNG image, Original Size and HTML

Answer

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Here is a static image of the calculated structure:



.

4.8.3: Predicting binding interactions and complex formation with AlphaFold

Exercise 4.8.3

Use AlphaFold3 from this link to determine the structure of a complex between a protein and a dsDNA from just their sequence. Then compare the computed structure with a very similar X-ray structure. This example comes from a Nature paper that used AlphaFold to predict the structure of the complex. Here is the reference:

Abramson, J., Adler, J., Dunger, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). https://doi.org/10.1038/s41586-024-07487-w

Protein: AMP-binding protein-catabolite gene activator from Rhizobium meliloti (strain 1021) (Ensifer meliloti) (Sinorhizobium meliloti), uniprot: Q92SD2. Here is the protein sequence:

MAEVIRSSAFWRSFPIFEEFDSETLCELSGIASYRKWSAGTVIFQRGDQGDYMIVVVSGRIKLSLFTPQGRELMLRQHE
AGALFGEMALLDGQPRSADATAVTAAEGYVIGKKDFLALITQRPKTAEAVIRFLCAQLRDTTDRLETIALYDLNARVAR
FFLATLRQIHGSEMPQSANLRLTLSQTDIASILGASRPKVNRAILSLEESGAIKRADGIICCNVGRLLSIADPEED

dsDNA sequence: Here are sequences of both DNA strands (5'-3'):

5'-CTAGGTAACATTACTCGCG-3' (19 mer) and 5'-GCGAGTAATGTTAC-3' (14 mer)

In AlphaFold 3, input the entry type, copies (number of molecules in the structure), and the sequences, as shown in the figure below.

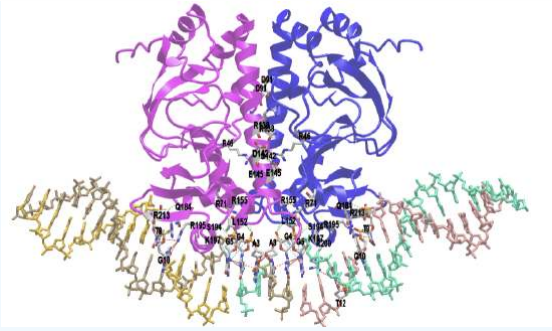
In the actual crystal structure, a protein dimer binds the dsDNA

Answer

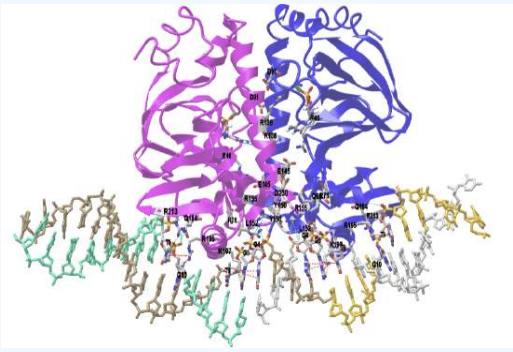
iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Here is a static image of the calculated structure:



The figure below shows an interactive iCn3D model of the actual X-ray crystal structure of the capsid trimer from Clr-cAMP-DNA complex (7PZB) for comparison. Note that cAMP could not be included in the AlphaFold 3 predicted structure of the complex using the version available.



NCBI iCn3D Figure: X-ray crystal structure of the capsid trimer from **Clr-cAMP-DNA complex (7PZB)**. (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YZJitoVd7Y2LAA

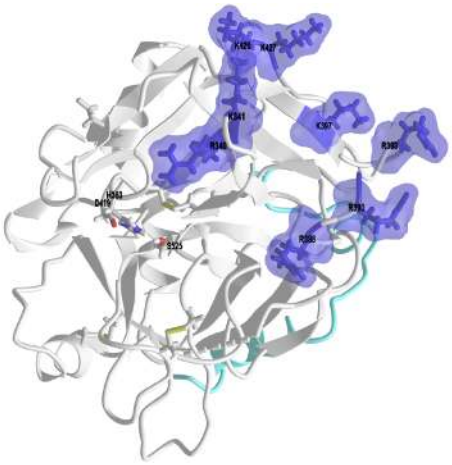
.

The program can be used to determine the structure of protein complexes as well as in this example from Chapter 4.14, the human sperm proteins and egg protein complex predicted by AlphaFold.

4.8.4: Designing target proteins de novo (ex., making novel binders) with RFDiffusion

The figure below shows an interactive iCn3D model of the X-ray crystal structure of human thrombin (3U69). Thrombin is a serine protease. It cleaves other proteins using its active site serine as a nucleophile. Other examples of serine proteases are the gut proteases chymotrypsin and trypsin, which are used in digestion. Thrombin cleaves a limited repertoire of proteins, mostly involved in blood clotting and its control. A main substrate is the protein fibrinogen, which, after cleavage of small peptides, forms fibrin that associates to form a fibrin clot. In addition to binding at the active site, specific substrates like fibrinogen interact with an anionic "exosite" on thrombin near the active site. This additional binding site limits the specificity of thrombin to specific substrates.

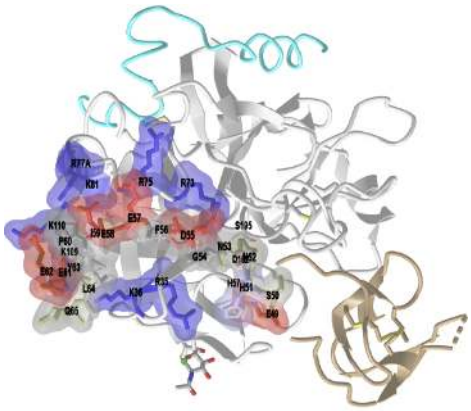
The figure below shows an interactive iCn3D model of human thrombin (3U69). Three active site residues (catalytic triad) involved in cleaving fibrinogen and other clotting proteins are shown in CPK-colored sticks. The blue side chains and surface areas represent key positively charged residues in "Exosite 1" of thrombin, involved in binding the anionic region of fibrinogen and a modulator of thrombin activity, thrombomodulin.



NCBI iCn3D Figure: X-ray crystal structure of human thrombin (3U69). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PkxukNXxPGuHa7

Organisms that need liquid blood for food have exploited Exosite 1 and created peptide inhibitors that bind to it and inhibit fibrinogen binding, for example. One such organism is a leech, which secretes a protein in its saliva called **hirudin**. It has a long C-terminal anionic tail that binds in the exosite groove and inhibits the binding of fibrinogen.

The figure below shows an interactive iCn3D model of the X-ray crystal structure of the hirudin-thrombin complex (4HTC). Key anionic residues in the C-terminal tail of hirudin are shown as red sticks and red surfaces interacting with the positive residues (blue sticks and surfaces) in Exosite 1 of thrombin.



NCBI iCn3D Figure: X-ray crystal structure of the hirudin-thrombin complex (4HTC). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...EyESJzsUV85Ex5.. To change the background color on the popup choose Style, Background, Transparent.

Synthetic inhibitors of thrombin that target exosite 1 of thrombin have been made. One contains a negatively charged DNA (an aptamer) covalently linked to an active site inhibitor, creating an "EXosite and ACTive site (EXACT) inhibitor".

? Exercise 4.8.4

In the next exercise, you will create a novel peptide inhibitor that targets exosite 1 of thrombin using RFDiffusion. This should be a simple example as the peptide is expected to have glutamates and aspartates, mimicking the hirudin C-terminal end.

You will run RFDiffusion from a commercial company called Neurosnap. First, set up an account (free, five models/month)

a. Go to RFDiffusion-v2
b. Separately, download this pdb file for human thrombin to your computer: 3U69

c. Open RFdiffusion-v2 and input the following data, which is shown in red. Leave the rest alone. H stands for the Heavy Chain in the PDB file.

Binder Input Chain	H
Binder Length Maximum	20
Binder Length Minimum	17
Binder ROG	false
Binding Pocket Residue End	1
Binding Pocket Residue Start	1
Fixed Residue End	20
Fixed Residue Start	10
Hotspots	H340,H341,H388,H390,H393,H397,H426,H427
Input Structure	input_structure.pdb

d. Run the program.

When completed, the output page has several tabs: Data and Visuals, Config, and Files. View the results in the Data and Visuals tab

Answer

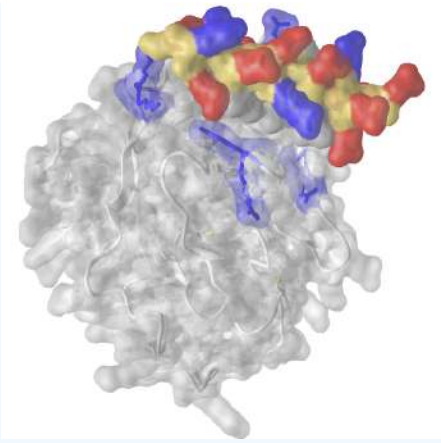
iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Note that the amino acid numbering of the hotspots changed in the output files from RFDiffusion, as shown in the table below.

Initial PDB numbering system for Heavy Chain (RU69)	RFDiffusion numbering for Heavy Chain Hot Spots
R340	20
K341	21
R388	68
R390	70
R393	73
K397	77
K426	106
K427	107

Here is a static image:



The light chain is not shown. The heavy chain H is shown in gray. The hot spots on thrombin (H340, H341, H388, H390, H393, H397, H426, and H427) are colored by charge (blue stick and surface). The binder is shown as a gold surface with the charged side chains shown in red (negative) and blue (positive). The sequence of RFDiffusion binder is EEEEKELELLREEIEKLEKE, which has 11 glutamates (E) but also 3 lysines (K), perhaps to maintain some charge balance.

Here is an animation from the output to show how RFDiffusion produces the final peptide at the hotspot residues.

Here are the stats for the top-ranked binder

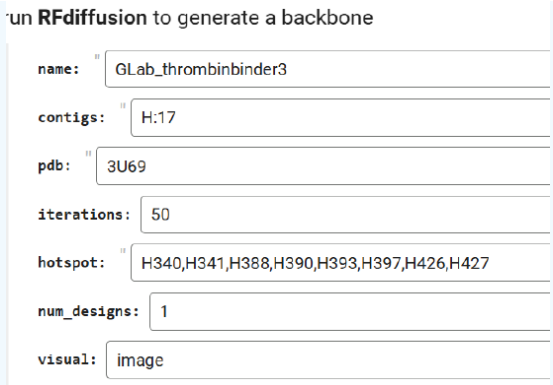
Rank	MPNN Score	RMSD	Mean pLDDT	Max PAE	pTM
1	1.14	19.69	92.98	23.48	0.91

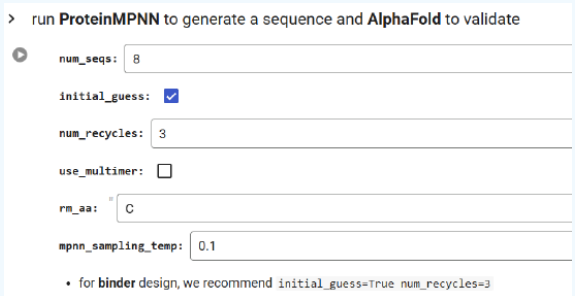
Watson, Joseph et al. “Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models” bioRxiv.Org. doi: https://doi.org/10.1101/2022.12.09.519842. Neurosnap Inc. - Computational Biology Platform for Research. Wilmington, DE, 2022. https://neurosnap.ai/.

? Exercise 4.8.5

Create a binder for human thrombin using Google Colab at this site (note: this link might be replaced in the future)

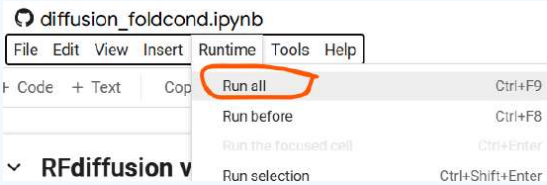
a. **Inputs**: Accept all defaults except for the sections below. For those, use the inputs shown. The goal is not to optimize the run but to show how easy it is to create a binder. The inputs allow you to create a 17 amino acid binder to the H (heavy) chain of human thrombin (3U69) at the positively charged Exosite 1 hot spots residues in the H chain, as described above.





b. **Run the program**. You can run sections individually or all at once. To run one section after the other, click the arrow for the first section as shown below, wait for it to complete, and then run the rest

The first section sets up the program and takes a few minutes to run. Alternatively, run all at once as shown in the image below.



The results should be downloaded to a .zip file in your computer's download folder.

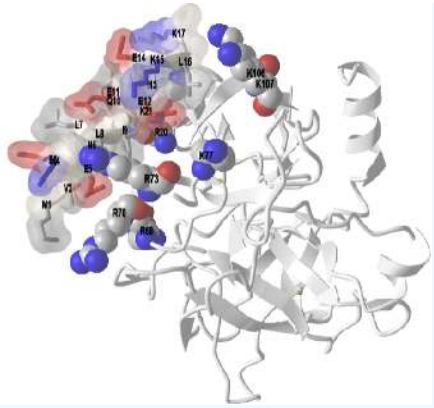
c. **Visualizing the results:** Unzip the folder and open the "best_design0.pdb file in iCn3D. Display and render your file in iCn3D.

Answer

iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Note that the number system for the H chain changed as the 1st run. Here is a static image:



The gray chain is the heavy chain of thrombin. The hot spots on this chain are shown in spacefill, CPK colors, and labeled. The binder is shown as a transparent surface with side chains colored by charge (red = positive, blue = negative).

4.8.5: Predicting Protein Function through Structure with FoldSeek

FoldSeek can be run directly on a local PDB file or one retrieved from the PDB through this link. However, it can also be run through iCn3D, adding many more possibilities for rendering and analyses.

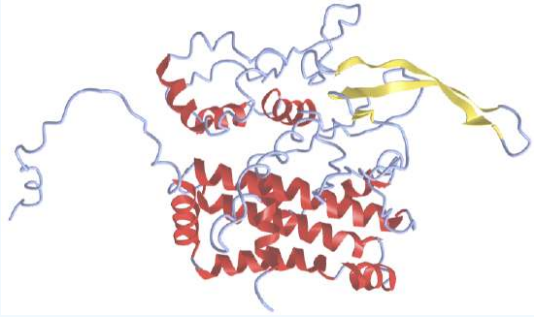
Use iCn3D to run FoldSeek

? Exercise 4.8.6

We will explore the structural similarities of a newly predicted viral protein from the Wuhan insect virus from an uncertain family to the structure of all proteins in the databases using FoldSeek using iCn3D.

First, download this file: hypothetical_protein__YP_009329883__Wuhan_insect_virus_23__1923727.pdb

a. Open iCn3D
b. **File, Open File, PDB appendable**, and load the downloaded file
c. **Color, Secondary, Sheets in Yellow**. Note that the protein is predominantly alpha-helical with some beta-sheets. Here is a static view other viral protein.



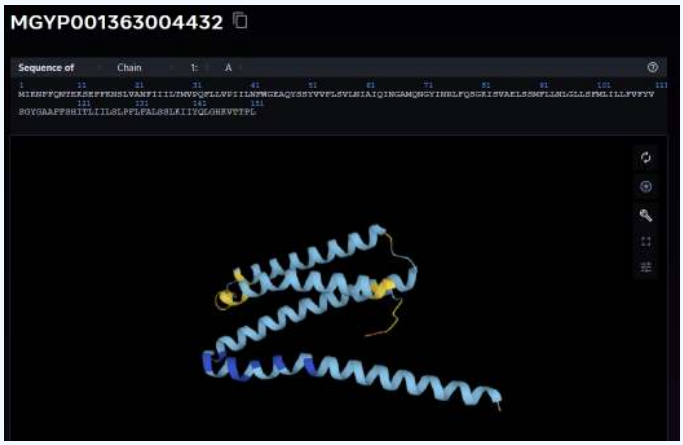
Now let's use FoldSeek within iCn3D to find similar structures (not sequences) with the databases (PDB and AlphaFold).

d. **File, Search Similar, FoldSeek (PDB and AlphaFold)**
e. Click Submit

A new tab with FoldSeek appears with the best statistical structural fits. The statistics shown in this table include the probability (0-1) that the structural match of the query (the protein studied) to the target (what it might resemble) is not due to chance. The other statistic is the expected or E-value, which gives the probability (0-1) that the structures (or sequences) match just by chance. The closer the E-value is to zero, the more significant the match. (See this PDB page describing E-values for sequence alignments.)

f. Which hit is best? How good is the quality of the fit?

g. If you click the blue hyperlink for the best hit, you will get the structure from the ESM database, and you can download it. The database protein structure is shown below.

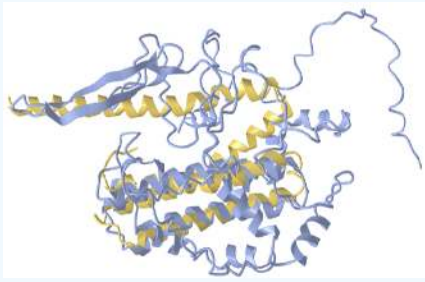


Instead, click the staggered three lines (☰) to the right to get the alignment with the hypothetical protein.

This pops up:

The query (blue, the protein studied) and the target (orange, what it might resemble)

h. Click the PDB link to download a PDB file of the two aligned protein segments. Model them in iCn3D to replicate the structure above.



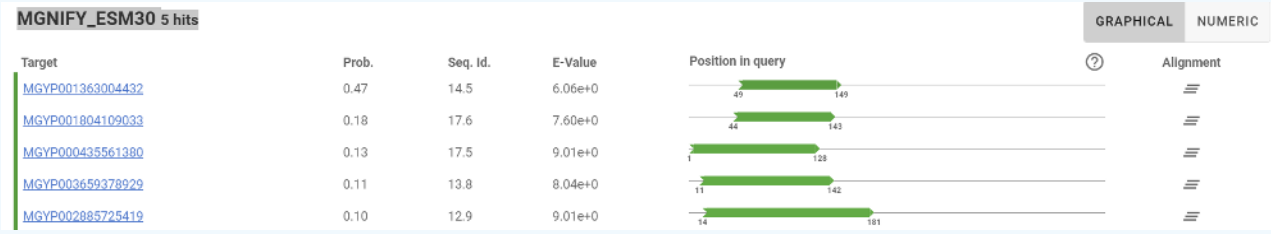
Query (hypothetical_protein__YP_009329883__Wuhan_insect_virus_23__1923727) is blue, Target (MGYP001363004432) is orange

This hypothetical viral protein structure is very dissimilar to any structure in the databases except for the helical bundles similar to protein MGYP001363004432 in the **ESM Metagenomic Atlas database.**

Answer

f. You will see that for this particular structure, there are just five "hits" in only one database, mgnify_esm30, from the ESM Metagenomic Atlas Database described above.

Here are the five hits.



MGNIFY_ESM30 5 hits

GRAPHICAL NUMERIC

Target	Prob.	Seq. Id.	E-Value	Position in query	Alignment
MGYP001363004432	0.47	14.5	6.06e+0	49 149	
MGYP001804109033	0.18	17.6	7.60e+0	44 143	
MGYP000435561380	0.13	17.5	9.01e+0	1 128	
MGYP003659378929	0.11	13.8	8.04e+0	11 142	
MGYP002885725419	0.10	12.9	9.01e+0	14 181	

None of the structural alignments are good statistically.

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

This hypothetical viral protein structure is very dissimilar to any structure in the databases except for the helical bundles similar to protein MGYP001363004432 in the **ESM Metagenomic Atlas database.**

Query (hypothetical_protein__YP_009329883__Wuhan_insect_virus_23__1923727) is blue, Target (MGYP001363004432) is orange

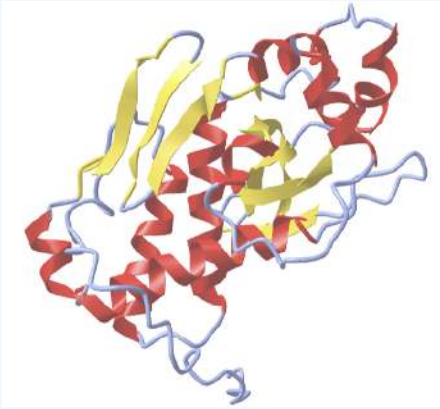
? Exercise 4.8.7

Now, let's try a new viral protein structure whose function we can likely infer from the FoldSeek structural overlay.

Download this protein structure (whose function has been surmised through programs like FoldSeek: matrix_protein__YP_001531158__Marburg_marburgvirus__11269.pdb.

a. Open iCn3D
b. **File, Open File, PDB appendable**, and load the downloaded file
c. **Color, Secondary, Sheets in Yellow**. Note that the protein is predominantly alpha-helical with some beta-sheets. Here is a static view other viral protein.

Here is a static image:



e. Now repeat the procedure above and use FoldSeek within iCn3D (as described in Exercise 6 above) to find proteins of similar structure and presumably function.

f. Model one of the top hits in iCn3D

Answer

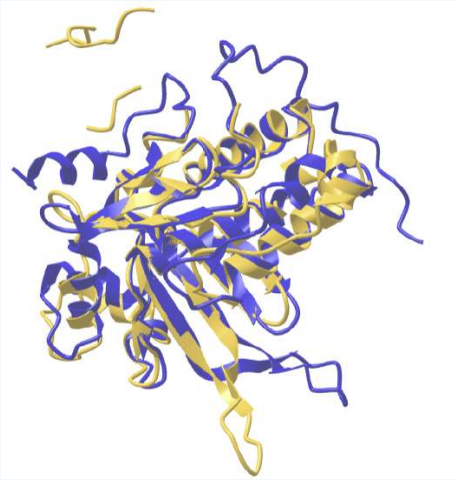
The optimal target is VP24 of the Marburg virus (Uniprot ID: **Membrane-associated protein VP24 - P35256**). Both query and target proteins are from the Marburg virus, in the Filoviridae family, including the Ebola virus. These are deadly viruses that cause hemorrhagic fevers and have high death rates. Here is one example of an almost-perfect hit.

Target	Description	Scientific Name	Prob.	Seq. Id.	E-Value	Position in query	Alignment
4or8-assembly1_B	Crystal structure of Marburg virus VP24	Marburg virus - Musoke, Kenya, 1980	1.00	94.1	9.90e-32		

The query (blue, the protein studied) and the target (orange, what it might resemble)

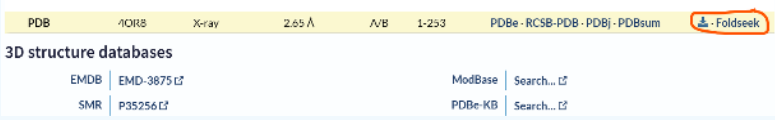
iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder. A static image is shown below.



Query (matrix_protein__YP_001531158__Marburg_marburgvirus__11269) is blue, Target (VP24 of the Marburg virus) is orange

From the Uniprot page for the VP24 protein, FoldSeek can be run to determine which proteins in the database are similar. In this case, VP24 is the query and FoldSeek finds targets of similar structure.



The best hits are in the PDB database:

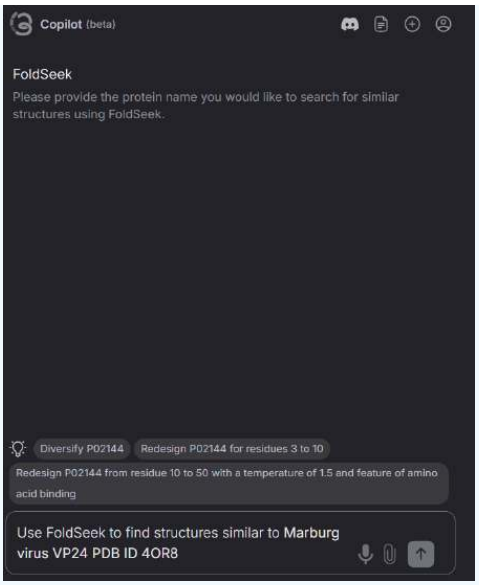
Target	Description	Scientific Name	Prob.	Seq. Id.	E-Value
4or8-assembly1_A	Crystal structure of Marburg virus VP24	Marburg virus - Musoke, Kenya, 1980	1.00	100	1.03e-46
4or8-assembly1_B	Crystal structure of Marburg virus VP24	Marburg virus - Musoke, Kenya, 1980	1.00	94.7	4.12e-37
6ehm-assembly1_C	Model of the Ebola virus nucleocapsid subunit from recombinant virus-like particles	Ebola virus - Mayinga, Zaire, 1976	1.00	37	7.36e-19
4u2x-assembly3_C	Ebola virus VP24 in complex with Karyopherin alpha 5 C-terminus	Ebola virus - Mayinga, Zaire, 1976	1.00	36.5	6.94e-19
3vne-assembly1_A	Structure of the ebolavirus protein VP24 from Sudan	Sudan ebolavirus	1.00	35.5	3.25e-18
4d9o-assembly1_A	Structure of ebolavirus protein VP24 from Reston	Reston ebolavirus - Reston	1.00	36	1.55e-16
3vnf-assembly1_A	Structure of the ebolavirus protein VP24 from Sudan	Sudan ebolavirus	1.00	34.8	1.75e-16
4d9o-assembly1_B	Structure of ebolavirus protein VP24 from Reston	Reston ebolavirus - Reston	1.00	35	3.84e-15

Running FoldSeek using 310 Copilot

Exercise 4.8.8

Now run FoldSeek using 310 Copilot. This program is part of a suite of commercial programs from Open AI. These programs allow users to input questions and data as sentences (much like Chatbots like ChatGPT, Claude, Gemini, etc.) to address complex questions in biology and biochemistry.

a. Click try it now. Input the query shown in the bottom of the figure below. IN this case, you are asking to find structures similar to the VP24 of the Marburg using the PDB ID = 4or8



Answer

Here are the results:

name	subject id	evalue	tmscore
4or8_1	4or8 B	1.0	1.0
4or8_2	4or8 A	0.922	0.9277
4or8_3	6ehm C	0.821	0.8276

As of 11/21/24, you can't upload a local pdb file (as we did with iCn3D above) and run FoldSeek within Copilot.

4.8.6: Exploring protein complexes with FoldSeek Multimer

In Exercise 3 above, you used AlphaFold to predict the structure of protein complexes (in that particular example, a protein:DNA complex) from sequences. By analogy, FoldSeek Multimer can find the 3D structures of target complexes from the 3D structure of a known (query) complex. In short, AlphaFold can do large-scale sequence-to-structure comparisons, while FoldSeek Multimer can do large-scale structure-to-structure comparisons. Let's try an example with a simple complex, the hepatitis A virus C3 proteinase (1HAV), a homodimer of an A and B chain.

? Exercise 4.8.9

Go to the FoldSeek Multimer server

a. Load Accession 1hav and run the program.
b. Copy/Snip the top 4 results. Compare the 1st and 4th results
c. Click the alignment icons = for each and take a screen snips. =

Answer

b. Here is a snip of the top results.

Complex		Chain						
qTM	tTM	Chain paring	Scientific Name	Prob.	Seq. Id.	E-Value	Position in query	Alignment
0.34	0.13	B → ProtVar_P83110_Q96RQ3_A	Homo sapiens	1.00	10.2	2.13e-7		
0.34	0.19	B → ProtVar_O43464_O99988_A	Homo sapiens	1.00	11.4	1.31e-5		
0.34	0.23	A → ProtVar_P61009_P83110_B	Homo sapiens	1.00	9	5.00e-6		
0.34	0.16	A → LevyLab_P83110_V1_2_relaxed...	Homo sapiens	1.00	8.2	9.99e-6		
		B → LevyLab_P83110_V1_2_relaxed...	Homo sapiens	1.00	9.5	6.91e-7		

The top results show high-quality alignments of one chain with a human protein, but that human protein is not part of a similar dimer. Result 4 shows good alignment of both chains of the query with a human dimer.

Note the label for the top hit: B ➔ ProtVar_P83110_Q96RQ3_A. The P83110 is the Uniprot number for human serine protease HTRA3.

c. Hit 1 - alignment for just 1 chain of the query and target

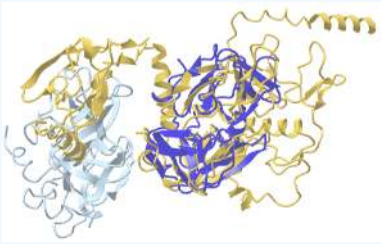
Query (blue, the protein studied) and the target (orange, what it might resemble)

The icons underneath the blue/orange models are explained below.

- PDB: select to download the combined file and model in iCn3D
- blue: toggle between the entire query and the aligned structure
- orange: toggle between the entire target and the aligned structure. Note that in this case, the target (human protein) is much bigger than the query (viral) protein so the alignment is only between part of the human protein.

iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder. A static image is shown below.

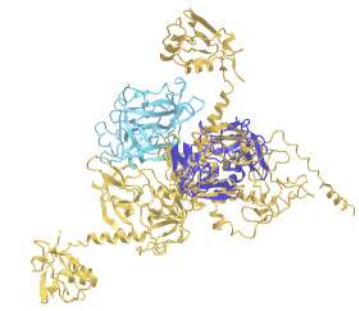


Query (blue, the protein studied) and the target (orange, what it might resemble)

Hit 4: Alignment of two chains (A and B) for the query and target

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder. A static image is shown below.

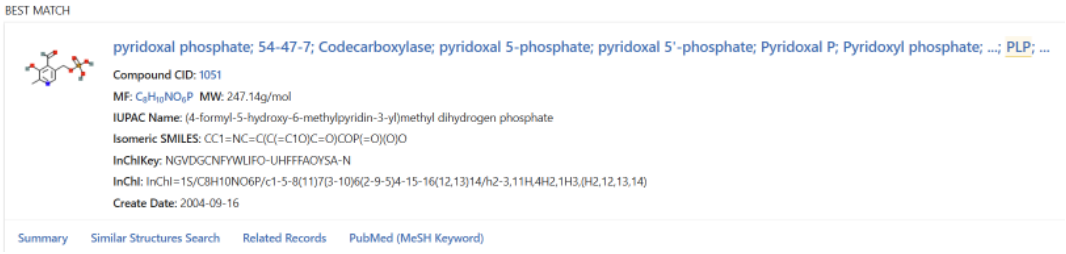


4.8.7: Docking of Small Molecules (Ligands)

Predicting binding interactions between small molecules and proteins is at the heart of the pharmaceutical industry. Previously, docking software for small molecule ligands and proteins was proprietary, costly, and difficult to run. Now, docking can be done online. In the next exercises, you will use three programs: 310 Copilot DiffDock (from a commercial company), DiffDock through Neuroapp, and SwissDock (Molecular Modeling Group, University of Lausanne, and the SIB Swiss Institute of Bioinformatics).

The inputs for a docking protein include the protein (typically the PDB ID) and the ligand, which can be represented as the actual structure, but more often a code in the SMILES or InChi formats. For these exercises, you will dock the small molecule pyridoxal phosphate to a low molecular weight protein tyrosine phosphatase, a protein that cleaves phosphorylated tyrosines in proteins.

The input representation for PLP can be obtained through PubChem shown below.



Docking using 310 Copilot

? Exercise 4.8.10

Docking using Copilot - DiffDock

a. Go to PubChem and get the SMILES representation for PLP.
b. Open 310 Copilot
c. Input this text: Run a docking experiment with the protein PDB ID = 5JNR and the small molecule SMILES=CC1=NC=C(C(=C1O)C=O)COP(=O)(O)O
d. When the docked structure is shown, select Download PDB
e. Save as 310CoPilotDock5JNR_PLP.pdb

Now model the results in iCn3D as follows:

f. Open iCn3D

g. **File, Open File, PDB appendable**, and load 310CoPilotDock5JNR_PLP.pdb

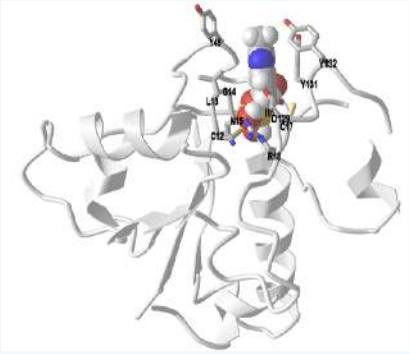
h. **Color, Secondary, Sheets in Yellow**. Note that the protein is predominantly alpha-helical with some beta-sheets. Here is a static view of the viral protein.

Answer

iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Here is a static view.

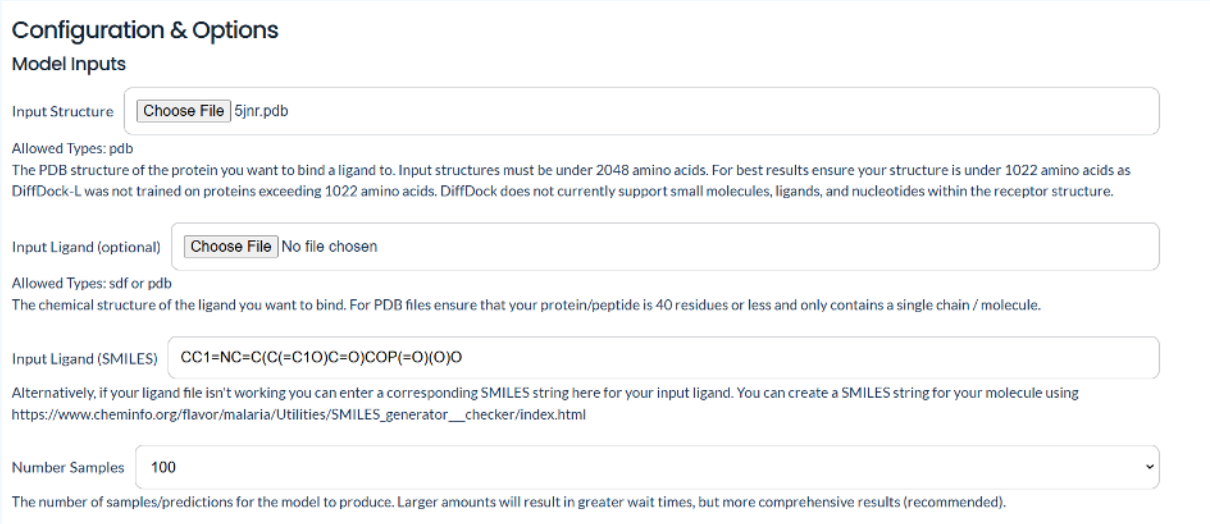


Docking using DiffDock in Neurosnap

? Exercise 4.8.11

Now, use DiffDock through Neurosnap to dock the ligand and protein.

a. Complete the menu as shown below.



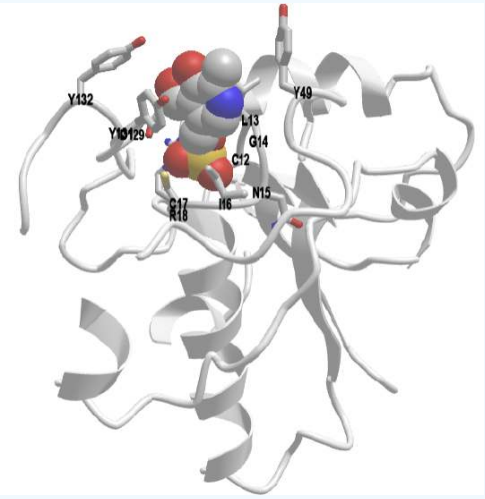
The output folder contains a separate PDB file and 100 SDF files for the various ligand "poses." Display them in iCn3D.

a. Open iCn3D

b. **File, Open File, SDF**, and choose the top-ranked SDF file (rank1_confidence-0.09.sdf)
c. File, Open File, PDB Appendable, and choose the proteins_no_ligands pdb file.
d. Render it as you see fit

Answer

Here is a screen snap of the docked structure.



iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Docking using SwissDock

Exercise 4.8.12

Now use SwissDock to dock PLP to a different low molecular weight protein tyrosine phosphate, PDB ID = 1xww. This has a SO_4^{2-} in the active site, which must be removed before the docking. The program uses two different methods, AutoDock Vina and Cavity Prioritization. Use the second one, which seems to give better results for this protein.

a. Input the values shown in the box below and run the docking experiment. When you input the target PDB file (1xww), choose the prompts shown below.

Query

Ligand	CC1=NC=C(C(=C1O)C=O)COP(=O)(O)O
Target	1xww_modified.pdb
Method	Attracting Cavities 2.0
Date	November 23, 2024, 1:57 pm UTC

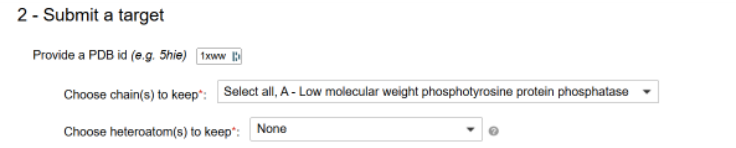
Parameters:

Box center:	13 - 0 - 18	Sampling exhaustivity:	medium	Number of RIC:	4
Box size:	20 - 20 - 20	Cavity prioritization:	buried		

If you publish these results, please, cite the following papers:

Bugnon M, Röhrig UF, Goullieux M, Perez MAS, Daina A, Michielin O, Zoete V. SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic Acids Res.* **2024**

Röhrig UF, Goullieux M, Bugnon M, Zoete V. Attracting Cavities 2.0: improving the flexibility and robustness for small-molecule docking. *J. Chem. Inf. Model.*, **2023**



b. Download the result as a zip file.

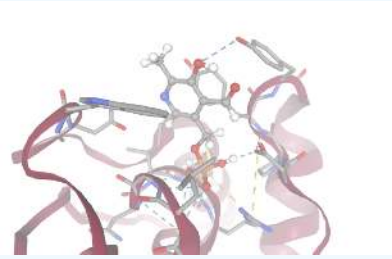
View and render the docked ligand in iCn3D. First, you must modify one of the files to view the docking results in iCn3D.

c. Choose Export Results as Zip file
d. Open the Zip folder and extract all.
e. Open the result.dock4 file in a simple text editor. Save just the first PLP coordinates by finding the 1st line with TER (terminate). Delete everything after that. Save the file as result0dock4.pdb file into the File subfolder you just extract
f. Open iCn3D
g. **File, Open File, PDB appendable**, and choose both the file you just made (results.dock4.pdb) and the receptor.pdb load it.
h. Render the image as you wish.

Answer

Here is the link to the actual SwissDock docking results page.

Here is a snip of the best-docked structure.



Top Pose:

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

This page titled 4.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **5.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

5: Protein Purification

5.1: Protein Purification
5.2: Cell Disruption
5.3: Fractionation
5.4: Electrophoresis
5.5: Electrophoresis
5.6: Gel Exclusion Chromatography
5.7: Ion Exchange Chromatography
5.8: Affinity Chromatography
5.9: High Performance Liquid Chromatography (HPLC)
5.10: Histidine Tagging
5.11: Quantification of Protein Concentration
5.12: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)

5: Protein Purification is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.1: Protein Purification

Learning Goals (ChatGPT o1, 1/25/25)

- **Explain the Rationale for Protein Purification:**
  - Describe why isolating a single protein from complex cellular mixtures is essential for structural and functional studies.
- **Differentiate Between Preparative and Analytical Purification:**
  - Distinguish the objectives, techniques, and quality control measures used in preparative versus analytical purification.
- **Outline Sample Processing Techniques:**
  - Identify various cell lysis methods (e.g., freeze/thaw, sonication, detergent treatment) and discuss how centrifugation separates soluble proteins from cellular debris.
- **Understand Precipitation and Dialysis Methods:**
  - Explain the role of ammonium sulfate precipitation and dialysis in concentrating proteins and removing contaminants.
- **Apply Chromatography Principles:**
  - Compare and contrast different chromatographic techniques (size exclusion, ion exchange, hydrophobic interaction, and affinity chromatography) based on protein properties such as size, charge, hydrophobicity, and binding affinity.
- **Interpret Chromatographic Data:**
  - Analyze elution profiles, calculate distribution coefficients, and evaluate parameters like void volume and stationary phase volume.
- **Quantitatively Evaluate Purification Efficiency:**
  - Calculate and interpret metrics such as total protein, total activity, specific activity, yield, and fold-purification to assess the success of a purification scheme.
- **Describe the Theory of Electrophoresis:**
  - Explain how charged particles move under an electric field, and relate factors like electrophoretic mobility to charge, size, and shape (including concepts like the Stokes radius).
- **Distinguish Native and Denaturing Electrophoresis:**
  - Compare techniques such as native PAGE and SDS-PAGE, and understand how SDS alters protein charge and conformation to facilitate molecular weight-based separation.
- **Integrate Isoelectric Focusing and Two-Dimensional Electrophoresis:**
  - Understand how pH gradients in isoelectric focusing separate proteins by their isoelectric points and how 2-DE combines IEF with SDS-PAGE for high-resolution analysis.
- **Interpret Gel Staining and Detection Methods:**
  - Evaluate different protein detection methods (Coomassie Blue, silver staining, autoradiography, Western blotting) and their applications in visualizing protein bands.
- **Relate Biochemical Principles to Instrumentation:**
  - Explain how HPLC, FPLC, and mass spectrometry (e.g., MALDI-TOF MS) integrate with protein purification protocols to provide structural and functional insights.
- **Understand the Role of Protein Tagging and Peptide Synthesis:**
  - Describe how molecular tags (e.g., His, FLAG, GFP) and solid-phase peptide synthesis facilitate purification and identification of proteins.
- **Critically Analyze Experimental Design:**

- Identify potential sources of error in purification protocols (e.g., proteolysis, non-specific binding) and propose strategies to optimize yield and purity.

- **Apply Theoretical Concepts to Practical Scenarios:**
  - Use mathematical relationships (such as $U = Q/(6\pi\eta Rs)$) to predict and analyze protein migration behavior in electrophoretic systems.

These goals aim to provide a comprehensive understanding that links the theory and practice of protein purification and electrophoresis, preparing students for advanced research and practical applications in biochemistry.

5.1.1: Introduction

A protein or other biological macromolecule must be purified before being rigorously studied structurally and functionally. Purifying a single protein from a mixture of as many as 10,000 other cellular or tissue proteins, each containing the same 20 naturally occurring amino acids linked in different orders, is difficult. Proteins differ in size (how many amino acids), charge (how many positively and negatively charged amino acids), sequence, and presence of specific binding sites on the proteins. Any technique used to purify protein must be based on these inherent differences. Once the protein is purified, it must be analyzed by a spectral or electrophoretic technique.

Protein purification is a series of processes intended to isolate and purify a single protein or complex from cells, tissues, or whole organisms. Protein purification is vital for characterizing the function, structure, and interactions of the protein of interest. Separation steps usually exploit differences in protein size, physical-chemical properties, binding affinity, and biological activity.

Protein purification is either *preparative* or *analytical*. **Preparative purifications** aim to produce a relatively large quantity of purified proteins for subsequent use. Examples include the preparation of commercial products such as enzymes (e.g., lactase), nutritional proteins (e.g., soy protein isolate), and certain biopharmaceuticals (e.g., insulin). Many steps and quality control are required to remove other host proteins and biomolecules that could threaten the patient's health. **Analytical purification** yields a relatively small amount of protein for various research or analytical purposes, including identification, structural characterization, and studies of protein structure, post-translational modifications, and function.

The choice of a starting material is key to the design of a purification process. In plants or animals, a particular protein is usually not distributed uniformly throughout the organism; different organs or tissues have higher or lower protein concentrations. Using tissues or organs with the highest concentration reduces the volume required to produce a given amount of purified protein. A protein of low abundance or high commercial value is often made using recombinant DNA technology. These techniques will be discussed in greater detail in Chapter 5.

5.1.2: Sample Processing

If the organism does not secrete the protein of interest into the surrounding solution, the first step of each purification process disrupts the cells that contain it. Depending on how fragile the protein is, one of several techniques could be used, including repeated freezing and thawing, sonication, homogenization by high pressure (French press), homogenization by grinding (bead mill), and permeabilization by detergents (e.g., Triton X-100) and/or enzymes (e.g., lysozyme). Finally, cell debris can be removed by centrifugation, leaving the proteins and other soluble compounds in the supernatant.

Proteases released during cell lysis cleave the proteins in the solution. As the protein of interest may be sensitive to proteolysis, it is important to proceed quickly and conduct many steps at low temperatures to reduce unwanted proteolysis. Alternatively, one or more protease inhibitors can be added to the lysis buffer immediately before cell disruption. Sometimes, it is also necessary to add DNase to reduce the viscosity of the cell lysate caused by a high DNA content.

5.1.3: Centrifugation

Centrifugation is a process that uses centrifugal force to separate mixtures of particles of varying masses or densities suspended in a liquid. When a vessel (typically a tube or bottle) containing a mixture of proteins or other particulate matter, such as bacterial cells, is rotated at high speeds, the inertia of each particle yields a force in the direction of the particle's velocity proportional to its mass. The tendency of a given particle to move through the liquid because of this force is offset by the resistance the fluid exerts on the particle. The net effect of "spinning" the sample in a centrifuge is that massive, small, and dense particles move outward faster

than less massive particles or particles with more "drag" in the liquid. When suspensions of particles are "spun" in a centrifuge, a "pellet" may form at the bottom of the vessel, enriched for the most massive particles with low drag in the liquid.

Non-compacted particles remain mostly in the liquid "supernatant" and can be removed from the vessel, separating the supernatant from the pellet. The angular acceleration applied to the sample determines the centrifugation rate, typically measured in units of *g*, the gravitational constant. If samples are centrifuged long enough, the particles in the vessels will reach equilibrium, in which they accumulate at a point in the vessel where their buoyant density is balanced by centrifugal force. Such an "equilibrium" centrifugation can allow extensive particle purification.

In ***sucrose gradient centrifugation***, a linear concentration gradient of sugar (typically sucrose, glycerol, or a silica-based density gradient medium, such as Percoll) is generated in a tube, with the highest concentration at the bottom and the lowest at the top. Percoll is a trademark owned by GE Healthcare companies. A protein sample is then layered on the gradient and spun at high speeds in an ultracentrifuge. This causes heavy macromolecules to migrate toward the bottom of the tube faster than lighter material. During centrifugation, in the absence of sucrose, as particles move farther from the center of rotation, they experience greater centrifugal force (the farther they move, the faster they move). However, the useful separation range within the vessel is restricted to a small observable window. A properly designed sucrose gradient will counteract the increasing centrifugal force so that the particles move proportionately to their time in the centrifugal field. After separating the protein/particles, the gradient is fractionated and collected. These are described in Figure 5.1.1.

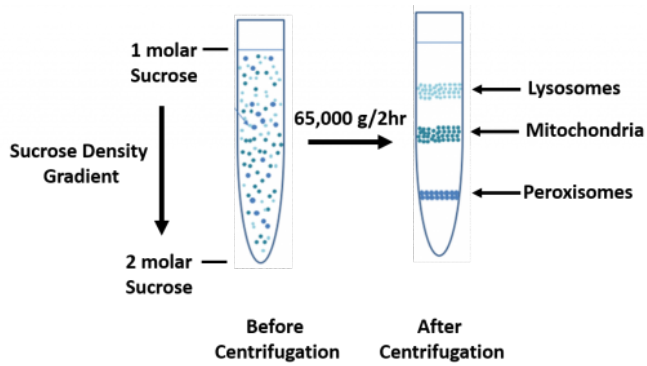


Figure 5.1.1: Sucrose Density Gradient. *Image derived from Michel Awkal*

5.1.4: Precipitation and Differential Solubilization

In bulk protein purification, precipitation is a common first step to isolate proteins using a salt such as ammonium sulfate $(NH_4)_2SO_4$. Ammonium sulfate is often used because it is highly soluble in water, is relatively free of temperature effects, and is typically not harmful to most proteins. Proteins are precipitated by $(NH_4)_2SO_4$ in their native state, which is important if you need the protein for structure/function studies. Furthermore, dialysis can remove ammonium sulfate as described in Figure 5.1.2.

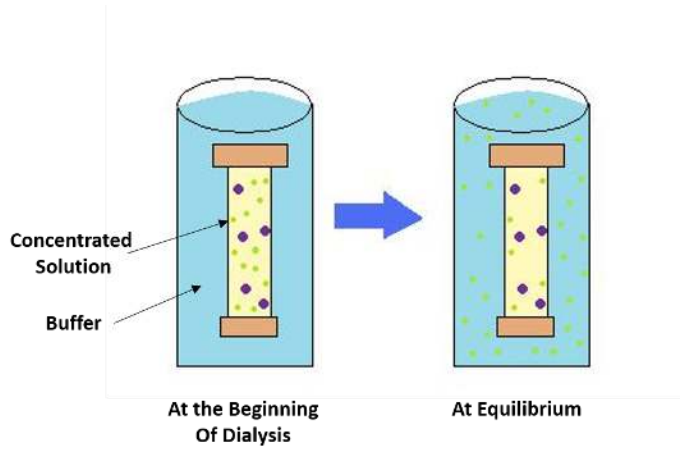


Figure 5.1.2: Dialysis of a macromolecule (blue dots). *Image adapted from Gjk003*

The process of dialysis separates dissolved molecules by their size. The biological sample is placed inside a closed membrane, where the protein of interest is too large to pass through the membrane's pores, but smaller ions can easily pass through. As the solution reaches equilibrium, the ions become evenly distributed throughout the solution while the protein remains concentrated in the membrane. This reduces the suspension's overall salt concentration.

The mechanism underlying salt precipitation is quite complicated. High concentrations of sodium chloride don't precipitate protein. Other salts, such as guanidinium chloride, unfold proteins without causing precipitation. Salt ions interact with the protein and solvent water in somewhat complicated ways (which we will explore later). We will be satisfied with the empirical observation that ammonium sulfate is the salt of choice to precipitate and concentrate proteins from a solution. One advantage of $(NH_4)_2SO_4$ protein precipitation from solution is that it can be performed inexpensively at a large scale, making it useful early in the purification of proteins. Different proteins precipitate at different $(NH_4)_2SO_4$ concentrations, so differential precipitation is often used. $(NH_4)_2SO_4$ concentrations are increased stepwise until the protein of interest is precipitated.

Some proteins are not soluble in water. These include transmembrane proteins that span cell membranes and large fibrous proteins. Membrane proteins can be solubilized by adding detergents such as sodium dodecyl sulfate (SDS), which unfolds the proteins, or octylglucoside or Triton X-100, which maintains protein structure.

5.1.5: Chromatography

Chromatography is used in almost all protein purification methods and is the key to separating a given protein from the 1000s of different proteins in cells and tissues. The separation of proteins on a chromatography column depends on the type of column and the chemical/physical properties of the molecule. There are four main types of chromatography used to separate proteins:

- **size exclusion** chromatography, in which proteins can be separated according to their size/shape or molecular weight
- **ion exchange** chromatography, in which proteins are separated by their charge/isoelectric point;
- **hydrophobic interaction** chromatography (similar to reverse-phase columns for purifying organic molecules), in which they are separated based on their relative hydrophobicity
- **affinity** chromatography, in which proteins are separated based on binding to a ligand covalently attached to a column bead.

The purification protocol generally contains one or more chromatographic steps for preparative protein purification. The basic procedure in chromatography is to flow the solution containing the protein through a column packed with a chromatography resin selected to separate proteins based on a specific property of the protein. Different proteins interact differently with the column material. They can thus be separated by the time required to pass the column or the conditions required to elute the protein from the column. Usually, proteins are detected as they elute from the column by measuring absorbance at 280 nm, where aromatic amino acids absorb.

5.1.5.1: Size Exclusion Chromatography (also known as Gel Filtration Chromatography)

This method separates proteins based on size and shape. The chromatography beads have tiny openings/pores into which proteins smaller than the pore diameter can enter. Large proteins that can't enter the pores flow around the beads and elute faster than small ones that enter the pores. They diffuse out of the pores and enter the rest of the moving solvent before getting "trapped" again for a short time in more pores. Eventually, they move through the column and elute at much higher volumes than larger proteins, which can't enter the pores. Thus, proteins will be separated based on size, as illustrated in Figure 5.1.3. The eluate is collected in sequential test tubes (or fractions). The figure below shows the pores as channels that go through the bead. The openings in resin beads should be considered tiny crevices and indentations, not channels.

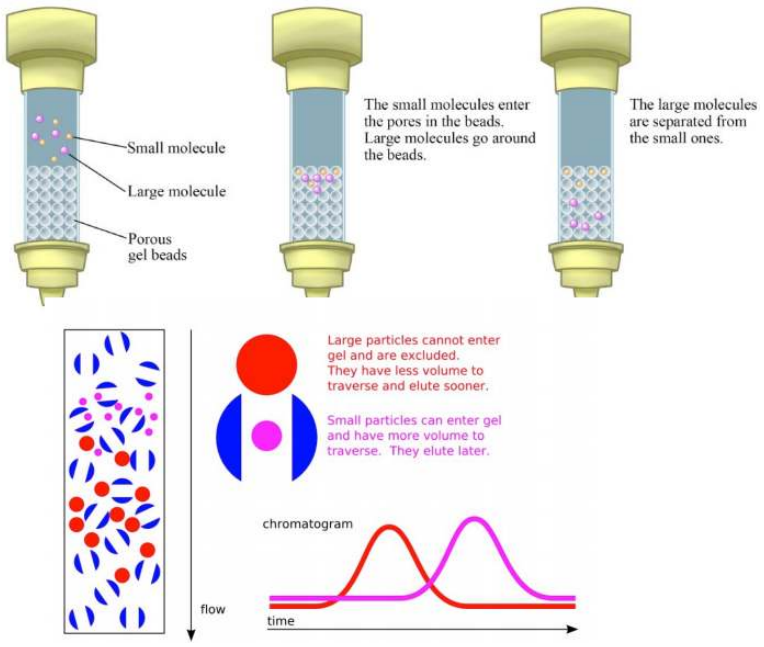


Figure 5.1.3: **Size Exclusion Chromatography.** *Image from Dr. Kevin Ahern and Indira Rajagopal*

Also known as gel filtration chromatography, it is a low-resolution isolation method involving beads with tiny “pores" that have a precise size distribution. The size is called an “exclusion limit," meaning molecules above a certain molecular weight will not fit into the tunnels. Molecules with sizes larger than the exclusion limit do not enter the tunnels and pass through the column relatively quickly by making their way between the beads. Smaller molecules, which can enter the tunnels, do so and, thus, have a longer path in passing through the column. Because of this, molecules larger than the exclusion limit will leave the column earlier, while smaller molecules that pass through the beads will elute from the column later. This method allows separation of molecules based on their size.

In any chromatography system, there is a mobile and a stationary phase. For size exclusion chromatography, the stationary phase is usually a polymerized agarose or acrylamide bead, which contains pores of various sizes filled with the solvent. Let's pretend that the solvent (typically an aqueous buffered solution) inside the bead is trapped there and doesn't exchange with the solvent moving around the bead, so it's part of the stationary phase. The mobile phase is the solvent used to elute the column, which flows around the bead. The chromatography beads are often supplied in a dried state, which must be rehydrated in the solvent before they are packed into the column. The volume of the agarose or acrylamide bead is very small compared to the volume of solution within its hydrated form.

5.1.5.2: Size and shape effects in size exclusion chromatography

Size-exclusion chromatography is so common that we will explore it in greater detail Several different column volumes can be defined as shown in Figure 5.1.4, where the packed chromatography beads are shown as circles.

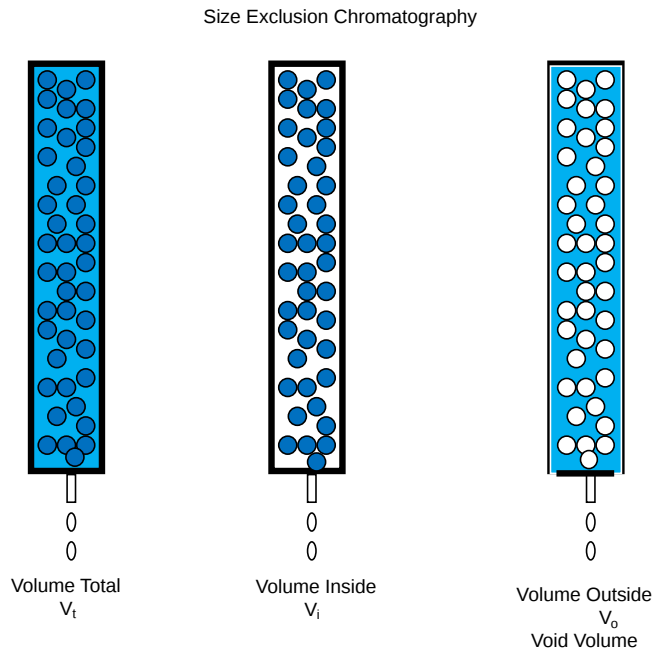


Figure 5.1.4: Define volumes in size exclusion chromatography

If we consider the mass of the beads to be *negligible relative to the volume of the beads*, the actual volume in the beads is mostly due to the trapped solution, which can be considered the "stationary" phase. The volume around the bead is called the **void volume**, V_o. It should be apparent that the volume inside the bead is given by

$$V_i = V_t - V_o \tag{5.1.1}$$

A solute elutes from the column in a broad peak. If the sample volume applied to the column is very small compared to V_t, the volume at which a solute elutes, V_e, is considered the center of the elution peak. This is true when $V_{sample} \gg V_e$.

Suppose we view this chromatography as a solute partitioning between the mobile and stationary phases (the basis of all chromatography). In that case, we might be interested in what fraction of the stationary phase, V_i, a solute might partition into. Such a ratio would be given by:

$$K = \frac{V_e - V_o}{V_t - V_o} \tag{5.1.2}$$

V_t-V_o (= V_{inside}) represents 100% of the stationary phase, where K is a **distribution coefficient**. Consider two cases:

1. A very large solute compared to the pore size of the bead: In this case, V_e-V_o = 0 since V_e would be equal to V_o. (The solute wouldn't "see" any of the V_i.) In this case, K = 0. The solute would elute from the column's void volume, since it is too large to partition into the bead volume. All solutes of molecular weight greater than or equal to the smallest solute that can't enter the gel beads will all elute in the void volume. Hence, solutes greater than this minimal size will co-elute from the column and not be separated. V_o is usually about 30-40% of the V_t.
2. A very small solute compared to the pore size. In this case, V_e-V_o = V_t - V_o, since V_e would be equal to V_t. The solute would "see" all of the solvent within the bead. In this case, K = 1. Like above, all solutes of MW equal to or less than the largest solute that can partition into the entire volume within a bead will co-elute at a volume near V_t.

Hence K is a **partition coefficient**, which varies from 0 - 1 and represents the fraction of V_i into which a solute could partition. However, this K is not exactly a partition coefficient since the actual volume of the gel matrix is assumed to be zero above. The graph in Figure 5.1.5 shows typical V_e as a fraction of V_t for solutes of different sizes (the x-axis is V_e/V_t).

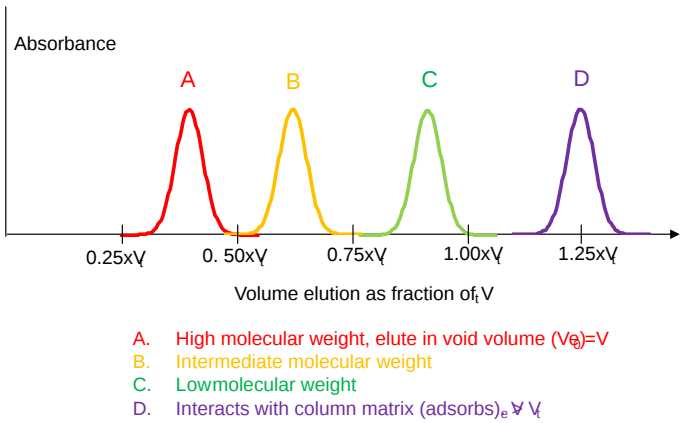


Figure 5.1.5: Elution volumes (V_e for different size macromolecules)

Large species that cannot enter the pores in the beads flow around it and elute in the void volume (V_0), which is about 35-40% of V_t (red bell-shaped curve). Very small species can partition into V_0 and V_i, so they elute near V_t (green bell-shaped curve). If a species adsorbs to the column bead via noncovalent interactions (e.g., hydrogen bonds or ion-ion interactions), it may elute after V_t (purple bell-shaped curve).

K depends on the solute's size and shape. The size and shape of an object determine its flow properties in a fluid. Frictional resistance (itself a force that acts in the opposite direction to the velocity, another vector quantity) can be shown to be proportional to the velocity.

$$F_f \propto v \tag{5.1.3}$$

or

$$F_f = -fv \tag{5.1.4}$$

where f is the frictional coefficient, which depends on the shape. The bigger the object, the more frictional resistance to movement. For a sphere, it can be shown that:

$$f = 6\pi\eta R_s \tag{5.1.5}$$

where η is the viscosity (a measure of the resistance to flow of a liquid - water has a low viscosity, real maple syrup has a high viscosity), and R_s (Stokes radius) is the radius of the hydrated sphere (the larger R_s, the larger the frictional coefficient, the larger the F_f which resists motion). For an irregularly shaped object, the Stokes radius is the radius of a sphere that would have the same frictional coefficient as the object. Hence, the R_s for a protein molecule that was not spherical would be much larger than the R_s for another protein molecule of identical molecular weight that was spherical. Hence, the V_e and the K values for a solute on a gel filtration column would best be related to the Stokes radius since R_s values consider both size and shape.

If you separate two proteins of equal mass but one is highly elongated, and the other is spherical, the elongated one, with a large R_S, would elute first (assuming that both don't elute together in the void volume, V_0.

Gel filtration can determine the molecular weight of an unknown, spherical (globular) protein compared to a standard curve generated from other globular proteins of known molecular weight. To ensure the proteins have the same "effective" shape, the proteins are eluted under denaturing conditions to remove shape-related contributions to the elution order.

5.1.5.3: Separation based on charge - Ion Exchange Chromatography

The chromatography resin in this type consists of an agarose, acrylamide, or cellulose resin or bead, which is derivatized to contain covalently linked positively or negatively charged groups. Proteins in the mobile phase will bind to the column via electrostatic interactions with the column's charged group. In a mixture of proteins, positively charged proteins will bind to a resin containing negatively charged groups, like the carboxymethyl group, CM ($-OCH_2COO^-$), or sulfopropyl, SP, ($-OCH_2CH_2CH_2SO_3^-$). In

contrast, the negatively charged proteins will pass through the column. The positively charged proteins can be eluted from the column with a mobile phase containing either a gradient of increasing salt concentration or a single higher salt concentration (isocratic elution). The most positively charged protein will be eluted last, at the highest salt concentration. Likewise, negatively charged proteins will bind to a resin containing positively charged groups, like the diethylaminoethyl group, DEAE ($-OCH_2CH_2NH(C_2H_5)_2^+$) or a quaternary ethyl amino group, QAE, and can be separated analogously.

Ion exchange chromatography separates compounds based on their ionic charge. The column used is selected according to its type and charge strength. Anion exchange resins carry a positive charge and are used to retain and separate negatively charged compounds (anions). In contrast, cation exchange resins carry a negative charge and are used to separate positively charged molecules (cations).

Before the separation begins, a buffer is pumped through the column to equilibrate the opposing charged ions. Upon sample injection, solute molecules will exchange with the buffer ions as each competes for the binding sites on the resin. The length of retention for each solute depends upon the strength of its charge. The most weakly charged compounds will elute first, followed by those with successively stronger charges. Because of the separating mechanism, pH, buffer type, buffer concentration, and temperature all play important roles in controlling the separation.

Figure 5.1.6 shows a ***cation exchange column***. The beads (brown) contain negatively charged functional groups that can bind positively charged protein (blue) or concentrated regions of positive charge on a protein.

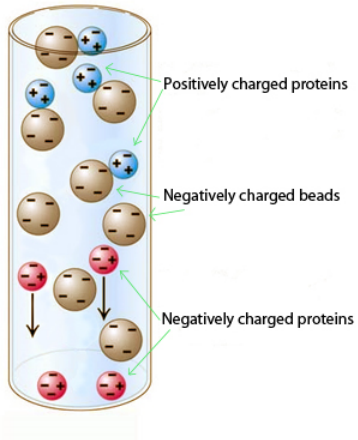


Figure 5.1.6: Ion Exchange Chromatography https://upload.wikimedia.org/wikiped...n_exchange.jpg. Jspiteri at English Wikibooks, Public domain, via Wikimedia Commons

Before loading the column with protein, the negatively charged beads would interact with positively charged counterions (often Na^+) from the column equilibration buffer. When the protein solution is introduced to the column, the positively charged protein will exchange with the bound Na^+ ions (hence the name cation exchanger). Conversely, an anion exchanger consists of positively charged beads, which exchange anions. Proteins bound via ion-ion interactions can be eluted by increasing the Na^+ concentration in the elution buffer stepwise or by using a salt concentration gradient. Ion exchange chromatography is a powerful protein purification tool and is frequently used in analytical and preparative separations.

5.1.5.4: Affinity Chromatography

In this technique, the chromatography resin is derivatized with a group that binds to a specific site on a given protein of interest. It may be a group that binds to the active site of an enzyme (such as benzamidine-agarose, which is used to purify trypsin) or an antibody that recognizes a specific amino acid sequence (an epitope) on a protein. For example, an antibody can be made to a specific peptide from albumin, the antibody covalently linked to agarose, and the antibody-agarose column then used to purify albumin specifically. This is a powerful technique since antibodies can be made that will bind selectively to a single protein. Knowing only the DNA sequence of a protein that has never been previously isolated, the amino acid sequence of the unknown protein can be derived from the DNA sequence. A 10-12 amino acid peptide from that protein can be synthesized in the lab (see the last section below), and an antibody raised against the peptide. The antibody will most likely bind to the unknown protein and the peptide, and could be used to purify the protein.

These features of affinity chromatography are illustrated in Figure 5.1.7.

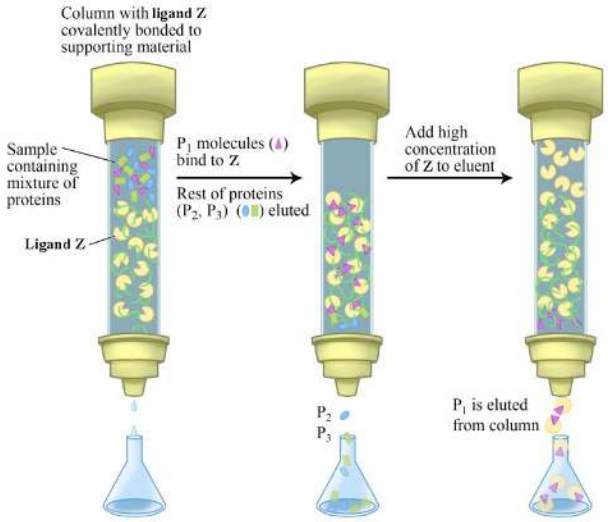


Figure 5.1.7: **Example of Affinity Chromatography.**

In this example in Figure 5.1.7, protein P1 has an affinity for ligand Z and will bind to the column, while proteins P2 and P3 will pass through the column. Protein P1 can then be eluted from the column using high concentrations of free ligand Z.

? In vitro peptide synthesis for antibody production

For more details on the chemistry of in vitro peptide synthesis, click the link below.

Answer

When making anti-peptide antibodies that recognize target proteins, or studying an isolated peptide on its own, it is more difficult to isolate and purify a peptide from its original protein than to synthesize it in the lab using solid-phase synthesis. We describe this technique below.

Peptides are chemically synthesized by the condensation reaction of the carboxyl group of one amino acid to the amino group of another. Two chemical challenges must be addressed. The formation of an amide bond between the carboxylic acid of one amino acid and the amine of the other is thermodynamically unfavorable, so the carboxyl end must be activated, typically by reaction with a reagent such as dicyclocarbodiimide. Secondly, reactive functional groups on the side chains and the amine of the carboxyl group-activated amino acid must be protected from unwanted reactions. Chemical peptide synthesis most commonly starts at the carboxyl end of the peptide (C-terminus) and proceeds toward the amino-terminus (N-terminus). Protein biosynthesis in living organisms occurs in the opposite direction. Chemical synthesis enables the production of peptides incorporating unnatural amino acids, peptide/protein backbone modification, and D-amino acids.

The established method for producing synthetic peptides in the lab is known as solid-phase peptide synthesis (SPPS). Pioneered by Robert Bruce Merrifield, SPPS allows the rapid assembly of a peptide chain through successive reactions of amino acid derivatives on an insoluble porous support. The solid support consists of small polymeric resin beads functionalized with reactive groups (e.g., amine or hydroxyl groups) that bind to the nascent peptide chain. Since the peptide remains covalently attached to the support throughout the synthesis, excess reagents and side products can be removed by washing and filtration. This approach circumvents the comparatively time-consuming isolation of the product peptide from solution after each reaction step, which would be required in conventional solution-phase synthesis.

Each amino acid to be coupled to the peptide chain N-terminus must be protected on its N-terminus and side chain using appropriate protecting groups such as t-Boc (t-butyloxycarbonyl-, acid-labile) or fluorenylmethyloxycarbonyl (Fmoc, base-labile), depending on the side chain and the protection strategy used (see below).

The general SPPS procedure involves alternating cycles of N-terminal deprotection and coupling reactions. The resin can be washed between each step to remove byproducts. Figure 5.1.8 shows the mechanism for the solid phase synthesis of a

dipeptide.

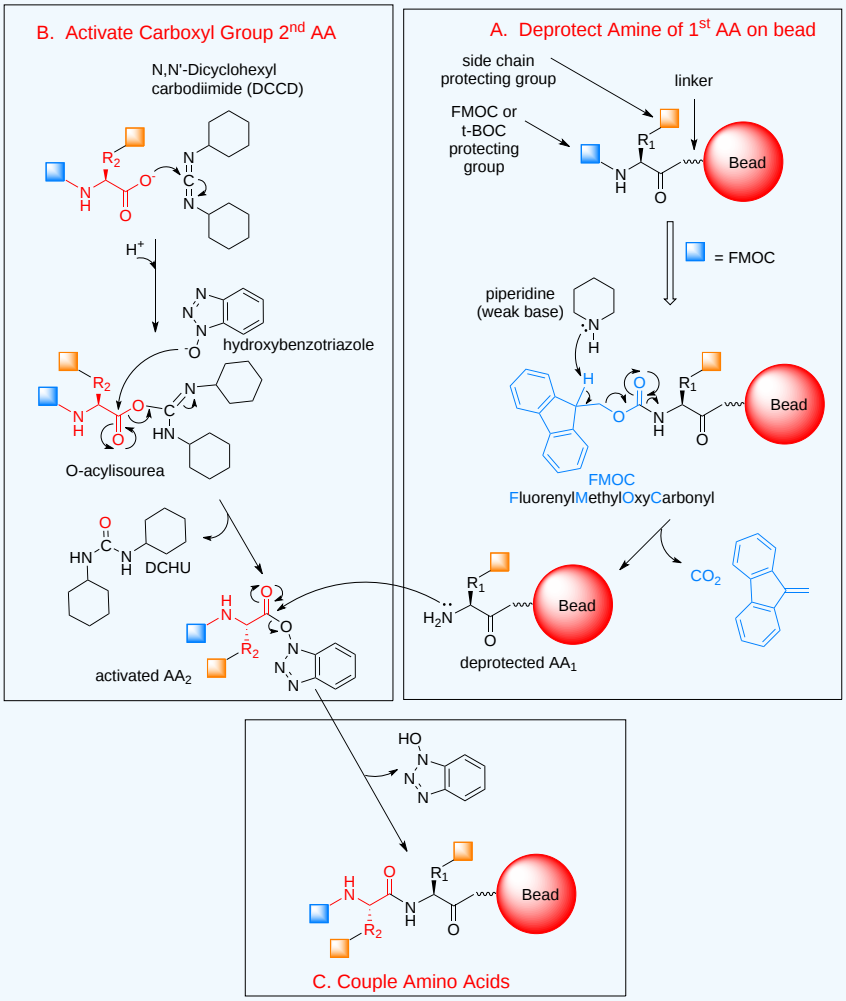


Figure 5.1.8: Solid phase peptide/protein synthesis

A. Deprotection of AA_1: The first amino acid is coupled to the resin or purchased pre-coupled. The amine terminus contacting an FMOC group is deprotected with piperidine. The hydrogen abstracted from the FMOC is acidic, as its negatively charged conjugated base is aromatic since the negative charge on that C becomes sp^2 hybridized to create the aromatic anion. The weak base piperidine is used to avoid side reactions.

B. Activation of AA_2: The carboxyl group of AA_2 reacts with a carbodiimide, which is attacked by the carboxylate of AA_2, leading to the formation of an isourea derivative. This can react with a second nucleophilic catalyst (which is regenerated in step C), hydrobenzotriazole (HBT), to form the activated HBT ester and the very stable urea derivative.

C. Coupling Reaction: The activated AA_2 now reacts with the amine of the solid phase N-terminal deprotected AA1 to form the peptide bond.

This cycle repeats until the desired sequence has been synthesized. At the end of the synthesis, the crude peptide is cleaved from the solid support while simultaneously removing all protecting groups with a strong acid reagent, such as trifluoroacetic acid. The crude peptide can be precipitated from a non-polar solvent like diethyl ether to remove organic-soluble by-products and then purified using reversed-phase HPLC. The purification process, especially of longer peptides, can be challenging because small amounts of several byproducts, which are very similar to the product, must be removed. For this reason, so-called continuous chromatography processes, such as MCSGP, are increasingly used in commercial settings to maximize yield without sacrificing purity.

Reaction yields limit SPPS, and typically peptides and proteins with 70 or more amino acids push the limits of synthetic accessibility. Synthetic difficulty is also sequence-dependent; aggregation-prone sequences, such as those found in

amyloids, are typically difficult to make. Longer lengths can be accessed using ligation approaches, such as native chemical ligation, in which two shorter, fully deprotected synthetic peptides can be joined in solution.

Cell proteins can be increasingly engineered by manipulating their gene to contain a molecular tag, either a small peptide or a protein for which antibodies are commercially available. The tag is expressed at either the target's N- or C-terminal end to avoid interfering with the folding of the target protein. Examples of peptide tags include His (sequence HHHHHH), FLAG (DYKDDDDK), and HA (YPYDVPDYA). The HA tags derive from the influenza hemagglutinin protein. A small protein, such as the green Fluorescent Protein (GFP), can also be used as a tag. The resulting fusion protein of GFP connected to the target protein can also allow the target protein to be localized and followed by confocal fluorescence microscopy within the cell. Chromatography resins with covalently attached antibodies to the His, FLAG, HA peptide tags, and GFP are commercially available as affinity chromatography resins, as shown in the right-hand side of Figure 5.1.9 below.

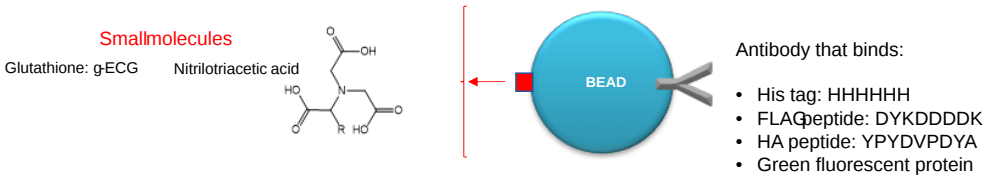


Figure 5.1.9: Protein Tags and small molecule bead ligands for affinity chromatography

Affinity reagents other than antibodies can be attached to the beads, as shown on the left-hand side of Figure 5.1.9. Two, in particular, are Ni-Nitrilotriacetic acid (Ni-NTA) and the short peptide glutathione (γ-gluatmylcysteinylglycine). They also bind tagged proteins. The Ni-Nitrilotriacetic binds the His tag by chelating the nickel ion with the six histidine imidazole groups on the His-tagged protein. (Note that His tags can also be bound to anti-His tag antibody beads.) Glutathione binds to a protein tag, Glutathione S-transferase (GST), which is fused to the target.

The structure of the Ni^{2+}-NTA complex attached to a bead and imidazoles (on a His_6 tag) is shown in Figure 5.1.10 below.

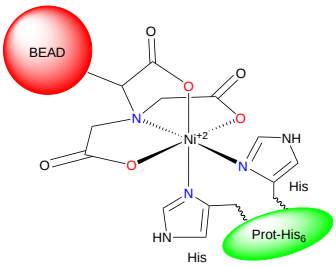


Figure 5.1.10: Ni^{2+}-NTA - imidazole complexes (after Wegner and Spatz, https://doi.org/10.1002/anie.201210317)

Figure 5.1.11 shows an interactive iCn3D model of the Ni^{2+} NTA complex.

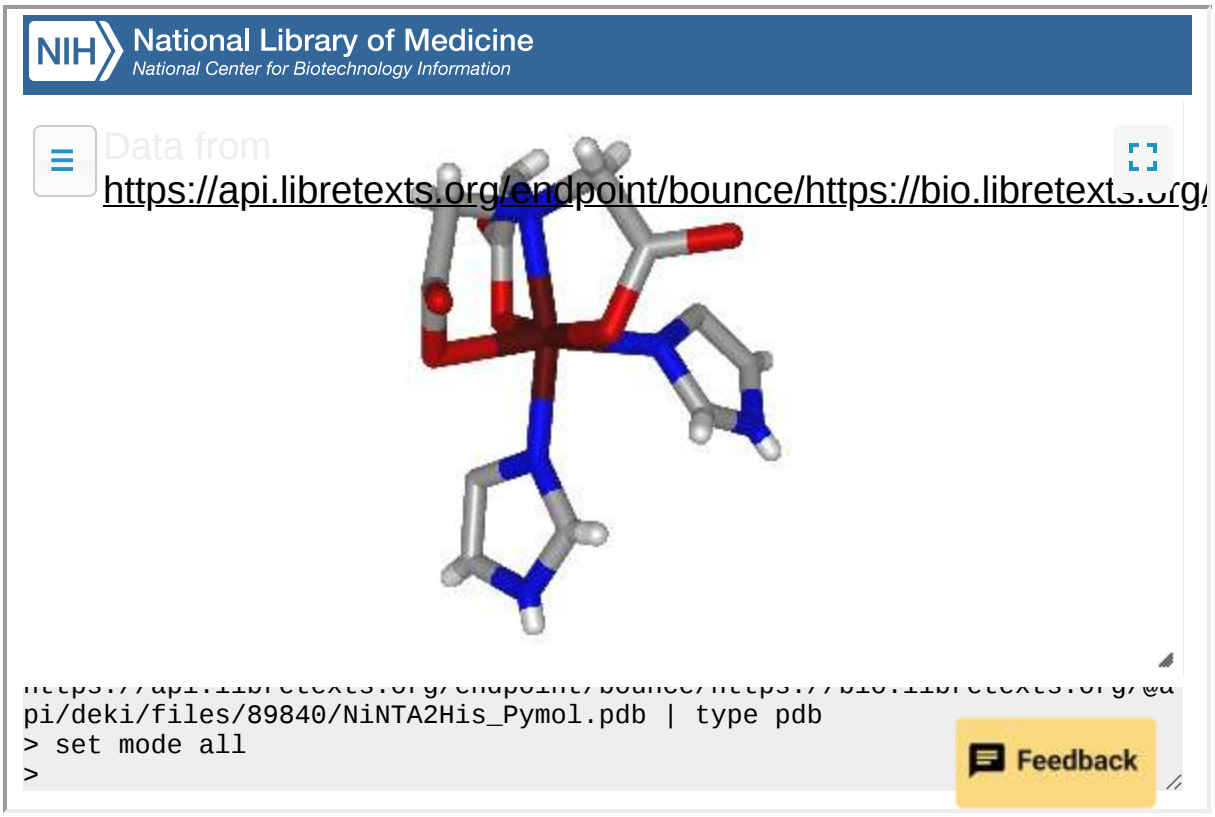


Figure 5.1.11: iCn3D model of the Ni^{2+}-NTA complex. To see the Ni^{2+} in spacefill, follow these commands in the iCn3D window. Click the = menu icon and then:

- **Select, Select on 3D**, choose Atom
- Alt-click the central Ni ion
- **Style, Chemical, Sphere**

Without a His_6 tag, two water molecules occupy adjacent axial and equatorial positions (shown interacting with imidazole rings of His in the images above) in the generally octahedral complex. One ring C in the iCn3D models shows a carbon with an unfilled valence. The bead is covalently attached to that carbon.

The His tag, probably the most widely used, binds strongly to divalent metal ions such as nickel and cobalt. The protein can be passed through a column containing Ni-nitrilotriacetic acid. All untagged proteins pass through the column. The protein can be eluted with imidazole, which competes with the imidazole side chain on the His tag for binding to the column, or by decreasing pH (typically to 4.5), which decreases the affinity of the tag for the resin. While this procedure is generally used to purify recombinant proteins with an engineered affinity tag (such as a 6xHis tag), it can also be used for natural proteins with an inherent affinity for divalent cations.

5.1.5.5: Hydrophobic Interaction Chromatography (HIC)

HIC media is similar to reverse phase chromatography in which a matrix like silica (very polar with exposed OH groups) is derivatized with ester or ether links from the silica surface hydroxyls to nonpolar molecules, usually containing 8 or 18 carbons in the acyl or alkyl chain. Proteins with exposed hydrophobic groups would preferentially bind to the bead. The interactions between the protein and the derivatized beads increase when high concentrations of salt are added to the aqueous solution, making water more polar. This would shift the equilibrium towards binding the surface-exposed nonpolar region on the protein to the nonpolar C8 or C18 chains. The ionic strength of the buffer is then reduced to elute proteins in order of increasing hydrophobicity, as shown in Figure 5.1.11.

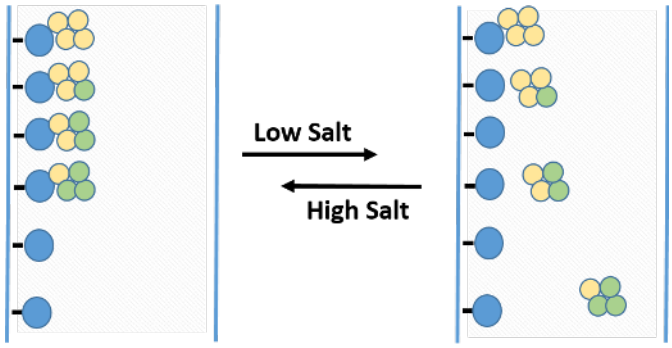


Figure 5.1.11: Hydrophobic Interaction Chromatography.

The column matrix, shown in blue, has a hydrophobic ligand covalently attached. In high salt conditions, proteins bind to the matrix with differing affinities, with more hydrophobic proteins (shown in yellow) binding more tightly than more hydrophilic proteins (shown in green). When the salt concentration decreases, more hydrophilic proteins are released first, followed by more hydrophobic proteins.

5.1.5.6: High Performance Liquid Chromatography (HPLC) and Fast Protein Liquid Chromatography (FPLC)

High-performance liquid chromatography, or high-pressure liquid chromatography (HPLC), applies high pressure to drive solutes through the column faster than gravity-driven solvent flow. The small and close-packed packing beads allow for greatly increased resolution. Because of the close packing of the small beads, no flow would occur with an external pump. The most common form of HPLC is "reversed phase" HPLC, where the column packing material is hydrophobic. The proteins are eluted by a gradient of water and increasing amounts of an organic solvent, such as acetonitrile. The proteins elute in order of their hydrophobicity. After purification by HPLC, the protein is dissolved in a solution containing only volatile compounds and can be easily ***lyophilized (freeze-dried)***. HPLC purification often results in protein denaturation and is thus not applicable to proteins that do not spontaneously refold.

Due to the drawbacks of HPLC, an alternative technique using lower pressure was developed called ***Fast Protein Liquid Chromatography (FPLC).*** In FPLC, the mobile phase is an aqueous solution, or "buffer". A positive displacement pump controls the buffer flow rate and is normally kept constant. In contrast, the buffer composition can be varied by drawing fluids in different proportions from two or more external reservoirs. The stationary phase is a resin composed of beads, usually of cross-linked agarose, packed into a cylindrical glass or plastic column. Depending on the application, FPLC resins are available in a wide range of bead sizes and surface ligands.

An ion exchange resin is typically chosen in the most common FPLC purification systems, as shown in Figure 5.1.12.

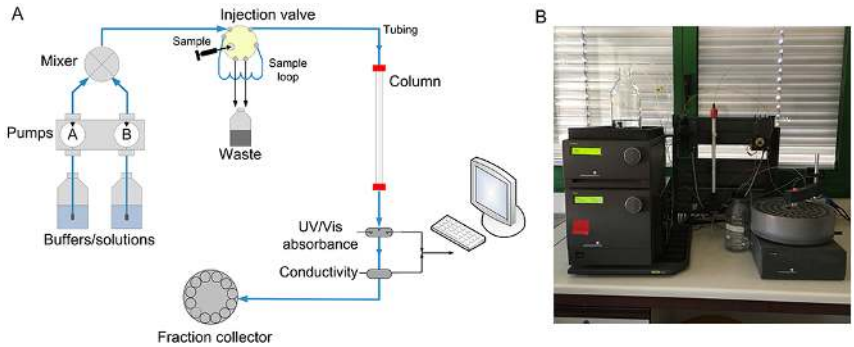


Figure 5.1.12: **Typical FPLC System.** A. Scheme of basic components and typical flow path for a chromatography system. B. Picrue of GE Healthcare AKTA FPLC apparatus. *Image provided by LaVerde, V., Dominici, P. and Astegno, A. (2017)* Bio-protocol *7(8): e2230.*

A mixture containing one or more proteins of interest is dissolved in 100% buffer A and pumped into the column. The proteins of interest bind to the resin while other components are carried out in the buffer. The total buffer solution flow rate is kept constant; however, the proportion of Buffer B (the "elution" buffer) is gradually increased from 0% to 100% according to a programmed concentration change (the "gradient"). Buffer B contains high concentrations of the exchanger ion. Thus, as Buffer B's concentration increases, bound proteins will dissociate due to ionic interactions between the exchanger ion and the column matrix,

and they will appear in the eluant. The eluant passes through two detectors, which measure salt concentration (by conductivity) and protein concentration (by absorbance at 280 nm). As each protein is eluted, it appears in the eluant as a "peak" in protein concentration and can be collected for further use.

5.1.6: Purification Scheme

During the protein purification process, a quantitative system is necessary to determine the total amount and concentration of the target protein at each step, its biological activity, and its overall purity. This will help guide and optimize the purification method being developed. Ineffective separation techniques can be disregarded, and other techniques that give higher yields and retain the protein's biological activity can be adopted.

Thus, each step in the purification scheme is quantitatively evaluated for the following parameters: total protein, total activity, specific activity, yield, and purification level. The terms are described below.

Pretend you are a researcher who wants to isolate a novel, unknown protein from a bacterial culture. You grow 500 ml of the bacteria overnight at 37 °C and harvest the bacteria by centrifugation. You remove the culture broth and retain the bacterial pellet. You then lyse the bacteria using freeze/thaw in 10 mL of reaction buffer. You then centrifuge the lysed bacteria to remove insoluble material and retain the supernatant containing soluble proteins. Your protein of interest has a biological activity you can measure using a simple assay that causes a color change in the reaction mixture, as illustrated in Figure 5.1.13. You also note that this reaction rate increases with increasing concentrations of your protein supernatant.

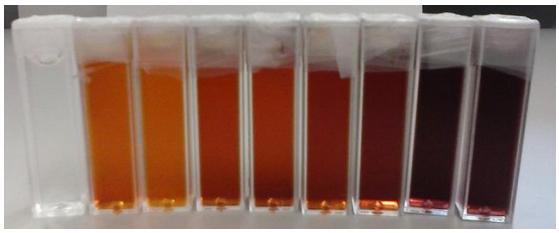


Figure 5.1.13: **Example of a Chemical Reaction that causes a color change from orange to brown depending on increasing concentration.** *Image from: Ludwig, N., et. al. (2015) on* Research Gate

At this point, you can measure your baseline concentrations for the first purification step (bacterial lysis and centrifugation to remove insoluble proteins and other cellular debris).

Total Protein is calculated by measuring the concentration in a fraction of the sample and then multiplying that by the total volume of your sample. In this case, you are starting with 10 mL of supernatant. In a typical protein concentration assay, you will use 50-200 µL of sample. For example, if you calculate that there is 7.5 µg/µL in your initial assay, you would need to convert that value into mg/mL and then multiply it by 10 mL for a total of 75 mg of protein in 10 mL of supernatant (Table 5.1.1).

Total Activity is measured as the enzyme activity in the assay multiplied by the sample volume. For example, you might use 5 to 50 µL of the sample in your biological reaction. If you calculated the activity in your assay to be 2.5 units/µL, this would be equivalent to 2,500 units/mL or 25,000 units/10 mL of supernatant. Note that the ***enzyme unit***, or international unit for the enzyme (symbol U, sometimes also IU), describes the ***enzyme's catalytic activity***. 1 U (µmol/min) is defined as the amount of the enzyme that catalyzes the conversion of one micromole of substrate per minute under the specified conditions of the assay method.

Specific Activity is measured by dividing the Total Activity by the Total Protein. In our example, 25,000 units divided by 75 mg of protein = 333.3 units/mg.

Yield measures the biological activity retained in the sample after each purification step. The amount in the first step is set to be 100%. All subsequent yield steps will be evaluated using the first purification step. It is calculated by dividing the current step's total activity by the first step's total activity and multiplying it by 100.

The ***Purification level*** evaluates the purity of the protein of interest by dividing the specific activity measured after each purification step by that of the first purification step. Thus, the first step always has a value of 1. A typical purification analysis scheme is shown below in Table 5.1.1.

Purification Step	Purification Method	Total Protein (mg)	Total Activity (units)	Specific Activity (units/mg)	Yield (%)	Purification Level
1	Cell Lysis - Centrifugation	75	25,000	333.3	100	1
2	Salting Out	54	23,416	433.6	93.7	1.3
3	Ion Exchange Chromatography	12.3	21,227	1725.8	84.9	5.2
4	Size Exclusion Chromatography	0.93	18,633	20,035	74.5	60.1

Table 5.1.1: A typical purification analysis scheme.

Note that after each purification step, the ***total protein*** goes down, as you separate the target protein from other proteins in the mixture. ***Total Activity*** also goes down with each purification step, as some of your protein of interest is also lost at each purification step because (1) some protein will stick to the test tubes and glassware, (2) some protein won't bind with 100% efficiency to your column matrix, (3) some protein may bind too tightly to be removed from the column matrix during elution, and (4) some protein may be denatured or degraded during the purification process.

The overall ***percent yield*** **(or recovery)** for each purification step represents the amount of your protein of interest that is retained. Alternative purification methods should be explored if the percent yield is too low.

Note that in a good protein purification scheme, the ***specific activity*** should go up substantially with each level of purification, as the amount of your protein of interest makes up a greater percentage of the total protein within that fraction. If the specific activity only increases modestly within a purification step, or if it decreases during a purification step, this could indicate that (1) your protein of interest is being substantially lost at that step, (2) your protein of interest is being denatured or degraded and is no longer biologically active, or (3) that a required cofactor or binding protein is being reduced at that purification step. Additional experiments may be needed to determine which factors predominate, so that steps can be taken to reduce protein inactivation. For example, many proteins are temperature sensitive and will degrade or denature at room temperature. Completing purification steps on ice can often reduce degradation.

The **fold-increase** (specific activity at step n divided by the initial specific activity) or ***purification level*** should increase during purification. A high-quality purification procedure results in a high fold increase or purification level AND a high recovery or yield.

5.1.7: Electrophoresis: Separation and Analysis

In column chromatography, flow through the column is driven by hydrostatic pressure, causing flow from higher-pressure regions at the top of the column reservoir to lower-pressure regions at the bottom of the column (drops eluting from the bottom of the column). Ultimately, the hydrostatic pressure (in columns not driven by mechanical pumps) is due to gravity. However, proteins are charged particles and can be moved by an external electric field rather than a gravitational field. Electrophoresis is the movement of charged particles in an electric field. As we will show below, the movement of a charged protein within a static matrix in the presence of an external electric field depends on both size and shape. Electrophoresis can be used for both analytical and preparative separations of proteins. The most common uses are for analytical separations.

5.1.7.1: Theory

What determines how a protein moves in an electric field? Consider a charged particle (+Q) moving in an electric field (E) in a nonconducting medium, such as water. Suppose the particle is moving at a constant velocity toward the cathode (- electrode where cations go). In that case, the net force F_{tot} on the particle is 0 (since F=ma, and the acceleration (a) of the particle is zero at constant velocity). Two forces are exerted on the particle, one F_E, the force exerted on the charged particle by the field, which is in the direction of the motion (toward the cathode), and the other, F_f, the frictional force on the charged particle, which retards its motion toward the cathode, and hence is in the direction opposite to the motion (toward the anode (+) electrode). This is shown in the Figure 5.1.14:

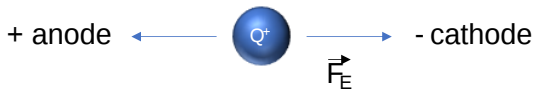


Figure 5.1.14: Movement of charge Q^+ in an electric field

Therefore:

$$\mathbf{F_{tot} = F_E + F_f} \tag{5.1.6}$$

where F_E, the electrical force, is

$$\mathbf{F_E = QE} \tag{5.1.7}$$

and

F_f, the frictional force, is

$$\mathbf{F_f = -fv} \tag{5.1.8}$$

In the last equation, v is the particle's velocity, and f is a constant called the frictional coefficient. This equation shows that the force F_f hindering motion toward the cathode is proportional to the particle's velocity. This is intuitive, since one would expect that the higher the velocity, the greater the F_f, which would hinder motion. The frictional coefficient depends on a molecule's size and shape. The larger the molecule, the larger the frictional coefficient (i.e., more resistance to the molecule's motion). It can be shown that the frictional coefficient for a spherical particle is given by

$$\mathbf{f = 6\pi\eta R_S} \tag{5.1.9}$$

where η is the viscosity (a measure of the resistance to flow of a liquid - water has a low viscosity, real maple syrup has a high viscosity), and R_s (Stokes radius) is the radius of the hydrated sphere (the larger R_s, the larger the frictional coefficient, the larger the F_f which resists motion toward the cathode). This equation should be intuitive based on your experiences.

When the velocity of the charged particles is constant (i.e, there is no net force on the particle), $F_e = F_f$, and using equations 3.3.6-8 gives

$$\mathbf{QE = fv} \tag{5.1.10}$$

Hence v/E = Q/f = U = the **electrophoretic mobility**, or

$$\mathbf{U = \frac{V}{E} = \frac{Q}{6\pi\eta R_S}} \tag{5.1.11}$$

Therefore, the electrophoretic mobility U is proportional to the **particle's charge density (Q/Rs)**, not just its size, as is the case for spherical proteins in size-exclusion chromatography. Macromolecules of different **charge densities** can thus be separated by electrophoresis. This discussion deals with the simplest case since, in reality, there are counterions in the solution (from salts), which would form a cloud around the charged macromolecule and partially shield the charged particle from the electric field E.

Modern-day electrophoresis is conducted in solid gels (such as polyacrylamide) formed from liquid acrylamide solutions after the addition of a polymerizing agent. The solid gel is porous to solute and solvent molecules. It serves as a medium for electrophoresis

while helping to eliminate convection forces in the liquid that interfere with the separation. Electrophoretic experiments have been conducted on the space station in weightless conditions to prevent such perturbations.

One complication in this idealized description of electrophoresis in polyacrylamide gels is that the gels contain pores through which the macromolecules can move. Think of the protein moving under an electric force through a "spider web-like" matrix. As in gel chromatography, smaller molecules can pass through the pores more readily than larger ones, so an additional sieving mechanism contributes to *effective* mobility. (Also, the gel could alter the local effective electric field.) The gel's sieving effect increases the resolving power of this technique.

It has been determined that the actual electrophoretic mobility of the protein, U, is a function of the mobility of the protein in a concentrated sucrose solution (Uo) and T, the total concentration of the acrylamide in the polymerized gel. The higher the concentration of acrylamide in the unpolymerized gel solution, the smaller the pores in the polymerized gel. An equation showing the relationship between U, Uo, and T is shown below:

$$\log U = \log U_0 - K_r T \tag{5.1.12}$$

where K_r is the slope of a log U vs T plot for a given protein. Since K_r is a function of the radius of the molecule, it is possible to determine the molecular weight of a protein molecule by performing several electrophoretic separations in gels of different acrylamide concentrations (T) and extrapolating results to T = 0, hence eliminating pore size effects. Problems arise, however, if the proteins are **not** spherical.

Is there any way to obtain molecular weight information, in addition to purity determination, on a single gel? What would result if two different proteins with the same molecular weight and total net charge but different shapes were run on a single acrylamide gel? The more elongated shape (large Stokes radius) would have lower electrophoretic mobility ($U = Q/6\pi\eta R_s$). A larger R_s would also cause the protein to enter the pores more slowly. Hence, electrophoretic mobility and sieving effects would cause this protein to run anomalously slowly and to appear to have a higher apparent molecular weight. Also, imagine two globular proteins of different sizes but with compensatory charge differences, which might allow them to migrate at the same speed through the gel.

An astute reader might quickly recognize a problem separating proteins by gel electrophoresis. Some proteins are negatively charged (pH > pI), some would be neutral (pH=pI), and the rest would be positive (pH < pI). Only some proteins would enter the gel and migrate toward the electrode at the bottom. Luckily, there is a way to eliminate both charge and shape effects in the electrophoresis of proteins: run the gel under denaturing conditions so that all proteins have the same charge density. The denaturant of choice for electrophoresis is usually sodium dodecyl sulfate (SDS), which is an ionic detergent with the structure $CH_3(CH_2)_{10}CH_2OSO_3^-$ (a single chain amphiphile). This detergent binds to and denatures most proteins, with about 1.4 g SDS binding per gram of protein (about 1 SDS/2 amino acids). Since there is one negative charge/SDS, the binding of SDS masks any of the charges on the protein and gives all proteins an overall large negative charge. *Additionally*, SDS-protein complexes have been shown to have an elongated, cylindrical shape. Since the amount of SDS bound per unit protein mass is constant, the overall charge density on all proteins is similar, so the electrophoretic mobility is only determined by sieving effects.

SDS also eliminates shape differences among proteins as a variable, since all proteins have the same general rod-like shape. (The use of SDS is analogous to using 8M urea in the gel chromatographic separation of proteins to determine molecular weights.) Mobility becomes only a function of the molecular weight of the protein and not its shape. The molecular weight of an unknown protein can be determined by comparing the protein's position on an SDS polyacrylamide gel with a series of known molecular weight standards, from which a linear plot of the ln Mr vs R_f can be used to calculate unknown molecular weights. This is similar to the analysis in gel chromatography, where ln Mr is a linear function of K_{avg}, the distribution coefficient when the gel is run under denaturing conditions. However, some proteins run anomalously on such gels (due to incomplete or excessive SDS binding), so alternative molecular weight-determination techniques should be used in conjunction with this method.

Proteins are usually heated in SDS to 100 °C for 3 minutes in the presence of a reducing agent such as β-mercaptoethanol (βME) to denature them into rod-shaped molecules. Apparent molecular weight can be obtained under non-reducing conditions (without βME), but these should be considered just estimates. Running proteins, both in the presence and absence of a reducing agent, can provide important information about a protein's subunit structure. A multimeric protein whose subunits are held together by disulfide bonds can be resolved into component subunits by adding a reducing agent. If the subunits are held together only by

noncovalent interactions, they will separate under denaturing conditions (SDS), eliminating subunit interactions in the presence or absence of β-ME.

5.1.7.2: Electrolytic vs Galvanic Cells

Electrode nomenclature can be confusing. In the above example, in the presence of an existing electric field produced by a power source, the +Q particle, a cation, moved to the cathode, the negative electrode. In protein or DNA electrophoresis, you know that the negatively charged proteins or DNA fragments, which are anions, move to the bottom of the gel towards the anode, the positive electrode. However, when you studied galvanic cells (that spontaneously produce current without a power supply) in general chemistry courses, the cathode is the positive electrode, and the anode is the negative electrode. Instead of remembering the charge state of the electrode, it is better to focus on the redox reactions that occur at each electrode:

- It is always true that Oxidation occurs at the Anode - both start with a vowel - and Reduction occurs at the Cathode - both start with a consonant!

In galvanic cells, an electrical current is generated from a spontaneous set of redox half-reactions. In electrolytic cells, an external power supply is required to drive particle motion and the nonspontaneous redox reactions at the electrodes. In electrophoretic cells for protein analysis, the bubbles released at the electrodes are from the electrolysis of water - $2H_2O(l) \rightarrow 2H_2(g) + O_2(g)$. The oxidation number of H goes from +1 to 0 in the reaction, so H is reduced and is produced at the cathode. In SDS-PAGE, SDS-coated proteins (negatively charged) migrate to the anode, the positive electrode in this case. Commercial electrolytic cells can produce $Cl_2(g)$ and Mg(s) from the aqueous electrolyte $MgCl_2(aq)$. Figure 5.1.15 illustrates the differences between galvanic and electrolytic cells.

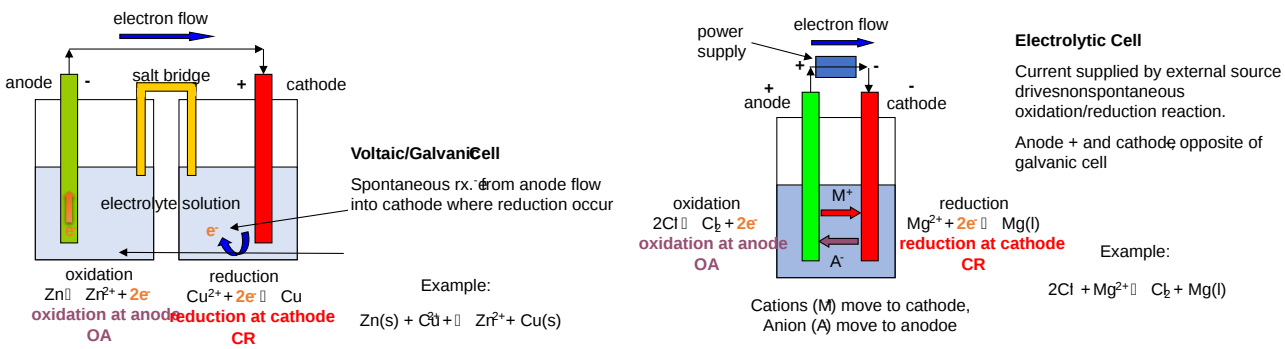


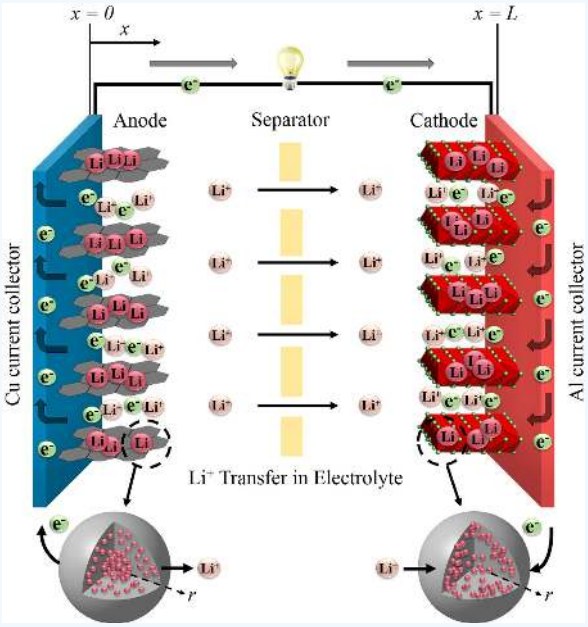
Figure 5.1.15: Galvanic vs electrolytic cells

? Lithium Ion Batteries

Optional: To support the material in Chapter 32: Biochemistry and Climate Change, to review your learning in previous courses, to prepare for more complex redox reactions (mitochondrial electron transport, photosynthesis, and nitrogen fixation) that we will encounter later in this book, and as we move to electric vehicles (EVs) to reduce greenhouse gases from burning fossil fuels, it's important to know how Li^+ ion batteries that power EVs work. All you need is basic general chemistry, which provides the foundation for understanding biochemical redox reactions.

Click the link below if you wish to learn more about lithium-ion batteries

The diagram below shows the components of a Li^+ ion battery that discharges to power electronic devices (such as a light bulb) and electric cars. Compare it to the spontaneously discharging galvanic/voltaic cell (left image above).



Schematic of the involved species/charge transports and electrochemical reactions in lithium-ion cells during discharge. Lithium atoms are shown as red spheres, Li^+ ions as light orange, and electrons as light green. Jiang, F., Peng, P. Elucidating the Performance Limitations of Lithium-ion Batteries due to Species and Charge Transport through Five Characteristic Parameters. *Sci Rep* **6**, 32639 (2016). https://doi.org/10.1038/srep32639. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

In a lithium-ion battery:

* The - electrode (anode, where oxidation occurs) is porous carbon graphite (parallel sheets of graphene) with Li atoms/ions on a conducting metal (Cu). This is where Li^+ ions and free electrons (e^-) are formed. The electrons move through the electric circuit to the + electrode (the cathode);

* A liquid electrolyte solution (replaces the "salt bridge" in the Cu/Zn^{2+} cell) allows the transfer of Li^+ to the + cathode (the tricky part to remember) to preserve charge neutrality.

* The + electrode (cathode, where reduction occurs) is a metal oxide (such as CoO_2) on a conducting metal (Al), which reduces the arriving Li^+ ions to neutral Li and intercalation into the CoO_2 polymer. This is the reaction shown in the figure above. Most references state that the Co(IV) in CoO_2 is reduced to Co(III), which is perhaps more likely given the low ionization energy of the gaseous form of Group 1 Li.

Here are the actual chemical reactions that occur at the anode and cathode during the discharge of the battery to power a light or an electric car:

Li atoms (neutral, red spheres) in the graphite anode are oxidized to Li^+ ions (yellow spheres) as shown in this reaction:

$$\text{Oxidation at anode: } LiC_6 \rightarrow C_6 + Li^+ + e^-$$

The Li^+ ions move from the anode through the electrolyte to the cathode, where they are reduced (like Cu^{2+} in the left figure above). Here is the reaction:

$$\text{Reduction (at cathode): } Li^+ + e^- + CoO_2 \rightarrow LiCoO_2$$

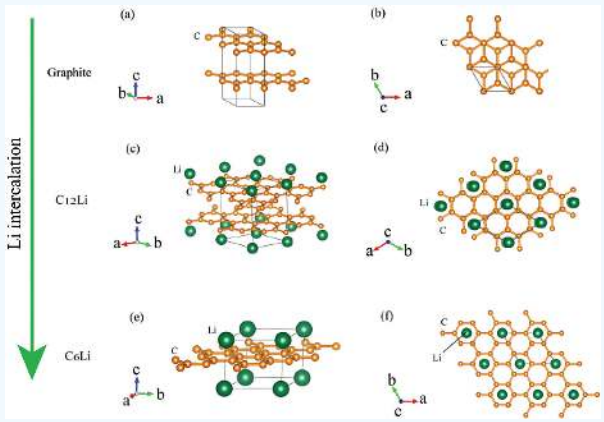
The last reaction could indicate that either Li^+ or Co(IV) is reduced to maintain charge balance. The oxidation number of Co in $CoCO_2$ is +4 (as each O is -2). In $LiCoO_2$, if Co is reduced, it becomes Co^{+3} as it gains an electron, with the +1 on Li balancing the net -4 on the two oxygens. If Li^+ is reduced, it becomes atomic Li^0. This presumes that the interactions are all ionic. So with charge states/oxidation numbers, this could be written as:

$$Li^+ + e^- + Co(IV)(O^{2-})_2 \rightarrow [Li^+Co(III)(O^{2-})_2 \text{ or } Li^0Co(IV)(O^{2-})_2]$$

A better way might be to envision a sharing of the electron between the Li^+ and Co(IV) ions. We'll address this more below.

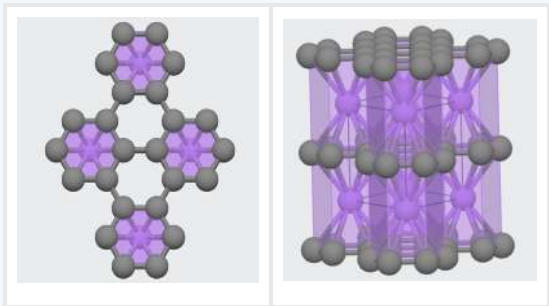
On charging (nonspontaneous, driven by plugging in the car), the reverse process occurs: reduced Li in $LiCoO_2$ donates an electron (as Zn in the left figure above) to form $Li^{+,}$ which moves to the graphite electrode. Here is an animation (which sometimes requires refreshing the browser) showing the structure and Li-ion diffusion out of a $LiCoO_2$ cathode (click the Lithium Ion Diffusion Icon).

The figure below shows the 3D structures of different graphite forms with intercalated Li in the anode.



The crystal structure of (a) and (b) graphite, (c) and (d) $C_{12}Li$, and (e) and (f) C_6Li. (b, d, and f) Show a c-axis view of graphite, $C_{12}Li$, and C_6Li, in which Li is located at the center of six-member carbon rings. ***Phys. Chem. Chem. Phys.***, 2017,**19**, 19058-19066. Creative Commons Attribution-NonCommercial 3.0 Unported

Here are two views of the LiC_6 anode structure (see this link for a 3D view). Carbons in graphite are shown as gray spheres, while Li is shown as purple spheres. Lithium migrates in a 2D plane between 2 graphene layers (graphene is a single layer of graphite). Here is another link for the structure of LiC_6.



The empirical formula for this structure is LiC_6, but it is also written as C_6Li. If we consider LiC_6 as purely an ionic compound, the lithium would appear to have a +1 charge and oxidation state, and each of the Cs would be -1/6 = -0.166, with the electron from Li entering the pi cloud of graphite. But didn't we say above that Li likely has an oxidation number/charge of 0 (i.e., it's a neutral atom) in the LiC_6 structure?

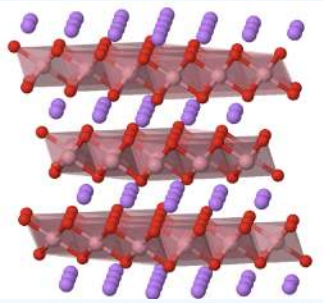
Here is some help for this problem. Assigning electrons to determine the charge state in metal ion complexes is arbitrary. It doesn't indicate the actual electron density around an atom (as in the case of H_3O^+, in which O has a formal charge of +1 but has almost all of the electron density). In the case of a transition metal complex, we can consider the bond between the metal and ligand to be a coordinate covalent bond. For a transition metal complex, the metal's oxidation number can be determined by first separating (in your mind) the ligand and its donating electrons from the metal. There are two ways to do this:

* send both electrons in the coordinate covalent bond to the withdrawn ligand (ie, a heterolytic cleavage). If the ligand is -NH_3, the removed ligand has a full lone pair (:NH_3) and is neutral. For a -CH_3 ligand, the removed ligand would be :CH_3^-, a carbanion with a negative charge.

* send one electron to the withdrawn ligand and one to the metal (i.e a homolytic cleavage). If the ligand is -CH_3, the removed ligand is $^{\cdot}CH_3$ (free radical and neutral).

For Li, a nontransition Group I metal, if we assign one electron in the bond to Li and the other to the separated C_6 electron cloud, the oxidation state of Li would be 0, consistent with the anode reaction shown above. Given the delocalized electron density from the pi clouds of the graphite interacting with Li, a coordinate covalent bond is a better descriptor than a simple ionic one. At the same time, given how easily Li gives up an electron, it makes sense to show both Li and Li^+ in the electrodes.

Here is a figure showing the $LiCoO_2$ cathode.



Lithium is shown in purple, oxygen in red, and Co in orange. https://www.chemtube3d.com/lib_lco-2/

Much research is devoted to finding new electrons and metal substitutes to make them cheaper and more environmentally favorable. Lithium manganese oxide ($LiMn_2O_4$) and lithium iron phosphate ($LiFePO_4$) are cathode alternatives that don't rely on expensive and toxic cobalt.

5.1.8: Polyacrylamide Gel Electrophoresis - PAGE

Electrophoresis is performed in a porous yet solid medium to eliminate any problems associated with convection currents. Such media are formed by solidifying a liquid agarose solution (used mostly for electrophoresis of DNA fragments and very large proteins) or by polymerizing a solution of acrylamide. Polymerization of acrylamide is initiated by the addition of ammonium persulfate in the presence of tetramethylenediamine (TEMED), along with a dimer of acrylamide (N, N'-methylene-bis(acrylamide) connected covalently between the amide nitrogens of the acrylamides by a methylene group. The structures of these compounds are shown in Figure 5.1.16.

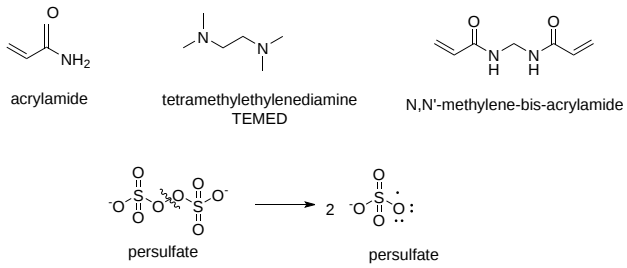


Figure 5.1.16: Structures of reagents use to make PAGE gels

As shown above, acrylamide's free-radical polymerization is initiated by adding ammonium persulfate, which generates free radicals upon dissolution in water.

The radical initiates polymerization of the acrylamide, as shown below. The TEMED, through its ability to exist as a free radical, acts as an additional catalyst for polymerization. A rigid gel is only formed, however, when N, N'-methylene-bis(acrylamide) is

added to the mixture during the polymerization, which cross-links adjacent acrylamide polymers as shown in Figure 5.1.17.

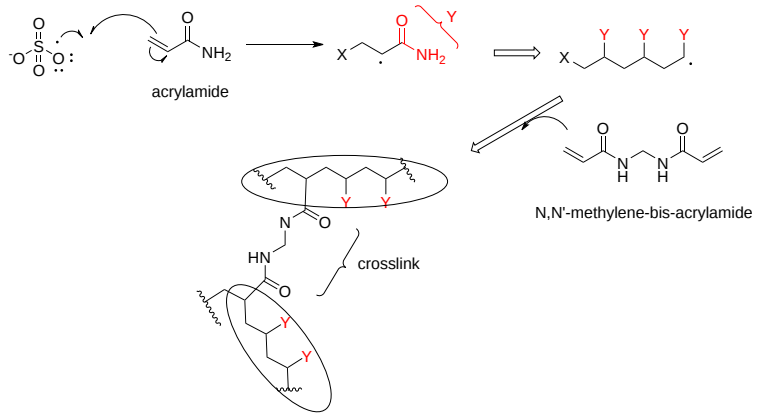


Figure 5.1.17: Polymerization of acrylamide

The amount of bisacrylamide added during polymerization controls the degree of cross-linking and, hence, the pore size of the polymerized gel. The effect of pore size is **OPPOSITE** to that in gel chromatography. In both cases, large proteins have difficulty entering the pore. In gel chromatography, large proteins partition preferentially into the mobile liquid phase (the void volume) and are eluted most **QUICKLY** from the column. In electrophoresis, large proteins that cannot readily enter the pores of the gel are not as readily transported by the electric field through the gel and elute most **SLOWLY**. Pore size can not be controlled as accurately as in the manufacture of gel chromatography resins.

How do proteins migrate through the gel? A viscous protein solution is layered on the top of the gel in a small well molded into the gel during the polymerization process. The bottom and top parts of the gel are inserted into reservoirs containing a buffered solution and the appropriate electrode. The electric field is applied, and the proteins migrate through the hydrated gel. The nature of the buffer solution in the reservoir and the polymerized gel is important. The buffer components must not bind to the proteins to be separated. Additionally, for native (non-denatured) gels, the medium pH must be adjusted so that the proteins have the appropriate charge to migrate in the expected direction.

There are many variations of electrophoresis commonly used. Gels can be polymerized in tubes or slabs, with or without denaturing agents. Additionally, a given slab might consist of two separate slabs, each polymerized on top of the other, with different acrylamide concentrations and pH values. The top part is the stacking gel, and the bottom is called the running gel. Other gels have a continuous gradient of acrylamide concentrations (from low at the top to high at the bottom). Most commercially available precast gels use continuous acrylamide concentration gradients. Figure 5.1.18 shows a gel placed in an electrophoresis chamber.

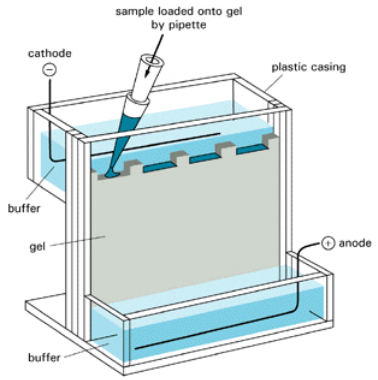


Figure 5.1.18: Electrophoresis gel and chamber https://openwetware.org/mediawiki/in...&oldid=1096467

Whether the gel has a continuous gradient or is discontinuous, the top part of the gel is a low concentration of acrylamide (2-4%), often in a Tris-HCl buffer solution (pH 6.5), usually 2 pH units below that used in the running gel. The lower part of the gel is 8-15% acrylamide, depending on the choice of gel, which is selected based on the molecular weight of the proteins to be separated. The upper buffer reservoir contains Tris buffer with a weak acid such as glycine (pKa2 = 9.6) to the same pH as the running gel.

Proteins electrophorese quickly through the low-concentration stacking gel at the top of continuous gradient gels and effectively "stack" as they hit the interface between the stacking and running gels or before they enter too far into the continuous gradient gel. This increases the compaction of proteins before they enter the "running" section of the gel, thereby improving resolution.

For discontinuous gels, how does this stacking process work? When the electrophoresis is started, glycine ions from the upper reservoir (at pH 8.7) enter the stacking gel since, at that pH, they have an average partial negative charge. The stacking gel buffer ions continue moving through the stacking gel, but when the glycine ions enter the pH 6.5 region of the stacking gel, they become zwitterions with a zero net charge and thus stop their motion toward the anode. The electrical resistance in the stacking gel increases as the number of ions moving through it decreases. To maintain a constant current throughout the circuit, there will be a *localized* increase in voltage *in the stacking gel* (by Ohm's Law, V = iR). This will cause the proteins to migrate quickly and all stack in a single, very thin disc right behind the Cl- ions in the stacking gel (which are in front because they have the highest charge density and electrophoretic mobility of any ion in the stacking gel). The proteins will not pass Cl- ions, since, if they did, they would immediately slow down, since they would no longer be in an area of diminished charged carriers and higher voltage. At the stacking gel/running gel interface, proteins cannot all migrate at the same rate due to the sieving effect of the more concentrated gel. They will be separated in the running gel. The glycine eventually enters the running gel, assumes its fully charged state at that pH (8.7), passes the proteins, and restores the charge deficiency that occurred in the stacking gel.

5.1.8.1: Detection of proteins in the gel:

Most proteins do not absorb at visible wavelengths of light and hence will not be visible during electrophoresis. To prevent proteins from eluting from the gel into the lower buffer reservoir, a small-molecular-weight anionic dye and bromophenol blue are added to the protein before it is loaded onto the gel. The electrophoresis is halted when the dye reaches the bottom of the gel. The gel assembly is removed from the electrophoresis chamber, the glass plates are separated, and the gel is washed into a series of solutions to render the banded proteins visible.

- Coomassie Brilliant Blue dye is the most common stain in labs. It is dissolved in a methanol/acetic acid solution, generating significant waste. Proteins bind this dye, resulting in a concomitant spectral shift in its absorbance properties. The methanol and acetic acid in the dye solution also help to "fix" the proteins in the gel and prevent their diffusion into the solution. After the gel is stained, the background stain is removed with acetic acid/methanol, leaving the blue-colored protein bands. Some proteins will not be stained with Coomassie blue. A "colloidal Coomassie Blue" stain doesn't use methanol or acetic acid, so it is commonly used. A simple water solution can destain the gel.
- Silver staining involves the reduction of Ag(I) to elemental silver and its deposition by protein in the appropriate reaction solutions, much as in a photographic process. (Remember, in the BCA assay, peptide bonds reduce Cu(II) to Cu(I), which is chelated to BCA.) A developer and fixer solution is required. This technique is 10-50 X more sensitive than Coomassie Blue staining. Figure 5.1.19 shows gels stained with Coomassie Blue (A)and silver staining (B).

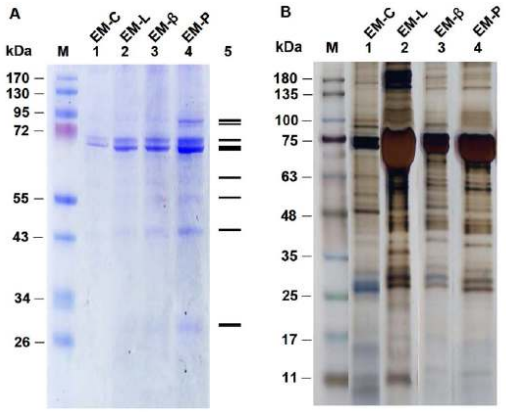


Figure 5.1.19: PAGE gels stained with Coomassie Blue (A) and silver stain (B) Chen Y-Y, Chen J-C, Lin Y-C, Kitikiew S, Li H-F, Bai J-C, et al. (2014) Endogenous Molecules Induced by a Pathogen-Associated Molecular Pattern (PAMP) Elicit Innate Immunity in Shrimp. PLoS ONE 9(12): e115232. https://doi.org/10.1371/journal.pone.0115232. Creative Commons Attribution License

- Modifying proteins with fluorescent or radioactive tags before electrophoresis increases sensitivity. After electrophoresis of a radiolabeled protein, the gel can be dried and overlaid with X-ray film for months, if necessary, to allow sufficient film

exposure by a low-concentration protein. This visualization technique is called autoradiography.

5.1.8.2: Variations on polyacrylamide gel electrophoresis:

Isoelectric focusing: In this technique, a pH gradient is set up within the polyacrylamide gel or strip. This is accomplished by pre-electrophoresing a series of low-molecular-weight molecules containing amino and carboxyl groups called **ampholytes,** each with a different isoelectric point. When subjected to an electric field, the most negative of the species will concentrate at the anode, while the most positive will concentrate near the cathode. The remaining ampholytes will migrate to their isoelectric points and establish a linear pH gradient in the gel.

Proteins initially in regions with a pH below their isoelectric point are positively charged and migrate toward the cathode. In contrast, those in a medium with a pH lower than its pI will be negatively charged and migrate towards the anode, as shown below in Figure 5.1.20. The migration will lead to a region where the pH matches the pI. There, the protein will have a net charge of zero and stop. Thus, amphoteric molecules are located in narrow bands where the pI coincides with the pH. In this technique, the point of application is not critical, as molecules will always move to their pI region. The stable pH gradient between the electrodes is achieved using a mixture of low-molecular-weight ampholytes whose pIs span a preset pH range.

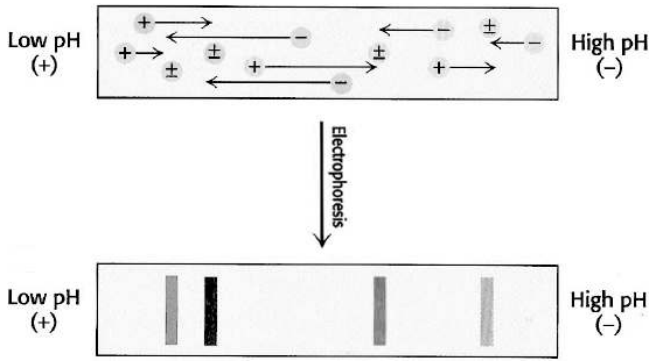


Figure 5.1.20: Isoelectric Focusing. A pH gradient is established in a gel before loading the sample. After the sample is loaded a voltage is applied. The protein will migrate to its isoelectric pH, at which they have no net charge. *Image from Magdeldin, S.*

2D electrophoresis: Two-dimensional gel electrophoresis (2-DE) separates a protein mixture according to two different molecular properties in each dimension. The most commonly used is based on a first-dimension separation by isoelectric focusing (IEF) and a second dimension by molecular weight using SDS-PAGE. A conditioning step is applied to proteins separated by IEF before the second-dimension run. This process reduces disulfide bonds and alkylates the resultant sulfhydryl groups of the cysteine residues. Concurrently, proteins are coated with SDS for separation based on molecular weight. After the IEF, the tube or strip is placed across the top of a slab gel and subjected to SDS-polyacrylamide gel electrophoresis at 90 ° to the direction of the initial isoelectric focusing experiment. Figure 5.1.21 shows a stained 2D electrophoresis gel.

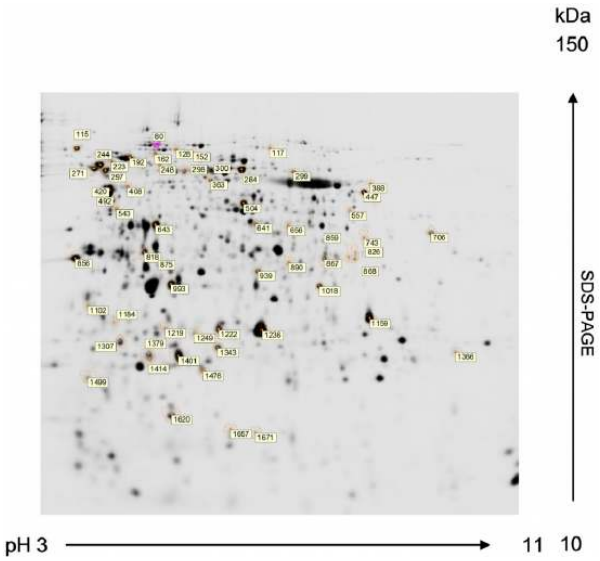


Figure 5.1.21: **Two-Dimentional Gel Electrophoresis.** *Image from Magdeldin, S*

In Figure 5.1.21, proteins of *Chlamydomonas reinhardtii are* resolved by 2-DE from preparative gels stained with MALDI-MS compatible silver reagent for peptide mass fingerprinting analysis. First dimension: isoelectric focusing in a 3-11 pH gradient. Second dimension: SDS-PAGE in a 12% acrylamide (2.6% crosslinking) gel (1.0 mm thick). Numbered spots marked with a circle correspond to proteins that are to be subsequently identified by MALDI-TOF MS. The MALDI-TOF MS analysis of protein sequences is discussed in more detail in Chapter 3.4.

One of the biggest problems in 2-DE is the analysis and comparison of complex protein mixtures. Currently, there are databases capable of comparing two-dimensional gel patterns. These systems enable automatic spot comparison to identify those needed for quantitative analysis precisely. Once interesting proteins are identified, they can be excised from gels, destained, and digested for mass spectrometry analysis. This technique is known as peptide mass fingerprinting. The ability to precisely determine molecular weight by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) and to search databases for peptide mass matches has enabled high-throughput protein identification. Proteins not identified by MALDI-TOF can be identified by sequence tagging or de novo sequencing using the Q-TOF electrospray LC-MS-MS.

Western blotting: After a standard SDS-slab electrophoresis experiment, the gel is overlaid with a piece of nitrocellulose membrane. The sandwich of gel and filter paper is placed back into an electrophoresis chamber such that the proteins migrate from the gel into the nitrocellulose, where they irreversibly bind. This is illustrated in the figure below. Note, however, that in the absence of staining, the protein bands in either the PAGE gel or Western blot would not be visible. Standards (lane 5) would be visible if they were labeled with chromophores, as shown in Figure 5.1.22.

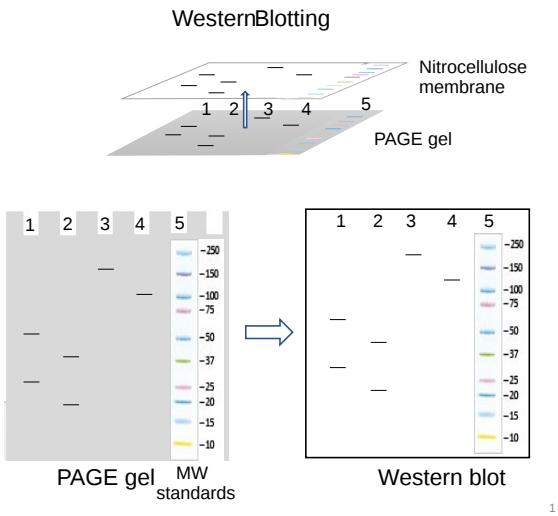


Figure 5.1.22: Western blotting of proteins

If a cell lysate were applied to a lane of a PAGE gel, the bands would appear as overlapping smears on the stained gel, regardless of the staining technique. Western blots are useful because specific bands can be visualized (stained) on the nitrocellulose membrane using a detection system linked to an antibody that recognizes a specific target protein. This is illustrated in Figure 5.1.23.

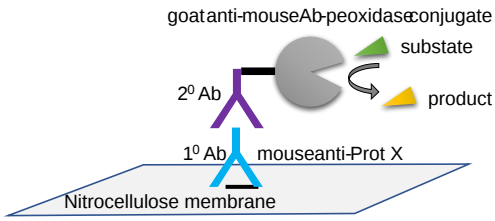


Figure 5.1.23: Developing a Western blot

3D electrophoresis: To detect specific proteins in a 2D electrophoresis experiment, a 3rd dimension of separation —Western blotting —could be performed on the PAGE gel, and the nitrocellulose membrane stained with an antibody specific to the target protein. That is illustrated in Figure 5.1.24.

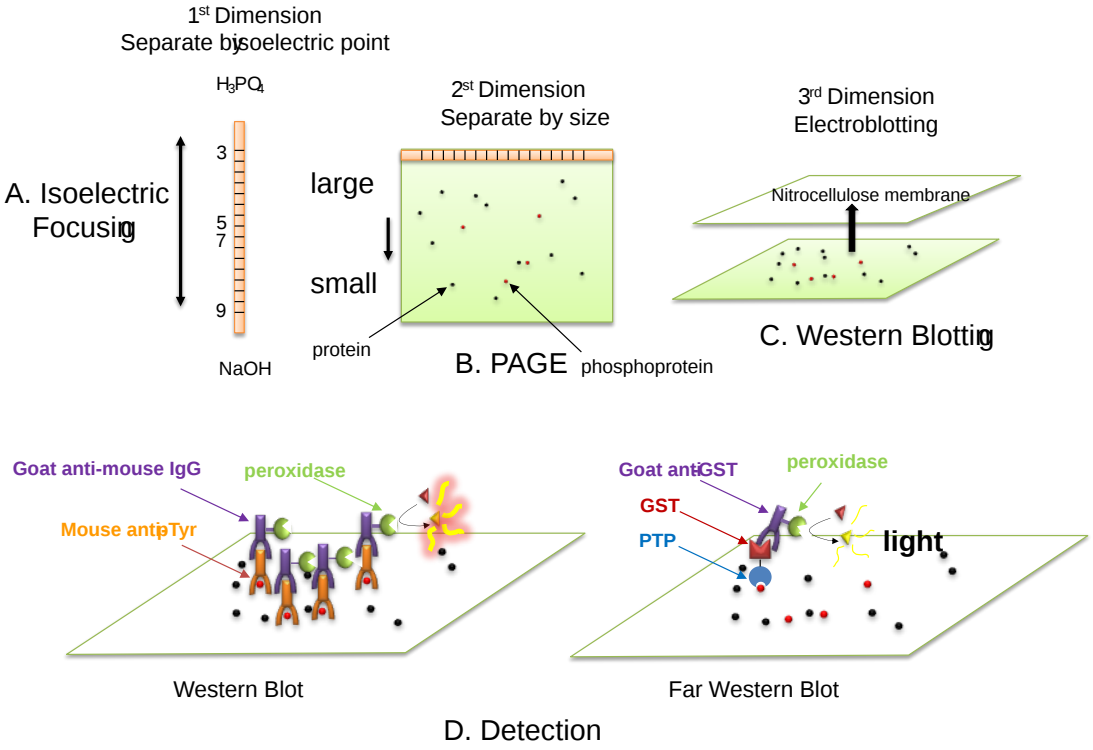


Figure 5.1.24: "3D" electrophoresis - Isoelectric focusing, PAGE and Western blotting

Part A, isoelectric focusing, is followed by a PAGE gel (B). The red dots represent proteins that have undergone a post-translational modification in which a phosphate group has been added to tyrosine side chains (for example). Western blotting is performed in panel C, and staining is performed in panel D. The left blot in D uses an antibody that recognizes phosphorylated tyrosine side chains on protein. The right blot in D is sometimes called a Far Western blot. If the protein on the nitrocellulose membrane retains some 3D native structure or can be induced to refold, it can be probed on the blot by a protein that binds to the native form of the protein on the blot. In the example shown in panel D above, the p-Tyr-protein target on the nitrocellulose membrane recognizes a PTP-GST fusion protein. GST is a protein tag for detection. PTP (a catalytically inactive form) is a protein tyrosine phosphatase, an enzyme that, in its active form, hydrolyzes p-Tyr on specific phosphorylated target proteins.

5.1.9: Summary

This chapter provides a comprehensive overview of protein purification and electrophoretic analysis, two essential methodologies in biochemistry for isolating and characterizing proteins. The chapter begins by emphasizing the necessity of protein purification, given the vast diversity of proteins in cells and tissues. It describes how inherent differences—such as size, charge, hydrophobicity, and binding affinity—are exploited to isolate a single protein from complex biological mixtures.

Sample Processing and Initial Fractionation:

The process starts with sample processing, where cells or tissues are disrupted using methods such as freeze/thaw cycles, sonication, or detergent treatment. Centrifugation is then used to separate soluble proteins from cellular debris. Special attention is given to preventing proteolysis by performing procedures at low temperatures and adding protease inhibitors.

Bulk Purification Techniques:

Precipitation using ammonium sulfate is introduced as an economical initial step to concentrate proteins while maintaining their native state, followed by dialysis to remove excess salt. The chapter then details various chromatographic techniques:

- **Size Exclusion Chromatography (Gel Filtration):** Separates proteins based on size and shape by exploiting differences in how they access the internal pore volume of the resin beads.

- **Ion Exchange Chromatography:** Separates proteins based on their net charge; proteins bind to charged resins and are eluted by altering the ionic strength or pH of the buffer.
- **Affinity Chromatography:** Uses specific ligand-receptor interactions (e.g., His-tag binding to Ni-NTA or antibody-antigen recognition) to achieve high selectivity in isolating the protein of interest.
- **Hydrophobic Interaction Chromatography (HIC):** Exploits differences in protein hydrophobicity by modulating salt concentrations to promote binding to nonpolar ligands on the resin. Advanced methods such as HPLC and FPLC are also discussed as ways to increase resolution and throughput in protein separation.

Quantitative Evaluation of Purification:
The chapter stresses the importance of tracking purification efficiency using quantitative metrics such as total protein, total activity, specific activity, yield, and fold purification. These measurements are crucial for optimizing purification protocols and ensuring the integrity of the target protein.

Electrophoretic Analysis:
The latter portion of the chapter focuses on electrophoresis, a technique that separates proteins based on their size, charge, and shape in an electric field. Both native and denaturing methods are covered:

- **SDS-PAGE:** By binding to proteins and imparting a uniform negative charge, SDS standardizes the electrophoretic mobility of proteins so that separation is primarily based on molecular weight.
- **Isoelectric Focusing (IEF):** Separates proteins based on their isoelectric points (pI) by establishing a pH gradient.
- **Two-Dimensional Electrophoresis (2-DE):** Combines IEF and SDS-PAGE to achieve high-resolution separation of complex protein mixtures. Detection methods, including Coomassie Blue and silver staining, as well as Western blotting and autoradiography for enhanced sensitivity and specificity, are also described.

This chapter integrates the theoretical and practical aspects of protein purification and analysis. It highlights the sequential nature of the purification process—from cell lysis to chromatography and electrophoretic separation—and underscores the importance of quantitative evaluation at each step to ensure high purity and activity of the isolated protein. This foundation is critical for subsequent structural, functional, and proteomic studies in biochemistry.

5.1.10: References

Molnar, C. and Gair, J. (2013) Antibodies. Chapter in Concepts in Biology, Published by B.C. Open Textbook Project. Available at: https://opentextbc.ca/biology/chapter/23-3-antibodies/

The Human Atlas Project. (2019) Methods. Available at: https://www.proteinatlas.org/learn/method

Uhlén M et al, 2015. **Tissue-based map of the human proteome.** *Science*
PubMed: 25613900 DOI: 10.1126/science.1260419

Thul PJ et al, 2017. **A subcellular map of the human proteome.** *Science.*
PubMed: 28495876 DOI: 10.1126/science.aal3321

Uhlen M et al, 2017. **A pathology atlas of the human cancer transcriptome.** *Science.*
PubMed: 28818916 DOI: 10.1126/science.aan2507

Ahern, K. and Rajagopal, I. (2019) Biochemistry Free and Easy. Published by Libretexts. Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_and_Easy_(Ahern_and_Rajagopal)/09%3A_Techniques/9.04%3A_Gel_Exclusion_Chromatography.

Magdeldin, S. (2012) Gel Electrophoresis - Principles and Basics. Published by InTech under Creative Commons Attribution 3.0. Available at: https://pdfs.semanticscholar.org/4b93/70ac3946cec6e12c369679c4178a5ef38e61.pdf

Structural Biochemistry/Proteins/X-ray Crystallography. (2018, November 19). *Wikibooks, The Free Textbook Project*. Retrieved 15:40, August 17, 2019 from en.wikibooks.org/w/index.php?title=Structural_Biochemistry/Proteins/X-ray_Crystallography&oldid=3488057.

UCD: Biophysics 200A (2019) "NMR Spectroscopy vs X-ray Crystallography", Chapter published in Current Techniques in Biophysics. Published by Libretexts and available at:

https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_200A_-_Current_Techniques_in_Biophysics/NMR_Spectroscopy_vs._X-ray_Crystallography

Wikipedia contributors. (2019, June 27). Protein purification. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 23:32, July 28, 2019, from en.Wikipedia.org/w/index.php?title=Protein_purification&oldid=903657925

Wikipedia contributors. (2019, February 15). Fast protein liquid chromatography. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 17:14, August 15, 2019, from en.Wikipedia.org/w/index.php?title=Fast_protein_liquid_chromatography&oldid=883530035

Wikipedia contributors. (2019, July 9). Protein mass spectrometry. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 15:27, August 16, 2019, from en.Wikipedia.org/w/index.php?title=Protein_mass_spectrometry&oldid=905547289

Wikipedia contributors. (2019, July 8). Peptide synthesis. In *Wikipedia*, *The Free Encyclopedia*. Retrieved 06:13, August 17, 2019, from en.Wikipedia.org/w/index.php?title=Peptide_synthesis&oldid=905401648

This page titled 5.1: Protein Purification is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

5.2: Cell Disruption

There are several ways to break open cells.

- Lysis methods include lowering the ionic strength of the medium cells are kept in. This can cause cells to swell and burst. Mild surfactants may be used to enhance the efficiency of lysis. Most bacteria, yeast, and plant tissues, which have cell walls, are resistant to such osmotic shocks, however, and stronger disruption techniques are often required.
- Enzymes may be useful in helping to degrade the cell walls. Lysozyme, for example, is very useful for breaking down bacterial walls. Other enzymes commonly employed include cellulase (plants), glycanases, proteases, mannases, and others.
- Mechanical agitation may be employed in the form of tiny beads that are shaken with a suspension of cells. As the beads bombard the cells at high speed, they break them open. *Sonication* (20-50 kHz sound waves) provides an alternative method for lysing cells. The method is noisy, however, and generates heat that can be problematic for heat-sensitive compounds.
- Another means of disrupting cells involves using a “cell bomb". In this method, cells are placed under very high pressure (up to 25,000 psi). When the pressure is released, the rapid pressure change causes dissolved gases in cells to be released as bubbles which, in turn, break open the cells.
- *Cryopulverization* is often employed for samples having a tough extracellular matrix, such as connective tissue or seeds. In this technique, tissues are .ash-frozen using liquid nitrogen and then ground to a fine powder before extraction of cell contents with a buffer.

Whatever method is employed, the crude lysates obtained contain all of the molecules in the cell, and thus, must be further processed to separate the molecules into smaller subsets, or fractions.

Contributors and Attributions

- Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

This page titled 5.2: Cell Disruption is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.1: Cell Disruption** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.
- **5.1: Cell Disruption** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.3: Fractionation

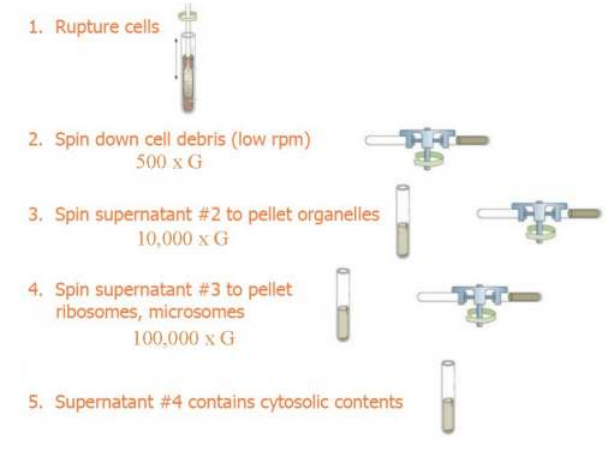


Figure 9.2.1: *Fractionation by centrifugation*

Many chromatographic techniques are performed in “columns." These are tubes containing the material (called the “support") used to perform the separation . Supports are designed to exploit the chemical, or size, differences of the many molecules in a mixture. Columns are “packed" (filled) with the support and a buffer or solvent carries the mixture of compounds to be separated through the support. Molecules in the sample interact differentially with the support and consequently, will travel through it with different speeds.

Contributors and Attributions

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

This page titled 5.3: Fractionation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.2: Fractionation** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.
- **5.2: Fractionation** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.4: Electrophoresis

Electrophoresis uses an electric field applied across a gel matrix to separate large molecules such as DNA, RNA, and proteins by charge and size. Samples are loaded into the wells of a gel matrix that can separate molecules by size and an electrical field is applied across the gel. This field causes negatively charged molecules to move towards the positive electrode. The gel matrix, itself, acts as a sieve, through which the smallest molecules pass rapidly, while longer molecules are slower-moving.

For DNA and RNA, sorting molecules by size in this way is trivial, because of the uniform negative charge on the phosphate backbone. For proteins, which vary in their charges, a clever trick must be employed to make them mimic nucleic acids - see polyacrylamide gel electrophoresis (PAGE) below. Different kinds of gels have different pore sizes. Like sieves with finer or coarser meshes, some gels do a better job of separating smaller molecules while others work better for larger ones. Gel electrophoresis may be used as a preparative technique (that is, when purifying proteins or nucleic acids), but most often it is used as an analytical tool.

Agarose Gel Electrophoresis

Agarose gel electrophoresis is a technique used to separate nucleic acids primarily by size. Agarose is a polysaccharide obtained from seaweeds (Figure 8.11). It can be dissolved in boiling buffer and poured into a tray, where it sets up as it cools (Figure 8.12) to form a slab. Agarose gels are poured with a comb in place to make wells into which DNA or RNA samples are placed after the gel has solidified. The gel is immersed in a buffer and a current is applied across the slab. Double-stranded DNA has a uniform negative charge that is independent of the sequence composition of the molecule. Therefore, if DNA fragments are placed in an electric field they will migrate from the cathode (-) towards the anode (+). The rate of migration is directly dependent on the ability of each DNA molecule to worm or wiggle its way through the sieving gel. The agarose matrix provides openings for macromolecules to move through. The largest macromolecules have the most difficult time navigating through the gel, whereas the smallest macromolecules slip through it the fastest.

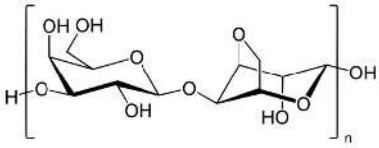


Figure 8.11 - Structure of the agarose polysaccharide. Wikipedia

Because electrophoresis uses an electric current as a force to drive the molecules through the matrix, the molecules being separated must be charged. Since the size to charge ratio for DNA and RNA is constant for all sizes of these nucleic acids, the molecules simply sort on the basis of their size - the smallest move fastest and the largest move slowest.

All fragments of a given size will migrate the same distance on the gel, forming the so-called “bands” on the gel. Visualization of the DNA fragments in the gel is made possible by addition of a dye, such as ethidium bromide, which intercalates between the bases and fluoresces when viewed under ultraviolet light (Figure 8.13) By running reference DNAs of known sizes alongside the samples, it is possible to determine the sizes of the DNA fragments in the sample. It is useful to note that, by convention, DNA fragments are not described by their molecular weights (unlike proteins), but by their length in base-pairs(bp) or kilobases (kb).

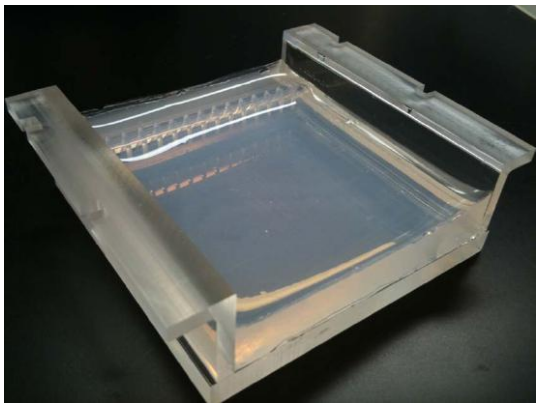


Figure 8.12 - Agarose gel electrophoresis separation of DNA - orange bands are DNA fragments. Wikipedia

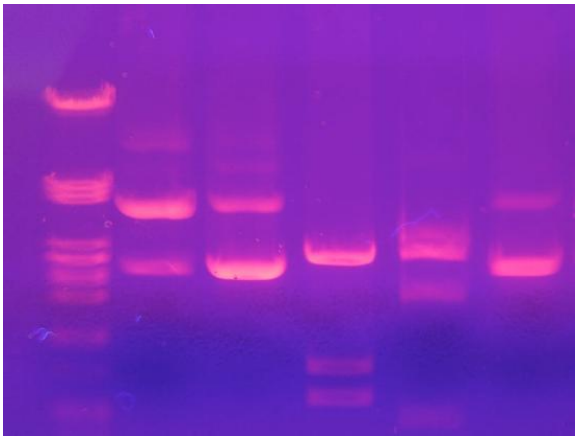


Figure 8.13 - DNA bands visualized with ethidium bromide staining. Wikipedia

Polyacrylamide gel electrophoresis (PAGE)

Like DNA and RNA, proteins are large macromolecules, but unlike nucleic acids, proteins are not necessarily negatively charged. The charge on each protein depends on its unique amino acid sequence. Thus, the proteins in a mixture will not necessarily all move towards the anode.

Additionally, whereas double-stranded DNA is rod-shaped, most proteins are globular (folded). Further, proteins are considerably smaller than nucleic acids, so the openings of the matrix of the agarose gel are simply too large to effectively provide separation. Consequently, unaltered (native) proteins are not very good prospects for electrophoresis on agarose gels. To separate proteins by mass using electrophoresis, one must make several modifications.

Gel matrix

First, a matrix made by polymerizing and cross-linking acrylamide units is employed. A monomeric acrylamide (Figure 8.14) is polymerized and the polymers are cross-linked using N,N'-Methylene-bisacrylamide (Figure 8.15) to create a mesh-like structure. One can adjust the size of the openings of the matrix/mesh readily by changing the percentage of acrylamide in the reaction. Higher percentages of acrylamide give smaller openings and are more effective for separating smaller molecules, whereas lower percentages of acrylamide are used when resolving mixtures of larger molecules. (Note: polyacrylamide gels are also used to separate small nucleic acid fragments, with some acrylamide gels capable of separating pieces of DNA that differ in length by just one nucleotide.)

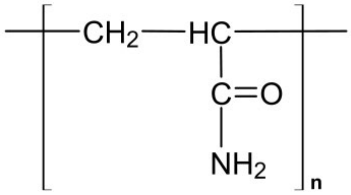


Figure 8.14 - Acrylamide monomer. Wikipedia

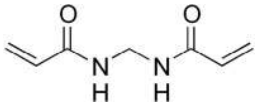


Figure 8.15 - N,N'-Methylenebisacrylamide - acrylamide crosslinking reagent. Wikipedia

Charge alteration by SDS

A second consideration is that proteins must be physically altered to "present" themselves to the matrix like the negatively charged rods of DNA. This is accomplished by treating the proteins with the anionic detergent, SDS (sodium dodecyl sulfate). SDS denatures the proteins so they assume a rod-like shape and the SDS molecules coat the proteins such that the exterior surface is loaded with negative charges, masking the original charges on the proteins and making the charge on the proteins more proportional to their mass, like the backbone of DNA.

Since proteins typically have disulfide bonds that prevent them from completely unfolding in detergent, samples are boiled with mercaptoethanol to break the disulfide bonds and ensure the proteins are as rod-like as possible in the SDS. Reagents like mercaptoethanol (and also dithiothreitol) are sulfhydryl-containing reagents that become oxidized as they reduce disulfide bonds in other molecules (see Figure 8.16)

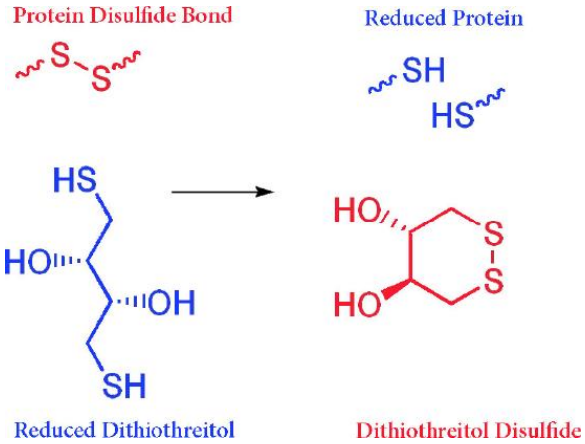


Figure 8.16 - Reduction of disulfide bonds by dithiothreitol. Wikipedia

Stacking Gel

A third consideration is that a "stacking gel" may be employed at the top of a polyacrylamide gel to provide a way of compressing the samples into a tight band before they enter the main polyacrylamide gel (called the resolving gel). Just like DNA fragments in agarose gel electrophoresis get sorted on the basis of size (largest move slowest and smallest move fastest), the proteins migrate through the gel matrix at velocities inversely related to their size. Upon completion of the electrophoresis, proteins may be visualized by staining with compounds that bind to proteins, like Coomassie Brilliant Blue (Figure 8.17) or silver nitrate.

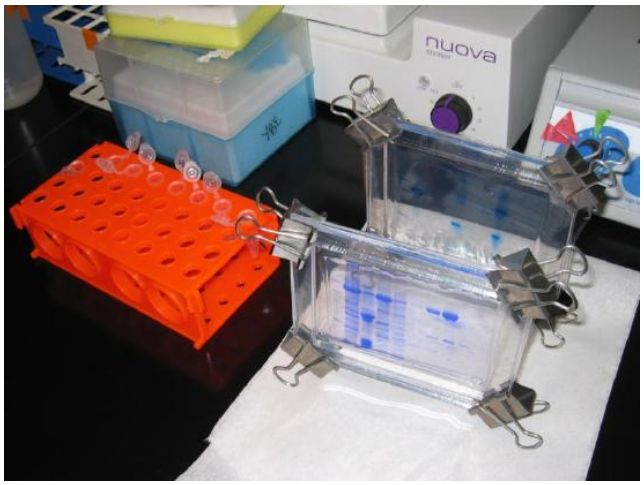


Figure 8.17 - Two SDS-PAGE gels - Proteins are the blue bands (stained with Coomassie Blue). Wikipedia

Non-denaturing gel electrophoresis

The SDS_PAGE technique described above is the commonest method used for electrophoretic separation of proteins. In some situations, however, proteins may be resolved on so-called “native” gels, in the absence of SDS. Under these conditions, the movement of proteins through the gel will be affected not simply by their mass, but by their charge at the pH of the gel, as well. Proteins complexed with other molecules may move as single entity, allowing the isolation of the binding partners of proteins of interest.

Isoelectric focusing

Proteins vary considerably in their charges and, consequently, in their pI values (pH at which their charge is zero). This can be exploited to separate proteins in a mixture. Separating proteins by isoelectric focusing requires establishment of a pH gradient in a tube containing an acrylamide gel matrix. The pore size of the gel is adjusted to be large, to reduce the effect of sieving based on size. Molecules to be separated are applied to the gel containing the pH gradient and an electric field is applied. Under these conditions, proteins will move according to their charge.

Positively charged molecules, for example, move towards the negative electrode, but since they are traveling through a pH gradient, as they pass through it, they reach a region where their charge is zero and, at that point, they stop moving. They are at that point attracted to neither the positive nor the negative electrode and are thus “focused” at their pI (Figure 8.18). Using isoelectric focusing, it is possible to separate proteins whose pI values differ by as little as 0.01 units.

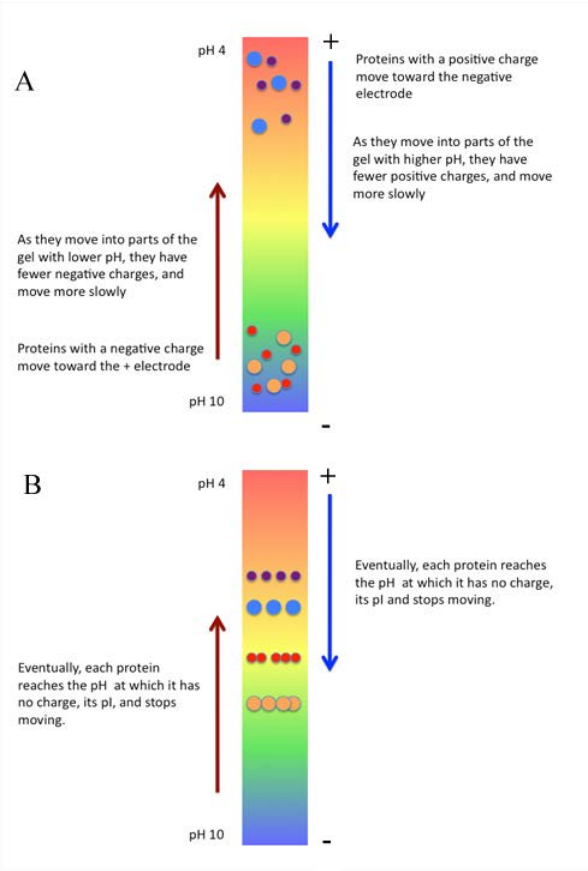


Figure 8.18 - Isoelectric focusing: A. At the start of the run; B. at the end of the run

2D gel electrophoresis

Both SDS-PAGE and isoelectric focusing are powerful techniques, but a clever combination of the two is a powerful tool of proteomics - the science of studying all of the proteins of a cell/tissue simultaneously. In 2-D gel electrophoresis, a lysate is first prepared from the cells of interest. The proteins in the lysate are separated first by their pI, through isoelectric focusing and then by size by SDS-PAGE.

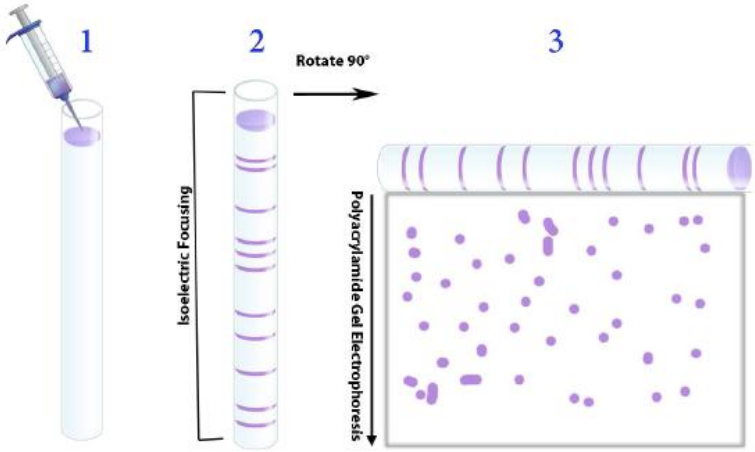


Figure 8.19 - Scheme for performing 2-D gel analysis. Image by Aleia Kim

The mixture of proteins is first applied to a tube or strip (Figure 8.19, Step 1) where isoelectric focusing is performed to separate the proteins by their pI values (Step 2). Next, as shown in the figure, the gel containing the proteins separated by their pIs is turned on its side and applied along the top of a polyacrylamide slab for SDS-PAGE to separate on the basis of size (Step 3). The proteins

in the isoelectric focusing matrix are electrophoresed into the polyacrylamide gel and separated on the basis of size. The product of this analysis is a 2-D gel as shown in Figure 8.20.The power of 2-D gel electrophoresis is that virtually every protein in a cell can be separated and appear on the gel as a spot defined by its unique size and pI. In the figure, spots in the upper left correspond to large positively charged proteins, whereas those in the lower right are small negatively charged ones. Every spot on a 2-D gel can be eluted and identified by using high throughput mass spectrometry. This is particularly powerful when one compares protein profiles between different tissues or between control and treated samples of the same tissue.

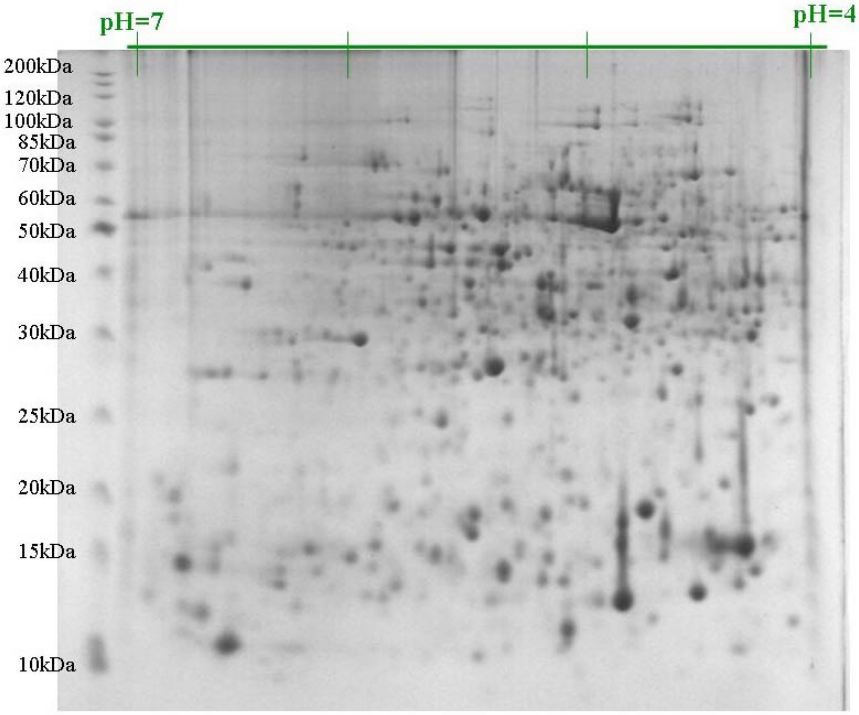


Figure 8.20 - Result of 2-D gel electrophoresis separation. Wikipedia

Protein profiles comparison

Comparison of 2-D gels of proteins from non-cancerous tissue and proteins from a cancerous tissue of the same type provides a quick identification of proteins whose level of expression differs between the two. Information such as this can be useful in designing treatments or in understanding the mechanism(s) by which the cancer develops.

This page titled 5.4: Electrophoresis is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **8.3: Electrophoresis** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **5.4: Electrophoresis** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

5.5: Electrophoresis

DNA molecules are long and loaded with negative charges, thanks to their phosphate backbones. Electrophoretic methods separate large molecules, such as DNA, RNA, and proteins based on their charge and size. For DNA and RNA, the charge of the nucleic acid is proportional to its size (length). For proteins, which do not have a uniform charge, a clever trick is employed to make them mimic nucleic acids.

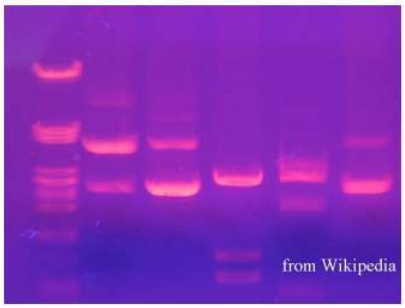


Figure 9.8.1: *DNA fragments separated by agarose gel electrophoresis*

Agarose

Agarose gel electrophoresis is a method for separating nucleic acids. It is worth noting that nucleic acids are the largest molecules found in cells, in some cases by orders of magnitude. Agarose provides a matrix which encases a buffer. The matrix provides openings for macromolecules to move through and the largest macromolecules have the most difficult time navigating, whereas the smallest macromolecules slip through it the easiest. Unlike column chromatography, electrophoresis uses an electric current as a force to drive the molecules through the matrix. Since the size to charge ratios for DNA and RNA are constant for all sizes of these nucleic acids, the size per force is also constant (since force is directly proportional to charge), so the molecules simply sort on the basis of their size - the smallest move fastest and the largest move slowest. Visualization of the DNA fragments in the gel is made possible by addition of a dye, such as ethidium bromide that fluoresces under ultraviolet light.

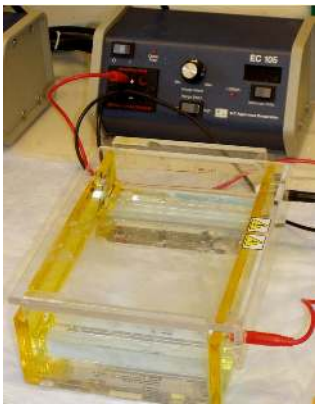


Figure 9.8.2: *Agarose gel electrophoresis*

SDS-PAGE

Like DNA and RNA, proteins are large macromolecules. Proteins, however, vary tremendously in their charge. Whereas double-stranded DNA is rod-shaped, most proteins are globular (folded up). Further, proteins are considerably smaller than nucleic acids, so the openings of the matrix of the agarose gel are simply too large to effectively provide separation. Consequently, unlike nucleic acids, proteins cannot be effectively separated by electrophoresis on agarose gels. To separate proteins by electrophoresis, one must make several modifications. First, a matrix made by polymerizing and crosslinking acrylamide units is employed. One can adjust the pore size of the matrix readily by changing the percentage of acrylamide in the gel. Higher percentages of acrylamide create smaller pores and are more effective in separating smaller molecules, whereas lower percentages of acrylamide reverse that. Second, proteins must be physically altered to “present" themselves to the matrix like the negatively charged rods of DNA. This is accomplished by treating the proteins with the detergent called SDS (sodium dodecyl sulfate). SDS denatures the proteins so they

assume a rod-like shape and the SDS molecules coat the proteins such that the exterior surface is loaded with negative charges proportional to the mass, just like the backbone of DNA. Third, a “stacking gel" may be employed at the top of the gel to provide a way of compressing the samples into a tight band before they enter the main polyacrylamide gel (called the resolving gel). Just as DNA fragments get sorted on the basis of size (largest move slowest and smallest move fastest), the proteins migrate through the gel matrix at rates inversely related to their size. Upon completion of the electrophoresis, there are several means of staining to visualize the proteins on the gel. They include reagents, such as Coomassie Brilliant Blue or silver nitrate (the latter is much more sensitive than Coomassie Blue staining and can be used when there are very small quantities of protein).

Figure 9.8.3: *SDS-Page Separation of proteins*

Isoelectric Focusing

Proteins vary considerably in their charges and, consequently, in their pI values (pH at which their charge is zero). Separating proteins by isoelectric focusing requires establishment of a pH gradient in an acrylamide gel matrix. The matrix’s pores are adjusted to be large to reduce the effect of sieving based on size. Molecules to be focused are applied to the gel with the pH gradient and an electric current is passed through it. Positively charged molecules, for example, move towards the negative electrode, but since they are traveling through a pH gradient, as they pass through it, they reach a region where their charge is zero and, at that point, they stop moving. They are at that point attracted to neither the positive nor the negative electrode and are thus “focused" at their pI. By using isoelectric focusing, it is possible to separate proteins whose pI values differ by as little as 0.01 units.

2-D Gel Electrophoresis

Both SDS-PAGE and isoelectric focusing are powerful techniques, but a clever combination of the two is a powerful tool of proteomics - the science of studying all of the proteins of a cell/tissue simultaneously. In 2D gel electrophoresis, an extract containing the proteins is first prepared. One might, for example, be studying the proteins of liver tissue. The liver cells are lysed and all of the proteins are collected into a sample. Next, the sample is subjected to isoelectric focusing as described earlier, to separate the proteins by their pI values. Next, as shown on the previous page, the isoelectric gel containing the separated proteins is rotated through 90º and placed on top of a regular polyacrylamide gel for SDS-PAGE analysis (to separate them based on size). The proteins in the isoelectric gel matrix are electrophoresed into the polyacrylamide gel and separation on the basis of size is performed. The product of this analysis is a 2D gel, in which proteins are sorted by both mass and charge.

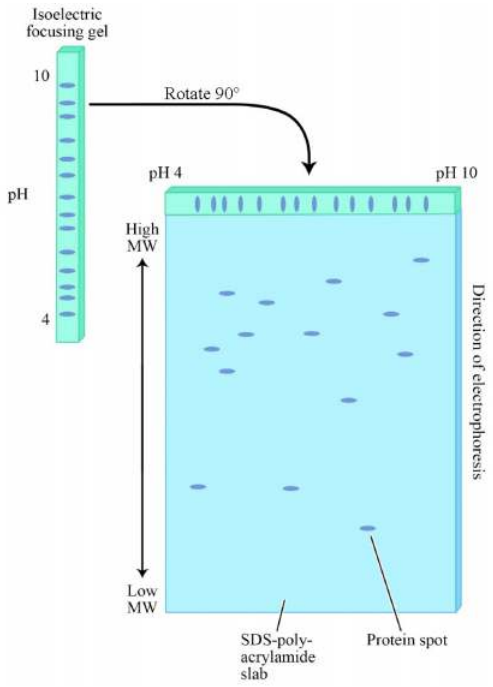


Figure 9.8.4: *2D Gel Electrophoresis*

The power of 2D gel electrophoresis is that virtually every protein in a cell can be separated and appear on the gel as a distinct spot. In the figure, spots in the upper left correspond to large positively charged proteins, whereas those in the lower right are small negatively charged ones. It is possible using high- throughput mass spectrometry analysis to identify every spot on a 2D gel. This is particularly powerful when one compares protein profiles between different tissues or between the samples of the same tissue treated or untreated with a particular drug. Comparison of a 2D separation of a non-cancerous tissue with a cancerous tissue of the same type provides a quick identification of proteins whose level of expression differs between them. Information such as this might be useful in designing treatments or in determining the mechanisms by which the cancer arose.

This page titled 5.5: Electrophoresis is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.8: Electrophoresis** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.6: Gel Exclusion Chromatography

Gel exclusion chromatography (also called molecular exclusion chromatography, size exclusion chromatography, or gel filtration chromatography) is a low resolution isolation method that employs a cool “trick." This involves the use of beads that have tiny “tunnels" in them that each have a precise size. The size is referred to as an “exclusion limit," which means that molecules above a certain molecular weight will not fit into the tunnels. Molecules with sizes larger than the exclusion limit do not enter the tunnels and pass through the column relatively quickly by making their way between the beads. Smaller molecules, which can enter the tunnels, do so, and thus, have a longer path that they take in passing through the column. Because of this, molecules larger than the exclusion limit will leave the column earlier, while those that pass through the beads will elute from the column later. This method allows separation of molecules by their size.

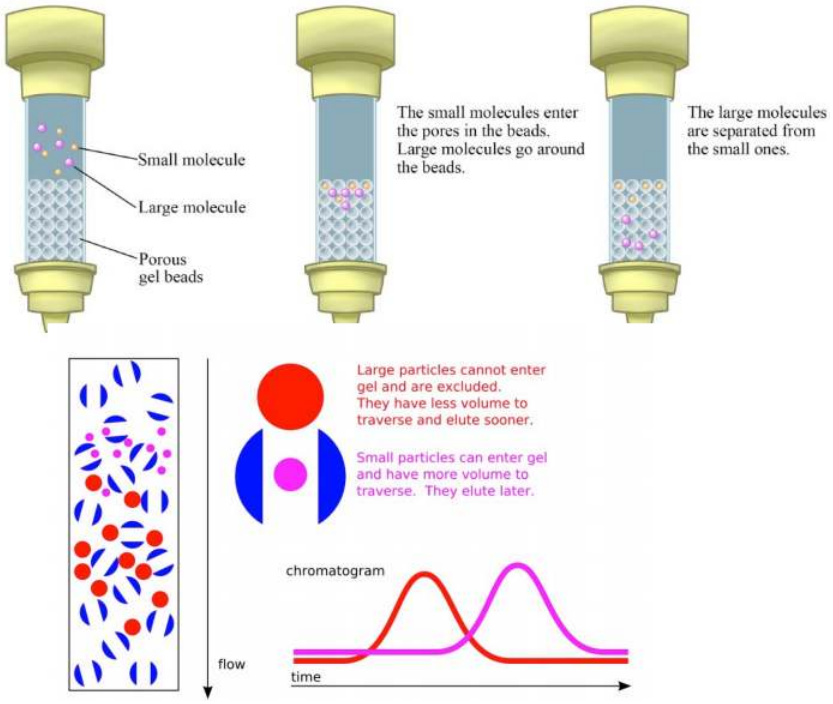


Figure 9.4.1: *Gel Exclusion Chromatography*

This page titled 5.6: Gel Exclusion Chromatography is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.4: Gel Exclusion Chromatography** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.7: Ion Exchange Chromatography

In ion exchange chromatography, the support consists of tiny beads to which are attached chemicals possessing a charge. Each charged molecule has a counter-ion. The figure shows the beads (blue) with negatively charged groups (red) attached. In this example, the counter-ion is sodium, which is positively charged. The negatively charged groups are unable to leave the beads, due to their covalent attachment, but the counter- ions can be “exchanged" for molecules of the same charge. Thus, a cation exchange column will have positively charged counter-ions and positively charged compounds present in a mixture passed through the column will exchange with the counter-ions and “stick" to the negatively charged groups on the beads. Molecules in the sample that are neutral or negatively charged will pass quickly through the column. On the other hand, in anion exchange chromatography, the chemical groups attached to the beads are positively charged and the counter-ions are negatively charged. Molecules in the sample that are negatively charged will “stick" and other molecules will pass through quickly. To remove the molecules “stuck" to a column, one simply needs to add a high concentration of the appropriate counter-ions to displace and release them. This method allows the recovery of all components of the mixture that share the same charge.

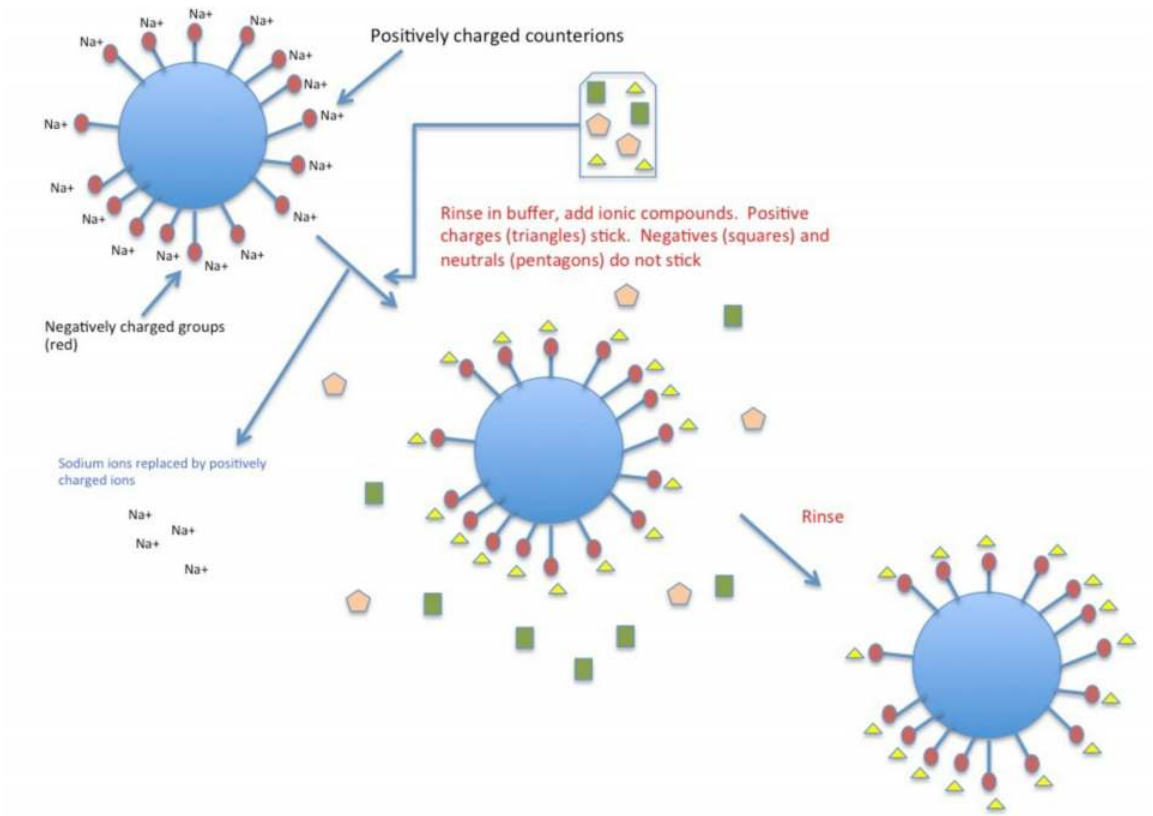


Figure 9.3.1: *Cation exchange chromatography*

This page titled 5.7: Ion Exchange Chromatography is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.3: Ion Exchange Chromatography** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.8: Affinity Chromatography

Affinity chromatography is a very powerful technique that exploits the binding affinities of target molecules (typically proteins) for substances covalently linked to beads. For example, if one wanted to separate all of the proteins in a sample that bound to ATP from proteins that do not bind ATP, one could covalently link ATP to support beads and then pass the sample through column. All proteins that bind ATP will “stick" to the column, whereas those that do not bind ATP will pass quickly through it. The proteins adhering to the column may then released from the column by adding ATP.

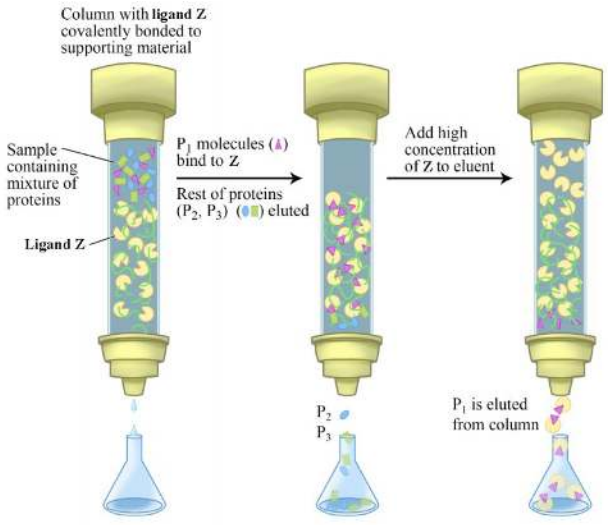


Figure 9.5.1: *Affinity Chromatography*

This page titled 5.8: Affinity Chromatography is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.5: Affinity Chromatography** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.9: High Performance Liquid Chromatography (HPLC)

HPLC (also sometimes called High Pressure Liquid Chromatography) is a powerful tool for separating smaller molecules based on their differential polarities. It employs columns with supports made of very tiny beads that are so tightly packed that .ow of solvents/buffers through the columns requires the application of high pressures (hence the name). The supports used can be polar (normal phase separation) or non-polar (reverse phase separation). In normal phase separations, non-polar molecules elute first followed by the more polar compounds. This order is switched in reverse phase chromatography. Of the two, reverse phase is much more commonly employed to due more reproducible chromatographic profiles (separations) that it typically produces.

This page titled 5.9: High Performance Liquid Chromatography (HPLC) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.6: High Performance Liquid Chromatography (HPLC)** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.10: Histidine Tagging

Histidine tagging is a powerful tool for isolating a recombinant protein from a cell lysate. It relies on using recombinant DNA techniques to add codons specifying a series of histidines (usually six) to the coding sequence for a protein. The protein produced when this gene is expressed has a run of histidine residues fused at either the carboxyl or amino terminus to the amino acids in the remainder of the protein. The histidine side chains of this “tag" have an affinity for nickel or cobalt ions, making separation of histidine tagged proteins from a cell lysate is relatively easy. Simply passing the sample through a column that has immobilized nickel or cobalt ions allows the histidine- tagged proteins to “stick," while the remaining cell proteins all pass quickly through. The histidine-tagged proteins are then eluted by addition of imidazole (which is chemically identical to the histidine side chain) to the column. Histidine tags can be cleaved off using endopeptidases.

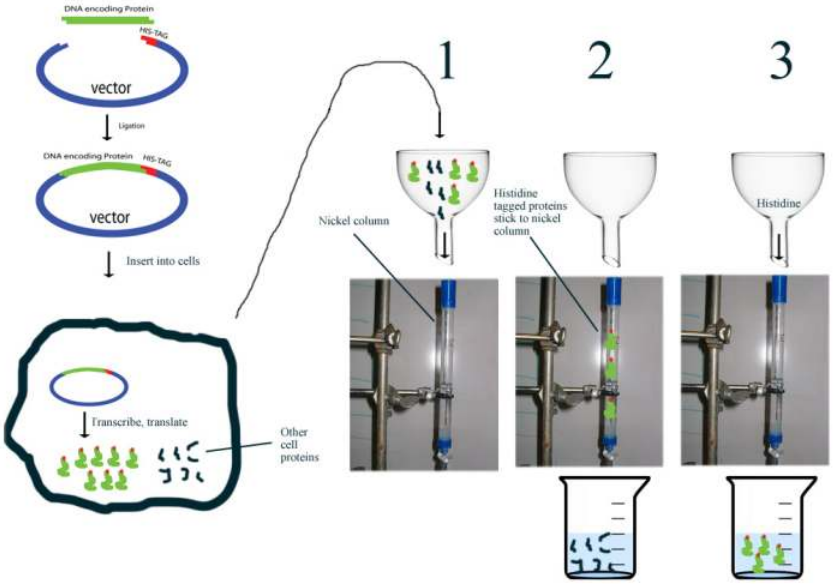


Figure 9.7.1: *Histidine Tagging*

This page titled 5.10: Histidine Tagging is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **9.7: Histidine Tagging** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.11: Quantification of Protein Concentration

Review: the Beer-Lambert Law

log(1/T) = ε*cl* = A
T = 1/10^{A}

Values for ε and *l*

- Path length ***l*** is usually in units of **cm**. (note: most spectrophotometers are designed to accept 1cm wide cuvettes)
- Molar extinction coefficient **ε** has units of M^{-1} cm^{-1} and is a constant of proportionality that relates the absorption of **molar** solutions
- Mass extinction coefficient **ε $^{1\%}$** refers to the absorbance of a 1% by mass solution. Typically this refers to an aqueous solution that we can take to have a density of 1000g/L. A 1% by mass aqueous solution would therefore refer to the dissolution of 10g/L, or a **10mg/ml** solution of the molecule of interest.
- Since the absorbance of a molecule is a function of the wavelength (i.e. the absorption is not equal for every wavelength) the extinction coefficient must also reference a wavelength. This is typically done using a subscript:

$\varepsilon^{1\%}_{280nm}$ = 14.5 g^{-1} L cm^{-1}

· In this case **a 10mg/ml solution of the molecule will have an absorbance reading of 14.5** (dimensionless units) at l = 280nm (the absorption at other wavelengths may not be known). The units of concentration are g/L, thus e will have dimensions of g^{-1} L cm^{-1}.

Why is it important to be able to quantitate protein concentration in a sample?

An important application of "Biotechnology" is the production of proteins as commercial products. Such products might have pharmaceutical applications (e.g. insulin, human growth hormone, tissue plasminogen activator, erythropoietin, blood clotting factor VIII.), industrial applications (e.g. subtilisin (an enzyme in detergents), 2,5-diketo-D-gluconate reductase (an enzyme in vitamin C production), as materials (e.g. silk protein in textiles, barnacle adhesion protein as a glue). In these cases, there are various aspects of successful production that require quantitation:

- How much of the protein can be produced (i.e. what is the efficiency of production)?
- How pure is the protein that is produced (industrial applications may require 90% pure, pharmaceutical applications may require 99.999% pure)

Such proteins may be isolated from natural sources (e.g. blood clotting factor VIII may be extracted from human blood), or they may be produced recombinantly (e.g. E. coli bacterial cells can be genetically engineered to produce human growth hormone). In both cases, it may be necessary to purify the protein using a series of fractionation steps. We will go into more detail about such fractionation steps in a later lecture, but the general idea is that a heterogeneous mixture of molecules can be fractionated based upon some physical property of the molecules. The following are properties that can be used to fractionate a heterogeneous mixture of biomolecules:

- Molecular mass (i.e. "big" molecules can be separated from "small" molecules)
- pKa (i.e. "acidic" molecules can be separated from "basic" molecules)
- Hydrophobicity (i.e. non-polar molecules can be separated from polar molecules)

For such fractionation steps involving proteins, we need to keep track of how much of the contaminating proteins went into one fraction and how much of our desired protein went into the other fraction. Although the details are somewhat more complicated than this simple description, **it is important to be able to quantitate protein concentration to be able to effectively purify a protein of interest**.

Once a protein is pure, it may be of considerable economic interest to be able to quantify the yield (and, therefore, be able to determine how much it cost to produce a given mass of protein). For example, the only source for human growth hormone (to treat small stature) used to be to extract it from human pituitary glands harvested from the brains of cadavers. Suffice it to say, this made the protein extremely expensive. Furthermore, the isolation from human tissues meant that the sample could also be potentially contaminated with human pathogens (hepatitis, CJD, AIDS, etc.). With the advent of genetic engineering, the production of human

growth hormone by bacterial cells (i.e. E. coli) meant that relative large quantities could be produced far cheaper (and with no threat of human pathogens).

Why not just weigh the protein?

- Most samples are typically quantities of milligrams or even micrograms, not grams, and thus, it is difficult to transfer and measure such small amounts
- Water is present in proteins, and it is extremely difficult to remove all the water (some water molecules hydrogen bond extremely tightly to proteins). Thus, the mass measurement would include some waters, and would increase the apparent mass of the protein

Absorbance spectra of biological molecules

Proteins

Proteins do not absorb in the visible wavelength unless they have a prosthetic group (e.g. Fe^{2+}) or an unnatural amino acid. However, the amino acids tryptophan, tyrosine and cysteine absorb light in the UV wavelength:

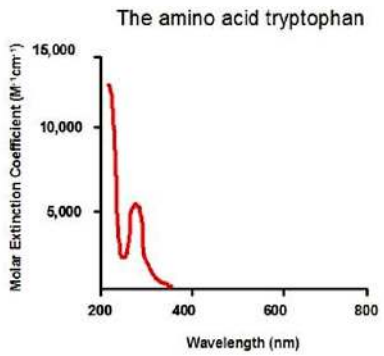


Figure 5.3.1: *Tryptophan absorption*

- Tryptophan has a peak of absorption at 280nm in the UV range
- This is a useful wavelength to quantitate the absorption of tryptophan
- Since the absorption is proportional to concentration, this is a useful way to quantitates protein concentration (for proteins containing Trp)

Nucleic acids

The aromatic rings in the bases of nucleic acids also absorb in the UV range:

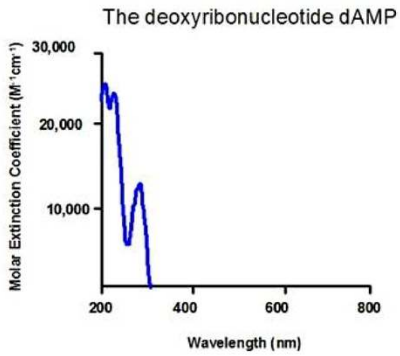


Figure 5.3.2: *dAMP Absorption*

- Each DNA and RNA base has a slightly different absorption spectrum
- 260 or 280nm is a typically useful wavelength to monitor concentration of nucleic acids

Note that samples of nucleic acids and proteins can both absorb at 280nm, therefore, samples of biological molecules should be **pure** in order to quantitate using UV absorption spectroscopy (any contaminating nucleic acids in a protein sample will increase the apparent absorbance, likewise for contaminating proteins in a nucleic acid sample).

Important aspects of quantification of proteins using UV absorbance

- If a protein contains Trp, Tyr or Cys residues it will absorb in the UV. If it does not contain these amino acids, it will not absorb UV light, and we cannot quantify it using this method
- Multiple Trp, Tyr or Cys residues will contribute to the Extinction coefficient for the protein. Thus, we need to know how many of these residues are present in the protein to know the correct extinction coefficient
- Nucleic acids (DNA, RNA) contaminant will also absorb UV light, as will other proteins with Trp, Tyr and Cys residues. Thus, the sample must be PURE to use UV absorption to quantify a protein

Molar extinction coefficients of Trp, Tyr and Cys amino acids:

Amino Acid	**E_{280nm} (M^{-1} cm^{-1})**
Trp	5690
Tyr	1280
Cys	120

Example 5.3.1: Bovine insulin

Bovine insulin contains 4 Tyr residues, 6 Cys residues and 0 Trp residues. We can determine the expected molar extinction coefficient at 280nm, E_{280nm}, by the following calculation:

E_{280nm} = (0)(5690) + (4)(1280) + (6)(120)

E_{280nm} = 5840 M^{-1} cm^{-1}

Thus, a 1.0M solution of pure bovine insulin would give an absorbance of 5,840 at 280nm (obviously, it would have to be diluted considerably to be read accurately).

A useful expression relating the parameters of E, concentration (C) and A are derived from the Beer-Lambert law (assuming 1cm path length):

A/E = C

For example, if a sample of bovine insulin was observed to give an absorbance at 280nm of 0.745 we could calculate the concentration to be:

0.745/5840 M^{-1} cm^{-1} = C

C = 1.28 x 10^{-4}M (note: cm^{-1} drops out with 1cm pathlength)

It should be noted that a deuterium lamp is required for UV spectrophotometry, as well as quartz cuvettes (since glass absorbs UV light)

Colorimetric (chromogenic) methods of protein concentration determination

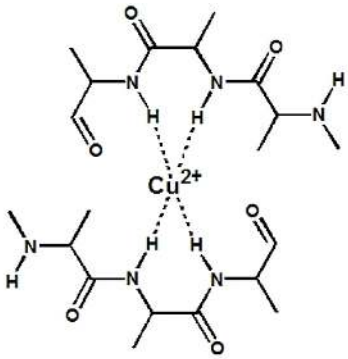
Chromo means color and genesis mean creation, so chromogenic means the "creation of color". Color means color (duh…) and metric means to measure, so colorimetric is to "measure color". Both terms refer to the same sort of thing in the present case - we can modify the protein sample with appropriate reagents so as to produce a color reaction (in visible spectrum) and measure protein concentration using a VIS spectrophotometer. Advantages are:

- Cheap lamp! (tungsten light bulb versus deuterium for UV)
- Cheap cuvette! (cheap glass or plastic versus quartz)
- Not contaminating absorbance from proteins or nucleic acids! (no absorption in VIS spectrum)

We will consider three methods: The **Biuret, Lowry and Bradford** methods of colorimetric determination of proteins.

Biuret

Under high pH (alkaline) conditions the copper II ion (Cu^{2+}) is believed to form a complex with peptide nitrogens of proteins:



Fiugre 5.3.3: *Cu^{2+} complex*

This complex absorbs light at 550nm and has the following useful properties:

- It is dependent upon at least a dipeptide structure (see above), thus, contaminating amino acids will not contribute to the 550nm absorption
- The binding depends upon the peptide backbone nitrogen and not the side chain functional group. Thus, the binding is independent of sequence
- Nucleic acids do not interfere since they don't share the peptide backbone structure
- However, ammonia and certain amines can interfere. Thus, ammonium sulfate salts, and the common biological buffer TRIS (tris hydroxymethyl amino ethane) will provide a false positive and cannot be present in the sample
- Also, the absorbtion is relatively weak, thus, the method is somewhat insensitive and requires a relatively high concentration of protein

Lowry

The Lowry method is a modification of the Biuret method. After treatment with Copper II, the protein is treated with phosphomolybdotungstate mixed acids (acidic compounds of molybdenum and tungsten ions). These acids are known as the **Folin-Ciocalteu (or just Folin) reagent**. The Folin reagent is added under alkaline conditions, and the Folin reagent is subsequently reduced by the the Copper ions as well as Tyr, Trp and polar amino acid side chains. The product of this reaction is **heteropolymolybdenum blue**, which absorbs strongly at 750nm. This assay has the following properties:

- More sensitive than the Biuret assay (can detect lower concentrations of protein)
- Somewhat dependent upon amino acid composition (i.e. relative concentrations of Tyr, Trp and polar amino acids)
- Absobtion reaction is linearly dependent upon protein concentration, but only at low concentrations of protein (i.e. the standard curve and assay must be performed at a low concentration regime).
- Sensitive to contaminants as with the Biuret method, as well as others related to the Folin reagent and redox reactions.
- More critical to timing and precision of person doing the assay

Bradford

A dye known as Coomassie Brilliant Blue was developed by the textile industry. It was noticed to stain skin as well as the textiles. Thus, this dye (which normally absorbs at 465nm) was known to bind to proteins and to absorb strongly at 595nm. The dye forms a wide variety of strong, but non-covalent, interactions including hydrogen bonding donor and acceptor interactions as well as hydrophobic (non-polar) interactions.

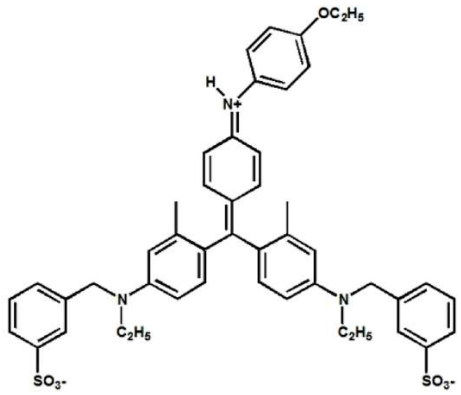


Figure 5.3.4: *Coomassie Brilliant Blue binding*

The method is quite simple: a single step in which the dye is added to the protein solution under acidic conditions, and then the absorbance is read at 595nm. However, the method has the following properties:

- The response is generally independent of the amino acid composition
- The assay is sensitive, but somewhat non-linear. Thus, a standard curve must always be performed (using known concentrations of pure protein)
- The dye stains pretty much everything, including cuvettes, floors, countertops. It can be a real mess if spilled (I know this from personal experience). And since it is a textile dye, if you get it on your clothes, you will need to learn to like blue polka dots.

This page titled 5.11: Quantification of Protein Concentration is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Michael Blaber.

- **5.3: Quantification of Protein Concentration** by Michael Blaber has no license indicated.
- **5.3: Quantification of Protein Concentration** by Michael Blaber has no license indicated.

5.12: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)

Learning Goals (ChatGPT o1, 1/30/25)

- **Fundamentals and Variants of ELISAs**
  - Explain the basic principles of enzyme-linked immunosorbent assays (ELISAs) and why they are widely used in biotechnology, clinical diagnostics, and pharmaceutical research.
  - Differentiate between traditional (direct/indirect) and sandwich ELISA formats, including how each type utilizes immobilized components (antigen or antibody) and the role of secondary labeled antibodies in detection.
- **Assay Design and Workflow**
  - Describe the step-by-step procedure of a traditional ELISA, including antigen immobilization, blocking, addition of primary antibody, washing steps, and introduction of the secondary labeled antibody.
  - Illustrate how competition assays work (e.g., competitive ELISA) and how the presence of analyte in a test sample affects the binding dynamics and signal output.
- **Detection Methods and Signal Generation**
  - Identify the different labels used in ELISAs (e.g., fluorophores, enzymes) and explain how their activity (via substrate conversion) is measured using spectrophotometric or fluorometric techniques.
  - Discuss the concept of signal amplification and how it enhances assay sensitivity.
- **Mathematical Modeling and Data Analysis**
  - Apply the Hill equation and the logistic (4-parameter) equation to describe and analyze ELISA standard curves, including the interpretation of parameters such as the Hill coefficient, L_{50} (or inflection point), minimum (a) and maximum (d) signal levels.
  - Evaluate how the slope of the dose–response curve (reflecting cooperativity or binding efficiency) influences data interpretation and quantitation of analyte concentration.
- **Quality Control and Assay Optimization**
  - Discuss key factors in ensuring ELISA reliability and validity, such as proper washing, blocking to minimize nonspecific binding, and the importance of including standard curves and appropriate controls.
  - Consider how variations in assay design (e.g., immobilization strategy, antibody pairing in sandwich ELISAs) can affect sensitivity, specificity, and dynamic range.
- **Lateral Flow Immunoassays (LFAs)**
  - Compare traditional ELISAs with lateral flow immunoassays, noting the differences in assay format (planar strip vs. microplate) and the role of capillary action in sample migration.
  - Explain how gold nanoparticles are used in LFAs for visual detection through surface plasmon resonance, including the underlying principles that enable observable color changes in positive tests.

By achieving these learning goals, students will be able to design, conduct, and critically evaluate immunoassays, and apply mathematical models to interpret experimental data, thereby bridging the gap between theoretical concepts and practical applications in biochemistry and related fields.

5.12.1: Introduction

Enzyme-linked immunosorbent assays (ELISAs) are used widely in biotechnology, pharmaceutical, and clinical medicine labs. At the same time, they appear to be underrepresented in chemistry and biochemistry curricula, even though their sensitivity, selectivity, and ease of use would argue for their widespread use.

ELISAs use primary antibodies specific to a target analyte (or antigen) as a central part of the assay. In a direct ELISA, the antigen is immobilized in wells of a multiwell plate or on a strip. An aqueous antibody solution is added, and an immobilized analyte

(antigen)-antibody complex forms. After washing the bound complex on the solid phase support to remove unbound species, a second labeled antibody is added for detection. This secondary antibody binds to the bound antibody's distal end (Fc domain) and not the analyte. The label on the second antibody can either be a fluorophore or an enzyme that can interact with an added substrate that will generate a colored solution. The color development is then measured with a fluorometer or spectrophotometer.

There are several variants of ELISAs, including the traditional ELISA, in which the antigen is bound or fixed to the surface of the solid support, or a sandwich ELISA, in which the antibody is bound to the surface. In the latter case, a second labeled antibody that binds to the antigen must bind at a different site (or epitope) on the antigen. For sandwich ELISAs, the target analyte must be large enough to accommodate two antibodies binding to different sites on the same molecule.

Cartoon diagrams showing the binding interactions in traditional and sandwich ELISAs are shown in Figure 5.12.1.

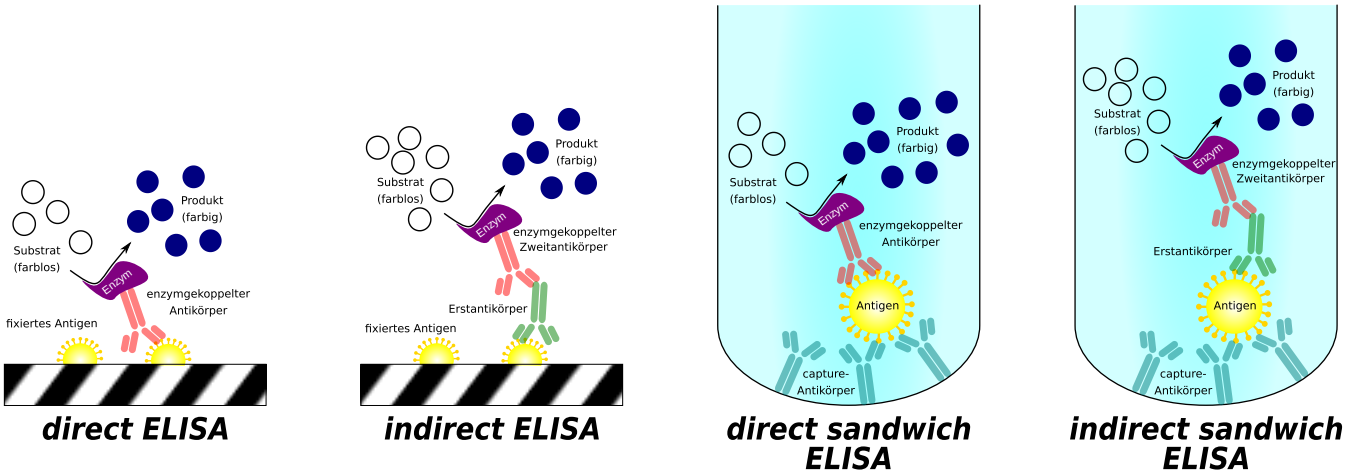


Figure 5.12.1: Binding interactions in traditional and sandwich ELISAs. Abbreviations are: fixiertes (fixed or immbolized), substrat (substrate), farblos (colorless), farbig (colored), enzymegekoppelter (enyzme coupled), antikörper (antibody), zweitantikörper (second antibody), erstantikörper (primary antibody). https://commons.wikimedia.org/wiki/File:ELISA.svg. Creative Commons Attribution 4.0 International license.

Note that both types have direct and indirect versions.

Figure 5.12.2 shows some of the steps in a traditional ELISA.

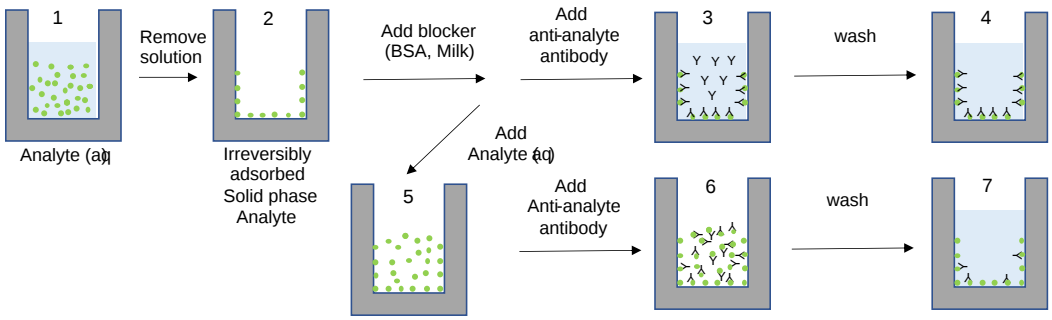


Figure 5.12.2: Steps in a traditional ELISA

In (1), an analyte is added to a well (outlined in gray). The solution is removed after a predefined time, and then a specific amount of analyte (such as protein) is irreversibly adsorbed (2). A blocker, such as bovine serum albumin (BSA) or milk, is added to bind any sites on the plate that could bind protein. A primary anti-analyte protein is added. Some bind to the immobilized protein, and the rest is free in solution. After washing, only the primary antibody-immobilized analyte remains.

Now, if a test solution (for example, from a patient's blood) is added after blocking but before adding the primary antibody, it remains in the well unbound (5). If a primary antibody is added to this well, the immobilized antigen competes with the free

antigen for binding, so less secondary antibody would be bound to the well (6). After washing, wells 4 and 7 remain. After additional rounds of washing and blocking, the secondary labeled antibody is added. Wells (such as 7) that contain patient analyte will bind less secondary antibody to the immobilized protein. When a substrate for the conjugated enzyme is added, the solution will be less colored after a defined incubation time.

A standard curve is made using a range of known analyte concentrations in assays. The more analytes in the standard, the greater the competition with the immobilized analyte for the solution-phase primary antibody. After washing away the solution-phase antigen:primary antibody complex, this would lead to a **lower** absorbance in the well with the higher solution-phase antigen in the sample.

Now, let's consider a sandwich assay. Instead of immobilizing a protein antigen, an antibody that binds the target antigen is immobilized. For example, an antibody against a surface SARS-CoV-2 protein can be immobilized. Next, a specimen (saliva, nasal swab) containing the target surface protein recognized by the immobilized antibody is added. The greater the viral load, the more SARS-CoV-2 binds to the antibody. Then, a second labeled antibody can be added that recognizes a different protein (the spike protein from SARS-CoV-2). After washing, the substrate is added, and the color development from the enzyme action on the substrate is measured. In this case, the more SARS-CoV-2 in the sample, the **higher** the signal (absorbance or fluorescence).

ELISAs have detection limits between 0.01 pg/mL and 100 ng/mL [1]. Although they are extensively used in health fields, they are not widely used in undergraduate biochemistry or chemistry courses, nor mentioned in the ACS's Guidelines and Supplements for either Analytical Chemistry or Biochemistry. Given their importance, we choose to discuss them here.

5.12.2: ELISA Data Analysis

The most difficult parts about ELISAs are understanding the chemical and mathematical equations, choosing/using modeling and analysis software, and evaluating validity/reliability. The typical data analysis is based on the generic Hill equation instead of the classical hyperbolic binding curve analysis we derived for a single ligand to a single site on a macromolecule. The Hill equation has more empirical parameters to use when fitting bind curves.

Here is the Hill equation that we studied before.

$$Y = \frac{L^n}{K_D + L^n} \tag{5.12.1}$$

For more complicated ELISA data, when a standard curve of known concentrations is used, and output signals (fluorescence, absorbance) vary from some minimum to a maximum value, a similar but more empirically useful **Logistic Equation** is used:

$$Y' = M\left(\frac{x^n}{k + x^n}\right) \tag{5.12.2}$$

where Y is the observable signal, and M is the maximal observable signal. The maximal signal might not be observed in an ELISA, as in binding a ligand to a macromolecule, when ligand concentrations >>K_D cannot be reached.

Let's use a variant of the Hill equation, using the L_{50} value, the ligand concentration at half-maximum binding.

$$\begin{aligned} 0.5 &= \frac{L_{50}^n}{K_D + L_{50}^n} \\ 1 &= \frac{2L_{50}^n}{K_D + L_{50}{}^n} \end{aligned} \tag{5.12.3}$$

hence

$$Y = \frac{L^n}{K_D + L^n} = \frac{L^n}{L_{50}^n + L^n}\left(\frac{\frac{1}{L^n}}{\frac{1}{L^n}}\right) = \frac{1}{\frac{L_{50}^n}{L^n} + 1} = \frac{1}{\left(\frac{L_{50}}{L}\right)^n + 1} \tag{5.12.4}$$

An analogous, somewhat similar equation can be derived from the logistic equation:

$$Y = d + \frac{a - d}{1 + \left(\frac{L}{c}\right)^b} \tag{5.12.5}$$

where four empirical parameters define the curve:

- a is the smallest measured absorbance value (blank);
- d is the largest absorbance value when Y = 1;
- c is the inflection point in the curve, which can easily be seen to occur when [L]= L_{50}= the ligand concentration at half-maximal saturation, and
- b is the slope of the curve at L_{50}, which is the Hill coefficient n. (for many ELISA curves ≈ 1).

Figure 5.12.3 shows an idealized graph of ELISA data.

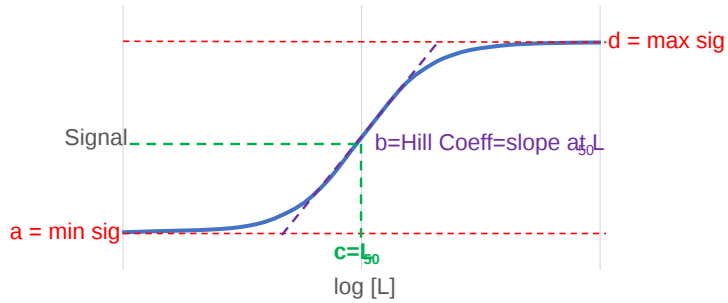


Figure 5.12.3: Idealized ELISA signal (fluorescence, absorbance) vs log [L] curve

The 4-parameter modified Logistic equation is ideal for fitting ELISA data.

An interactive active graph for the 4-parameter Logistic curve is shown in Figure 5.12.4. T

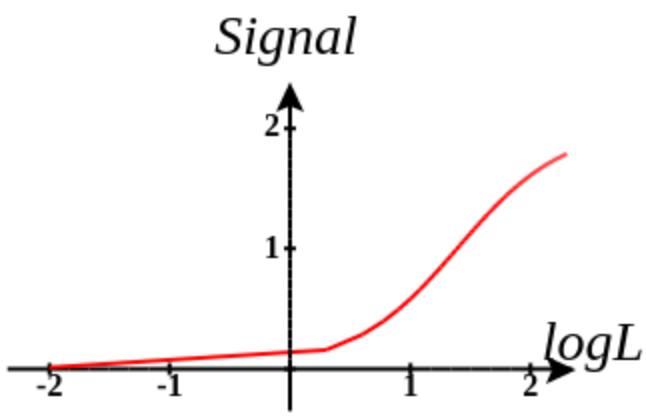


Figure 5.12.4: Interactive active graph for the four-parameter Logistic curve.

Two parameters, the minimum signal **a** and the maximum signal **d,** have been set to 0.01 and 2.0, respectively, and are not changeable in the figure. Note that the greater the value of **b** (slope of the curve at the inflection point), the more "sigmoidal" the semilog curve appears (similar to the Hill binding curves for hemoglobin).

5.12.3: Lateral flow immunoassays

During the COVID pandemic, home test kits (when they were available) were used. These tests are a modified version of the sandwich ELISA described above. They differ in two main ways. The assays were not done in wells but in a planar sheet, where the samples flowed laterally across the sheet. As it flows across a strip through capillary action, a sample containing the SARS-CoV-2 with its spike protein would first encounter a labeled antibody to the spike protein. It would then flow into a region that contained immobilized test (anti-spike protein) and control antibodies. The bound analyte would stop and build up to a sufficient concentration to see an observable colored band on the strip, but only if the sample contained the viral spike protein. These events are illustrated in Figure 5.12.5.

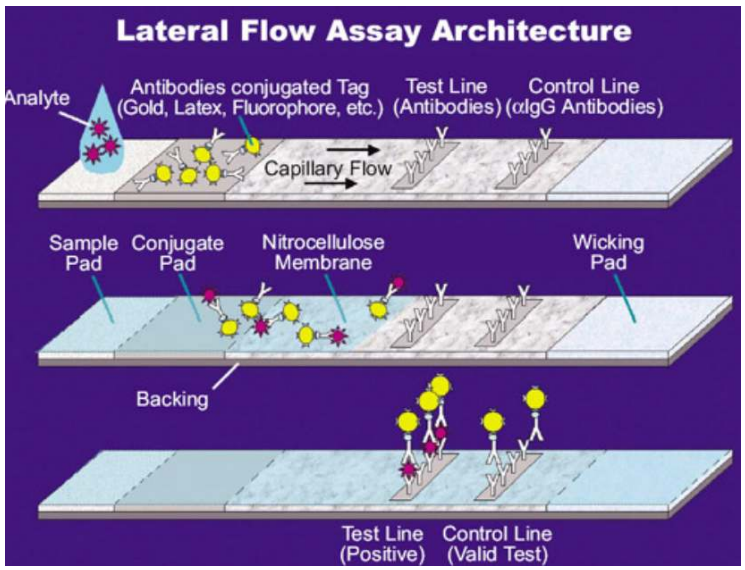


Figure 5.12.5: Lateral flow ELISA assay. https://en.Wikipedia.org/wiki/Latera...Flow_Assay.jpg

Figure 5.12.6 shows a lateral flow assay that detects the presence of serum or potentially salivary **antibodies** (either IgG or IgM) against the SARS-Cov2 proteins.

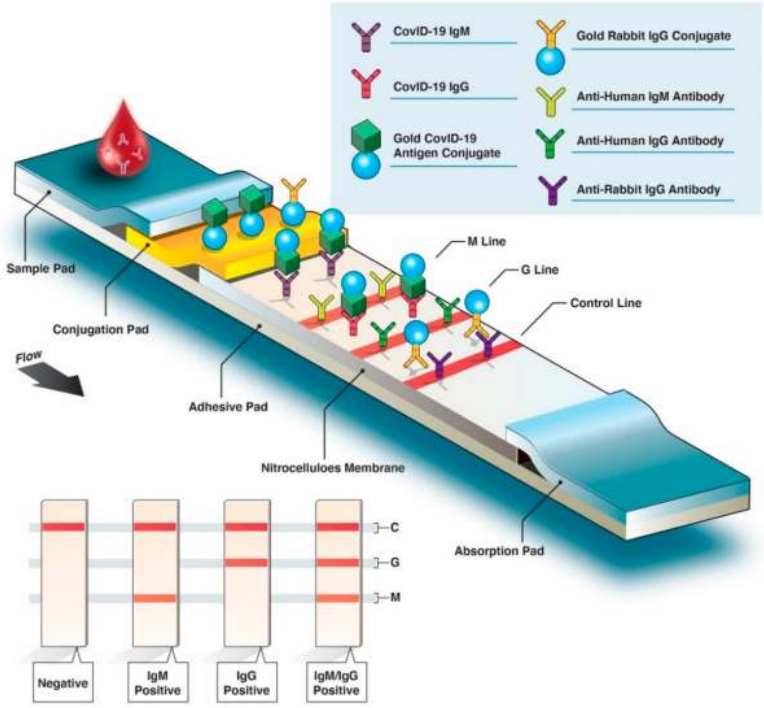


Figure 5.12.6 shows a lateral flow assay that detects the presence of serum or potentially salivary **antibodies** (either IgG or IgM) against the SARS-Cov2 proteins. https://commons.wikimedia.org/wiki/F...0453-g002.webp. Creative Commons Attribution 4.0 International license.

The green cube represents the virus with viral proteins. It has been labeled with a gold nanoparticle (blue sphere). Gold is widely used as a labeling reagent in lateral flow immunoassays because it is chemically inert and extremely stable. The concentrated gold particles found in positive samples at the end of the strip can be observed visually since the gold nanoparticles absorb light through surface plasmon resonance. In this process, light of specific tunable wavelengths (based on the nanoparticle size) is absorbed when it matches the oscillatory frequency of the electron clouds of the metal nanoparticle. Plasmons are the oscillations in the electrons (hence electromagnetic oscillation) that occur when the nanoparticle's surface interacts with photons, causing oscillations in the electrons at the same frequency as the light (resonance).

Gold nanoparticles (AuNPs) are used in a new (2025) assay called NasRED (Nanoparticle-Supported Rapid Electronic Detection). The assay can detect either an antibody to a virus or bacteria, or a bacterial or viral antigen. The assay is rapid, very sensitive, and cheap, allowing almost real-time detection (less than 15 minutes) of a sample. The AuNPs are precoated with biotinylated antigens (to detect antibodies) or antibodies (to detect viral proteins). A small volume sample containing the analyte is added. The system is mixed and centrifuged so the AuNPs are sedimented. Since the AuNPs are derivatized with multiple biotinylated sensors, the binding of antigen to the AuNPs is multivalent. Vortexing the sedimented sample resuspends the unbound AuNPs, leaving the much larger crosslinked AuNPs bound with the antigen precipitated. The absorbance of the supernatant with unbound AnNPs is determined. The more antigen in the sample, the fewer AuNPs in the supernatant, the lower the optical signal.

The assay requires no labeling of the sample. It's reported to be 3000x more sensitive than an ELISA and 100Kx more sensitive than a lateral flow assay. Imagine going to a doctor with an infection that could be viral or bacterial. NasRED for a panel of likely infectious agents could be used, allowing almost immediate diagnosis and prescription of the appropriate antiviral or antibacterial drug.

5.12.4: Summary

This chapter explores the principles, methodologies, and data analysis strategies of enzyme-linked immunosorbent assays (ELISAs), a key analytical tool widely used in biotechnology, pharmaceuticals, and clinical diagnostics. Despite their widespread

application in health-related fields, ELISAs have historically received less emphasis in traditional chemistry and biochemistry curricula. The chapter addresses this gap by detailing both the theoretical and practical aspects of these assays.

1. **Fundamental Principles and Formats**
   The chapter begins by introducing ELISAs as sensitive, selective, and user-friendly immunoassays that rely on specific antibody–antigen interactions. It reviews two primary formats:
   - **Traditional ELISAs:** Where the antigen is immobilized on a solid support and detected by a primary antibody followed by a secondary labeled antibody.
   - **Sandwich ELISAs:** Where an antibody is immobilized to capture the target antigen from a sample, and a second labeled antibody binds to a different epitope on the antigen.
     Both direct and indirect detection methods are discussed, emphasizing how enzyme or fluorophore labels generate measurable signals upon substrate conversion.
2. **Assay Workflow and Mechanism**
   Detailed procedural steps are outlined, including sample addition, antigen immobilization, blocking of non-specific binding sites, primary antibody incubation, and sequential washing steps that remove unbound species. The formation of an antigen–antibody complex, and subsequent binding of a labeled secondary antibody, set the stage for the detection phase. The resulting colorimetric or fluorescent signal is proportional to the amount of analyte present, enabling quantification via spectrophotometry or fluorometry.
3. **Data Analysis and Mathematical Modeling**
   A significant focus is placed on the analysis of ELISA data using mathematical models. The chapter contrasts classical hyperbolic binding models with the Hill equation and introduces a four-parameter logistic equation as an effective tool for fitting standard curves. Key parameters—such as the Hill coefficient (reflecting cooperativity), L_{50} (ligand concentration at half-maximal signal), and the minimum and maximum observable signals—are defined and discussed in the context of quantifying analyte concentrations.
4. **Lateral Flow Immunoassays**
   Extending the discussion to point-of-care diagnostics, the chapter describes lateral flow immunoassays (LFAs), a modified version of the sandwich ELISA format adapted for use in portable, rapid test kits. The design of LFAs, which utilizes capillary action on a planar strip and gold nanoparticles for visual detection via surface plasmon resonance, is explained. This section highlights the importance of LFAs during the COVID-19 pandemic and their role in detecting viral antigens and antibodies.
5. **Practical Applications and Educational Importance**
   The chapter concludes by underscoring the practical significance of ELISAs and LFAs in modern laboratory settings, while also advocating for their increased incorporation into undergraduate curricula. By mastering ELISA techniques and data analysis, students gain valuable skills that bridge biochemical theory with clinical and industrial applications.

Overall, the chapter provides a comprehensive overview of ELISA technology, from its foundational concepts and assay design to advanced data modeling and real-world applications, equipping students with a thorough understanding of this essential biochemical technique.

This page titled 5.12: Binding - Enzyme Linked Immunosorbant Assays (ELISAs) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

6: Enzyme Thermodynamics

6.1: Enzymes
6.2: Enzyme Commission Number
6.3: Endergonic and Exergonic Reactions
6.4: The Laws of Thermodynamics
6.5: Thermodynamics
6.6: Energy in Metabolism

6: Enzyme Thermodynamics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.1: Enzymes

Biological catalysts are called enzymes, and the overwhelming majority of enzymes are proteins. The exceptions are a class of RNA molecules known as ribozymes, of which most act upon themselves (i.e. part of the RNA strand is a substrate for the ribozyme part of the strand). In this book (and most textbooks in this field), unless otherwise specified, the term enzyme refers to one made of protein. Enzymes confer extraordinary specificity to a chemical reaction: a reaction that might occur between a variety of potential substrates in an uncatalyzed situation may only be allowed between two specific substrates when catalyzed by an enzyme. Enzymes allow cells to run chemical reactions at rates from a million to even a trillion times faster than the same reactions would run under similar conditions without enzymes. In some cases, the enzymes al - low reactions to proceed that would normally (i.e. sans enzyme) require more extreme temperature, pressure, or acidity/alkalinity. Finally, and perhaps most importantly for life, enzymes can be regulated. This is crucial for the cell, since it must be able to react to different situations, such as availability of energy, accumulation of toxic byproducts, the need to reproduce, etc. Not only can enzymes be modified either covalently or noncovalently to increase or decrease their activity, the cell can also regulate production of the enzymes, providing another level of control over particular cellular biochemical reactions.

enzyme classification

Enzymes have been cataloged and classified since the 1950's, during which time there was an explosion of enzyme discoveries and a need for a unified nomenclature and catalog. An International Commission on Enzymes was established and thus started the Enzyme List. All enzymes now have both recommended names for common usage, often reflecting historical naming, and a systematic name, which is highly specific. They also have a classification number based on their activity. The major classes of enzymes are

1. *Oxidoreductases* that carry out oxidation-reduction reactions,
2. *Transferases* that transfer functional groups,
3. *Hydrolases* that carry out hydrolysis reactions,
4. *Lyases* that eliminate groups to form double bonds,
5. *Isomerases* that rearrange the bonds in a molecule but do not add or remove atoms, and
6. *Ligases* that form bonds in reactions coupled to ATP hydrolysis.

As an example, DNA ligase (recommended name) catalyzes the formation of a phosphodiester bond between the 3' end of one DNA fragment and the 5' end of another. Its rather long and tedious systematic name is "poly(deoxyribonucleotide):poly(deoxyr ibonucleotide) ligase (AMP-forming)" and its classification number is 6.5.1.1. As a ligase, it is class 6; because it forms phosphoric ester bonds, it is subclass 5; the sub-subclass of 1 in this case is meaningless because it is the only sub-subclass of phosphoricester bond-forming ligases, but the final number designates the DNA ligase separately from other 6.5.1 enzymes such as RNA ligase, which is 6.5.1.3.

Enzymes are the most diverse type of protein in a cell. They vary not only in size, but also in the number of independently manufactured subunits that must come together to form an active enzyme, or holoenzyme. Part of the reason for requiring so many different enzymes is that they are usually very specific for their substrate molecules, and that specificity is based upon a combination of shape and charge. The interactions between substrate and enzyme are often likened to a lock and key or pieces of a jigsaw puzzle. If the substrate fits the shape of the enzyme's active site (the part of the enzyme that carries out the actual catalytic reaction), and the charges interact (e.g. positively charged amino acids on the enzyme lining up with negative charges on the substrate), then there may be further stabilization of the interaction by Van der Waals and hydrogen bond interactions. In fact, formation of a stable Enzyme-Substrate (ES) intermediate is energetically analogous to the transition state (Figure **6.1.2**) of reactions.

The specificity of enzymes is such that stereoisomers may not be recognized by some enzymes: for example, a protease (enzymes that chop up proteins into smaller pieces by hydrolyzing the peptide bonds between specific amino acids) such as trypsin can be stymied by the presence of a D-amino acid in place of the usual L-amino acid in a protein, even though it is a mirror image of the very same amino acid. This specificity means that enzymes are highly selective with respect to the reactions they catalyze, which means that specific reactions can be greatly enhanced without causing a general increase in many related chemical reactions.

Another implication of the high specificity is that enzymes can (and often do) have high affinity for their substrates without the problem of binding non-substrate molecules (other than specific inhibitors - see below).

If most biochemical reactions would proceed extremely slowly, if at all, without catalysis, enzymes are needed to lower the activation energy needed for chemical reactions to support life. Exactly how does an enzyme lower the activation energy of a reaction? What exactly does “activation energy” mean in the context of a cell? To understand this, there are two principles to keep in mind: first, when we talk about chemical reactions, generally, we are concerned with populations of substrate, product, and enzyme molecules, not individuals; and second, the reactions are generally taking place between molecules dissolved in the aqueous cytoplasm of the cell.

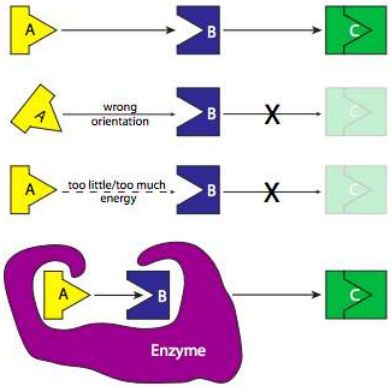


Figure 6.1.3. Enzymes can lower activation energy by binding substrates individually and bringing them together under optimal conditions to react.

Consider a reaction in which substrates A and B interact to form product C (Figure 6.1.3). If this reaction is not catalyzed, it depends on the happenstance that a molecule of A runs into a molecule of B in just the right orientation, and with the right amount of energy, to react and form the new molecule. We can conceptualize “activation energy” as the difficulty in getting A and B together perfectly so the reaction can proceed. How might an enzyme lower this activation energy? By making it easier for A and B to find each other with the right orientation and energy. So it could have binding sites for molecule A and molecule B, and once it has bound these two molecules, it changes its conformation, bringing A and B together under exactly the right conditions to react and form C. Once the reaction is complete, the product floats off because the enzyme has no affinity for it, and the enzyme returns to its initial shape, ready to bind more substrates.

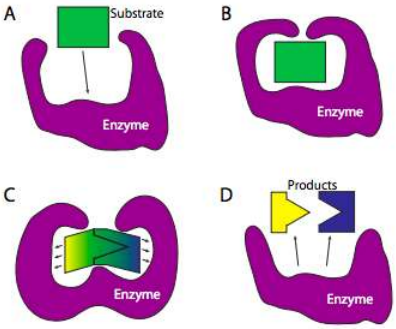


Figure 6.1.4. The enzyme (A) binds a specific substrate, leading to a conformational change (B) that stresses the molecule (C) until it breaks down into two smaller component molecules (D), which are released.

> Enzymes may also facilitate a chemical reaction by acting as a temporary holding site for an active group being transferred from one substrate to another. Alternatively, temporary formation of hydrogen bonds or even covalent bonds between the enzyme and substrate can alter chemical characteristics of the substrate to make it react more easily. An example of enzyme mechanisms on the molecular level is shown in Chapter 5: Figure 6.1.1.

Another example may be found with enzymes that break apart a molecule (Figure 6.1.4). In order for a molecule to break apart, it may need to collide with another molecule with sufficient energy to break one or more of its covalent bonds. An enzyme that

catalyzes the breakdown reaction might bind to the molecule, and in binding it, undergoes a conformational shift that bends or twists the molecule in such a way that the bonds in the substrate molecule are weakened or broken. These two examples oversimplify the chemistry of enzyme activity into a mechanical idea, but the general relationship in how an enzyme lowers activation energy for a reaction is accurate.

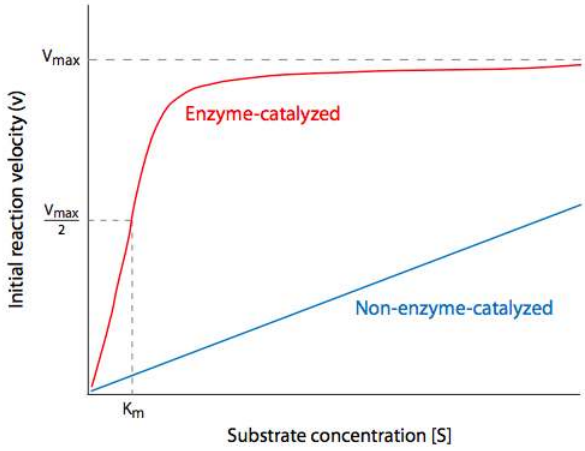


Figure 6.1.5. Saturation curve for a single-substrate enzyme-catalyzed reaction.

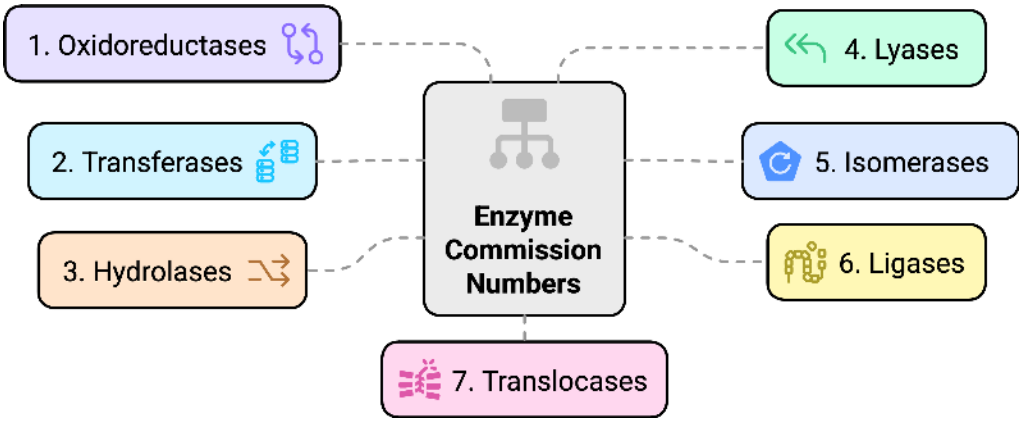
This page titled 6.1: Enzymes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by E. V. Wong via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Enzymes** by E. V. Wong is licensed CC BY-NC-SA 3.0. Original source: https://www.axopub.com/wp01/2012/02/26/download-cells-molecules-and-mechanisms/.

6.2: Enzyme Commission Number

6.2.1 Introduction

Enzymes are biological catalysts that speed up chemical reactions in living organisms. Scientists needed a way to organize the growing list of enzymes, leading to the creation of the **Enzyme Commission (EC) number system**. This system works like the **Dewey Decimal System** or the **Library of Congress classification**, but for enzymes. Each enzyme is assigned a unique four-part EC number based on the type of reaction it catalyzes.



The system, developed by the International Union of Biochemistry and Molecular Biology (IUBMB), provides a standardized way to classify and study enzymes.

6.2.2 The Seven Major Classes of the Enzyme Commission

EC Number	Class Name	Reaction Catalyzed	Example Reaction	Enzyme Examples (Trivial Names)	Enzyme Count (as of March 2025)
EC 1	**Oxidoreductases**	Catalyze oxidation-reduction (redox) reactions by transferring electrons or hydrogen atoms.	$AH + B \rightarrow A + BH$ (reduced) $2A + O_2 \rightarrow 2AO$ (oxidized)	dehydrogenase, oxidase	2,010
EC 2	**Transferases**	Transfer functional groups (methyl, acyl, amino, phosphate, etc.) between molecules.	$AB + C \rightarrow A + BC$	transaminase, kinase	2,069
EC 3	**Hydrolases**	Break chemical bonds by adding water (hydrolysis).	$AB + H_2O \rightarrow AOH + BH$	lipase, amylase, peptidase, phosphatase	1,357
EC 4	**Lyases**	Break bonds without hydrolysis or oxidation, often forming double bonds.	$X\text{-}A + B\text{-}Y \rightarrow A{=}B + X\text{-}Y$	decarboxylase	773
EC 5	**Isomerases**	Rearrange atoms within a molecule without changing its molecular formula.	$ABC \rightarrow BCA$	isomerase, mutase	320
EC 6	**Ligases**	Join two molecules by forming new bonds (C-O, C-S, C-N, or C-C) with ATP hydrolysis.	$X + Y + ATP \rightarrow XY + ADP + P_i$	synthetase	249

EC 7	**Translocases**	Move ions or molecules across membranes.	—	transporter	98

6.2.3 The Six Major Classes of Enzyme Reactions

6.2.3.1 EC 1: Oxidoreductases

Oxidoreductases catalyze oxidation-reduction (*redox*) reactions by transferring electrons or hydrogen atoms from one molecule to another. In these reactions, at least one substrate is oxidized while another is reduced. These enzymes are crucial in metabolic pathways such as cellular respiration.

Common oxidoreductases include **dehydrogenases, reductases, and oxidases**. A well-known example is **lactate dehydrogenase**, which plays a key role in glycolysis by converting pyruvate into lactate.

6.2.3.1.1 Example Reaction: Lactate Dehydrogenase

Pyruvate is reduced to lactate while NADH is oxidized to NAD^+:

$$\text{Pyruvate} + \text{NADH} + H^+ \rightleftharpoons \text{Lactate} + \text{NAD}^+ \tag{6.2.1}$$

6.2.3.2 EC 2: Transferases

Transferases move functional groups, such as methyl, acyl, amino, or phosphate groups, from one molecule to another. These enzymes often require cofactors that act as donors.

One important subclass of transferases is the **kinases**, which transfer phosphate groups. For example, **hexokinase** phosphorylates glucose, enabling it to enter metabolic pathways like glycolysis.

6.2.3.2.1 Example Reaction: Hexokinase

Glucose is phosphorylated using ATP:

$$\text{Glucose} + \text{ATP} \rightarrow \text{Glucose-6-phosphate} + \text{ADP} \tag{6.2.2}$$

6.2.3.3 EC 3: Hydrolases

Hydrolases break chemical bonds by adding water in a process called *hydrolysis*. They typically act on C-O, C-N, and C-S bonds, breaking down larger molecules into smaller components.

Hydrolases are widely involved in digestion and metabolism. A key example is **trypsin**, a proteolytic enzyme that breaks down proteins in the digestive system.

6.2.3.3.1 Example Reaction: Dipeptide Hydrolysis

Alanine dipeptide hydrolysis into two alanine molecules:

$$\text{Ala-Ala} + H_2O \rightarrow \text{Ala} + \text{Ala} \tag{6.2.3}$$

6.2.3.4 EC 4: Lyases

Lyases catalyze the removal or addition of functional groups to form or break double bonds, without using hydrolysis or oxidation. These enzymes can break bonds such as C=C, C=O, C=N, and C=S.

An example of a lyase is **aldolase**, which is essential in glycolysis for breaking down fructose-1,6-bisphosphate into smaller sugar molecules.

6.2.3.4.1 Example Reaction: Aldolase

Cleavage of fructose-1,6-bisphosphate into two three-carbon sugars:

$$\begin{aligned}&\text{Fructose-1,6-bisphosphate}\\&\rightarrow \text{Glyceraldehyde-3-phosphate} + \text{Dihydroxyacetone phosphate}\end{aligned} \tag{6.2.4}$$

6.2.3.5 EC 5: Isomerases

Isomerases rearrange atoms within a molecule to form an isomer. The chemical formula remains unchanged, but the structure is altered, allowing for different biochemical functions.

A common example is **glucose-6-phosphate isomerase**, which converts glucose-6-phosphate into fructose-6-phosphate in the second step of glycolysis.

6.2.3.5.1 Example Reaction: Glucose-6-Phosphate Isomerase

Conversion of glucose-6-phosphate to fructose-6-phosphate:

$$\text{Glucose-6-phosphate} \rightleftharpoons \text{Fructose-6-phosphate} \tag{6.2.5}$$

6.2.3.6 EC 6: Ligases

Ligases catalyze the joining of two molecules by forming new chemical bonds, often powered by ATP hydrolysis. They are responsible for the formation of carbon-carbon, carbon-nitrogen, carbon-oxygen, and carbon-sulfur bonds.

An essential ligase is **DNA ligase**, which is crucial in DNA replication and repair, joining broken DNA strands together.

6.2.3.6.1 Example Reaction: Pyruvate Carboxylase

Formation of oxaloacetate from pyruvate:

$$\text{Pyruvate} + CO_2 + \text{ATP} \rightarrow \text{Oxaloacetate} + \text{ADP} + P_i \tag{6.2.6}$$

6.2.3.7 EC 7: Translocases – An Extension to the System

In 2018, the Enzyme Commission introduced EC 7: Translocases, a new category for enzymes that facilitate the movement of molecules, ions, or macromolecules across membranes or within cellular compartments. Unlike other enzyme classes, translocases do not chemically modify their substrates. Instead, they catalyze translocation, often utilizing energy sources like ATP hydrolysis or ion gradients to drive transport.

The classification of translocases follows the general principles of enzyme nomenclature set by the Enzyme Commission:

- Only single catalytic entities should be classified as enzymes. Multi-enzyme systems that participate in transport should include “system” in their names (e.g., proton pump system).
- Translocases should not be confused with enzymes that modify their substrates. Unlike oxidoreductases (EC 1) or transferases (EC 2), translocases move substances without altering their chemical structure.
- Systematic names should reflect the molecule transported and the driving energy source. Examples include ATP-driven ion pumps and symporters using electrochemical gradients.

As research progresses, new translocases continue to be discovered and classified. Major enzyme databases, including Expasy, now recognize EC 7 translocases as an established extension to enzyme classification.

6.2.4 Understanding the EC Numbering System

The Enzyme Commission (EC) numbering system organizes enzymes into a hierarchy based on the reactions they catalyze. Each enzyme is assigned a unique four-part EC number (e.g., **EC 3.4.11.4**), where each level provides more specificity. Let’s explore how this system is structured.

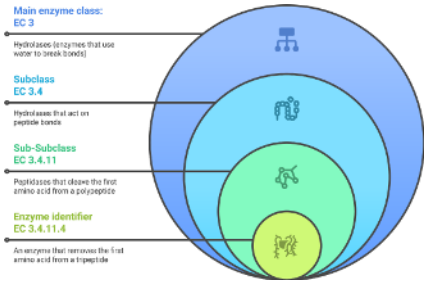
6.2.4.1 Four Parts of the EC Number

Each enzyme’s EC number consists of four parts, separated by periods:

EC Number Component	Meaning	Description
First digit	Main enzyme class (EC X)	Indicates the broad category of reaction catalyzed (e.g., oxidation-reduction, transfer, hydrolysis).
Second digit	Subclass (EC X.Y)	Specifies the type of reaction within the main class, often based on the type of bond or chemical group involved.
Third digit	Sub-subclass (EC X.Y.Z)	Further defines the reaction mechanism or the specific type of chemical transformation.
Fourth digit	Enzyme identifier (EC X.Y.Z.W)	Uniquely identifies the specific enzyme that catalyzes a reaction within the given subclass.

Let's examine the classification of a specific enzyme, EC 3.4.11.4, to see how it fits into the system:

- **EC 3** - Hydrolases (enzymes that use water to break bonds)
- **EC 3.4** - Hydrolases that act on peptide bonds
- **EC 3.4.11** - Peptidases that cleave the first amino acid from a polypeptide
- **EC 3.4.11.4** - An enzyme that removes the first amino acid from a tripeptide

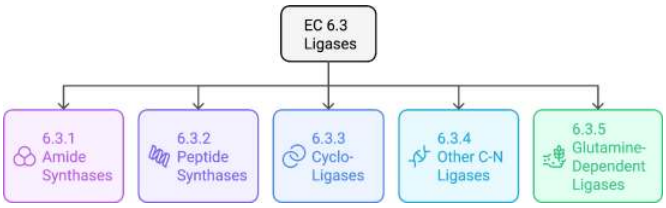


6.2.4.2 Example 1: Exploring an Entire Class (EC 2 - Transferases)

Now, let's look at an entire enzyme class, EC 2 (Transferases), which move functional groups between molecules. This class includes many subclasses, each defined by the type of group being transferred:

EC Number	Subclass Name	Description
EC 2.1	Transferring One-Carbon Groups	Transfers single-carbon units such as methyl groups
EC 2.2	Transferring Aldehyde or Ketone Groups	Transfers aldehyde or ketone groups between molecules
EC 2.3	Acyltransferases	Transfers acyl groups (e.g., acetyl, fatty acyl)
EC 2.4	Glycosyltransferases	Transfers sugar (glycosyl) groups
EC 2.5	Transferring Alkyl or Aryl Groups	Transfers alkyl or aryl groups (except methyl groups)
EC 2.6	Transferring Nitrogenous Groups	Transfers amino or related nitrogen-containing groups
EC 2.7	Transferring Phosphorus-Containing Groups	Transfers phosphate groups (e.g., kinases)
EC 2.8	Transferring Sulfur-Containing Groups	Transfers sulfur-based functional groups
EC 2.9	Transferring Selenium-Containing Groups	Transfers selenium-based functional groups

6.2.4.3 Example 2: A Deep Dive into a Subclass (EC 6.3 - Carbon-Nitrogen Bond Formation)



Finally, let's zoom into a specific subclass, EC 6.3 (Ligases that form carbon-nitrogen bonds). This subclass includes several enzyme sub-subclasses:

EC Number	Sub-subclass Name	Description
EC 6.3.1	Acid-Ammonia (or Amine) Ligases (Amide Synthases)	Form amide bonds using ammonia or amines
EC 6.3.2	Acid-Amino-Acid Ligases (Peptide Synthases)	Form peptide bonds between amino acids
EC 6.3.3	Cyclo-Ligases	Form cyclic molecules via carbon-nitrogen bonds
EC 6.3.4	Other Carbon-Nitrogen Ligases	Form C-N bonds but don't fit into other categories
EC 6.3.5	Carbon-Nitrogen Ligases with Glutamine as Amido-N-Donor	Use glutamine as a nitrogen donor for C-N bond formation

6.2.5 History of the EC Number System

6.2.5.1 The Road to Systematic Naming

- Early enzyme names were inconsistent, often based on the substrate or reaction type (e.g., urease, maltase).
- By the 1950s, biochemists realized the need for a standardized classification system.
- In 1956, the International Commission on Enzymes was formed, and in 1961, the first edition of the EC system was published.

6.2.5.2 Growth of the EC System Over Time

Year	Publication	Number of Enzymes Classified
1961	Report of the Enzyme Commission	712
1964	Enzyme Nomenclature	875
1972	Enzyme Nomenclature	1,770
1978	Enzyme Nomenclature	2,122
1984	Enzyme Nomenclature	2,477
1992	Enzyme Nomenclature	3,196
2013	—	5,212
2020	—	6,520
2023	—	6,754
2025	—	6,876

As research expands, new enzymes continue to be classified, with the list regularly updated by the IUBMB.

6.2.6 Additional Resources

For up-to-date enzyme classifications, visit:

- IUBMB Enzyme Database – Official classification list.
- Expasy Enzyme Database – Searchable enzyme database.
- The Enzyme Database – Comprehensive enzyme classification and nomenclature.
- Wikipedia - EC Number – General overview.

6.2: Enzyme Commission Number is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.3: Endergonic and Exergonic Reactions

Endergonic and exergonic reactions

Any system of molecules that undergoes a physical transformation/reorganization (aka. reaction) will have an associated change in internal energy and entropy. If we examine a single isolated reaction, in which unique reactants convert into unique products the Gibbs energy of the system will be dependent several factors, key among which are (a) the internal energy and entropy differences associated with the molecular rearrangements and (b) the degree to which the reaction is out-of-equilibrium.

If, for the sake of simplicity we begin by considering only the contribution of the molecular transformations in the system on ΔG, we conclude that reactions with $\Delta G < 0$, the products of the reaction have less Gibbs energy than the reactants. Since ΔG is the difference between the enthalpy and temperature-scaled entropy changes in a reaction, a net negative ΔG can arise in through changes largely of enthalpy, entropy or most often both. The left panel of Figure 1 below shows a common graphical representation of an **exergonic** reaction. This graph is called a reaction coordinate diagram. It plots Gibbs energy on the y-axis, and the x-axis in arbitrary units shows the progress of a reaction. With an exergonic reaction, the figure on the left shows two key things: (1) the difference between the free energy of the reactants and products is negative and (2) the progress of the reaction requires some input of free energy (shown as an energy "hill" or barrier). This graph does not tell us how the energy in the system redistributed, only that the difference between enthalpy and temperature-scaled entropy is negative. **Exergonic reactions** are said to occur spontaneously. Understanding which chemical reactions are spontaneous is useful for biologists who are trying to understand whether a reaction is likely to "go" or not.

It is important to note that the term "spontaneous"—in thermodynamics—implies nothing about how fast the reaction proceeds. The change in free energy only describes the difference between beginning and end states, NOT how fast that transition takes place. This is contrary to the everyday use of the term which usually carries the implicit understanding that something happens quickly. As an example, the oxidation/rusting of iron is a spontaneous reaction. However, an iron nail exposed to air does not rust instantly—it may take years.

A chemical reaction with a positive ΔG means that the products of the reaction have a higher free energy than the reactants (see the right panel of Figure 1). These chemical reactions are called **endergonic reactions**, and they are NOT spontaneous. An endergonic reaction will not take place on its own without transferring energy into the reaction or increase of entropy somewhere else.

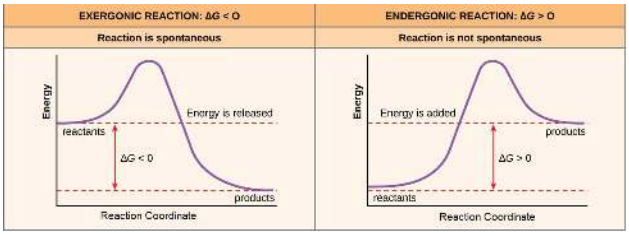


Figure 1. *Reaction coordinate diagrams of exergonic and endergonic reactions. Exergonic and endergonic reactions are characterized by changes in Gibbs energy. In the equilibrium state of an exergonic reaction, the Gibbs energy of the products is lower than that of the reactants. Meanwhile, the equilibrium state of an endergonic reaction in, the Gibbs energy of the products is higher than that of the reactants. Attribution: Marc T. Facciotti (own work)*

The building of complex molecules, such as sugars, from simpler ones is an anabolic process and is endergonic. On the other hand, the catabolic process, such as the breaking down of sugar into simpler molecules, is exergonic. Like the example of rust above, while the breakdown of biomolecules is generally spontaneous, these reactions don't occur need to occur instantaneously (quickly). Remember, the terms endergonic and exergonic only refer to the difference in Gibbs energy between the products and reactants; they don't tell you about the rate of the reaction (how fast it happens). Rate will be discussed in later sections.

An important concept in the study of metabolism and energy is that of chemical equilibrium. Most chemical reactions are reversible. They can proceed in both directions, often transferring energy into their environment in one direction and transferring energy in from the environment in the other direction. The same is true for the chemical reactions involved in cell metabolism, such as the breaking down and building up of proteins into and from individual amino acids, respectively. Reactants within a closed system will undergo chemical reactions in both directions until a state of equilibrium is reached. **Equilibrium** in a chemical

reaction is the state in which both reactants and products are present in concentrations that have no further tendency to change with time. Usually, this state results when the forward reaction proceeds at the same rate as the reverse reaction. NOTE THIS LAST STATEMENT! Equilibrium means that the relative concentrations of reactants and products are not changing in time, BUT it does NOT mean that there is no interconversion between substrates and products—it just means that when the reactant(s) are converted to product(s) that product(s) are converted to reactant(s) at an equal rate (see Figure 2). The state of equilibrium is also one of the lowest possible free energy states for the reaction and is a state of maximal entropy.

If a reaction is kept or started far out of equilibrium this state of the system also contributes to the overall Gibbs energy of a reaction. Either a rebalancing of substrate or product concentrations (by adding or removing substrate or product) or a positive change in free energy, typically by the transfer of energy from outside the reaction, can to move a reaction to an out-of-equilibrium state. Note that in a living cell, most chemical reactions do not reach a state of equilibrium—this would require that they reach their lowest free energy state, a state that is almost by definition incompatible with life. Energy is therefore required to keep biological reactions out of their equilibrium state. In this way, living organisms are in a constant, energy-requiring, uphill battle against equilibrium and entropy. This also means that the Gibbs energy of most biological reactions as they occur in the cell must also include a contribution from this out-of-equilibrium state. The Gibbs energy of these reactions, therefore, is often different from that reported under standard conditions.

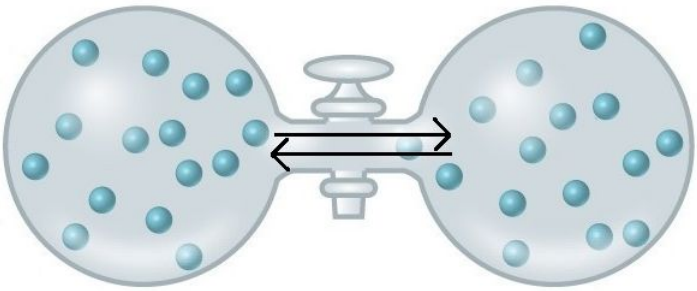


Figure 2. *At equilibrium, do not think of a static, unchanging system. Instead, picture molecules moving in equal amounts from one area to another. Here, at equilibrium, molecules are still moving from left to right and right to left. The net movement, however, is equal. There will still be about 15 molecules in each side of this flask once it reaches equilibrium. Source: https://courses.candelalearning.com/...apter/entropy/*

This page titled 6.3: Endergonic and Exergonic Reactions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Marc Facciotti.

- **6.1: Endergonic and Exergonic Reactions** by Marc Facciotti has no license indicated.

6.4: The Laws of Thermodynamics

Two fundamental concepts govern energy as it relates to living organisms: the First Law of Thermodynamics states that total energy in a closed system is neither lost nor gained — it is only transformed. The Second Law of Thermodynamics states that entropy constantly increases in a closed system.

More specifically, the First Law states that energy can neither be created nor destroyed: it can only change form. Therefore, through any and all processes, the total energy of the universe or any other closed system is constant. In a simple thermodynamic system, this means that the energy is transformed either by the transfer of heat energy (i.e. heating and cooling of a substance) or by the production of mechanical work (i.e. movement). In biological and chemical terms, this idea can be extended to other forms of energy such as the chemical energy stored in the bonds between atoms of a molecule, or the light energy that can be absorbed by plant leaves.

> Work, in this case, need not imply a complicated mechanism. In fact, there is work accomplished by each molecule in the simple expansion of a heated mass of gaseous molecules (as visualized by expansion of a heated balloon, for example). This is expressed mathematically as the Fundamental Thermodynamic Relation:
>
> in which E is internal energy of the system, T is temperature, S is entropy, p is pressure, and V is volume.
>
> Unlike the First Law which applies even to particles within a system, the Second Law is a statistical law — it applies generally to macroscopic systems. However, it does not preclude smallscale variations in the direction of entropy over time. In fact, the Fluctuation theorem (proposed in 1993 by Evans et al, and demonstrated by Wang et al in 2002) states that as the length of time or the system size increases, the probability of a negative change in entropy (i.e. going against the Second Law) decreases exponentially. So on very small time scales, there is a real probability that fluctuations of entropy against the Second Law can exist.

The Second Law dictates that entropy always seeks to increase over time. Entropy is simply a fancy word for chaos or disorder. The theoretical final or equilibrium state is one in which entropy is maximized, and there is no order to anything in the universe or closed system. Spontaneous processes, those that occur without external influence, are always processes that convert order to disorder. However, this does not preclude the imposition of order upon a system. Examining the standard mathematical form of the Second Law:

$\Delta S_{system} + \Delta S_{surroundings} = \Delta S_{universe}$,

where $\Delta S_{universe} > 0$

shows that entropy can decrease within a system as long as there is an increase of equal or greater magnitude in the entropy of the surroundings of the system.

The phrase “in a closed system” is a key component of these laws, and it is with the idea encapsulated in that phrase that life can be possible. Let’s think about a typical cell: in its lifetime, it builds countless complex molecules - huge proteins and nucleic acids formed from a mixture of small amino acids or nucleotides, respectively. On its surface, this example might seem to be a counterexample to the second law - clearly going from a mixture of various small molecules to a larger molecule with bonded and ordered components would seem to be a decrease in entropy (or an increase in order). How is this possible with respect to the second law? It is, because the second law applies only to closed systems. That is, a system that neither gains nor loses matter or energy.

> The “universe” is a closed system by definition because there is nothing outside of it.

A living cell is not a closed system: it has inputs and outputs. However, the second law is still useful if we recognize that the only way that it can be bypassed is through the input of energy. If a cell cannot take in food (input of matter and energy into the system) it dies, because the second law requires that everything eventually breaks down into more random/chaotic collections of smaller components. The order required to sustain life (think about all the different complex molecules that were mentioned in the previous

chapter) is phenomenal. The same thing applies on the organismal level (Figure 6.4.1) - without an input of energy (in the form of food molecules for animals or in the form of light for plants), the organism will die and subsequently decompose.

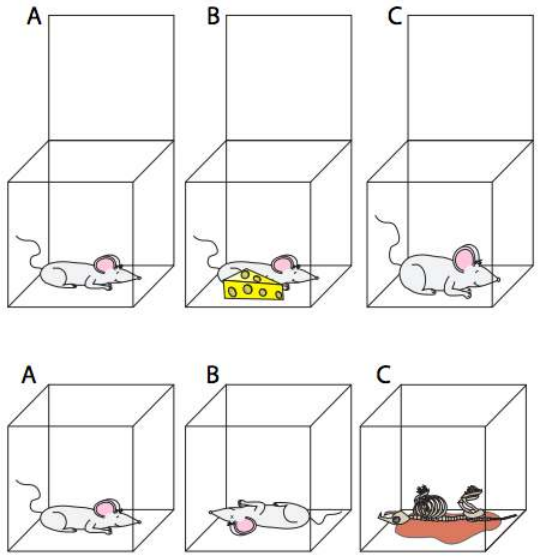


Figure 6.4.1. In the top panel, depicting an open system in which there are inputs for matter and energy (in the form of food), the box is open and air can be exchanged, food dropped in, etc, thus allowing the mouse to grow. However, in the bottom panel, depicting a closed system, the mouse does not have ready access any more oxygen than is in the box, nor does it have access to food. Without these inputs, the second law takes effect, and the mouse dies and decomposes into many smaller molecules.

Creating molecules from atoms costs energy because it takes a disordered collection of atoms and forces them, through chemical bonds, into ordered, non-random positions. There is likewise an energy cost to formation of macromolecules from smaller molecules. By imposing order in the system, there must be an associated input of energy. This happens at every level of the system: atoms to molecules, small molecules to macromolecules, groups of molecules to organelles, etc.

> As the polymerizing reaction reduces entropy, it requires energy generated (usually) by the breakdown of ATP into AMP and PPi, which is a reaction that increases entropy.

Where does that energy go? It ends up in the bonds that are holding the molecules or macromolecules in their ordered state. When such a bond is broken, and a molecule is turned back into a collection of atoms, energy is released. The energy in a chemical bond is thus potential energy - it is stored energy that, when released, has the ability to do work. This term, if you recall your high school physics, is usually learned along with kinetic energy, which is energy that is being used in the process of actually doing work (i.e. moving an object from one place to another). The classic example is the rock on the top of a hill: it has potential energy because it is elevated and could potentially come down. As it tumbles down, it has kinetic energy as it moves. Similarly in a cell, the potential energy in a chemical bond can be released and then used for processes such as putting smaller molecules together into larger molecules, or causing a molecular motor to spin or bend - actions that could lead to pumping of protons or the contraction of muscle cells, respectively.

Coming back to the second law, it essentially mandates that breaking down molecules releases energy and that making new molecules (going against the natural tendency towards disorder) requires energy. Every molecule has an intrinsic energy, and therefore whenever a molecule is involved in a chemical reaction, there will be a change in the energy of the resulting molecule(s). Some of this change in the energy of the system will be usable to do work, and that energy is referred to as the free energy of the reaction. The remainder is given off as heat.

The Gibbs equation describes this relationship as

$$\Delta G = \Delta H - T\Delta S \tag{6.4.1}$$

where ΔG is the change in free energy, ΔH is the change in enthalpy (roughly equivalent to heat), T is the temperature at which the reaction takes place, and ΔS is the change in entropy. As a matter of convention, release of free energy is a negative number, while

a requirement for input of energy is denoted with a positive number. Generally, a chemical reaction in which ΔG < 0 is a spontaneous reaction (also called an exergonic reaction), while a chemical reaction in which ΔG > 0 is not spontaneous (or endergonic). When ΔG = 0, the system is in equilibrium. ΔG can also be expressed with respect to the concentration of products and reactants:

$$\Delta G = \Delta G^o + RT \ln \left(\frac{[P_1][P_2][P_3]...}{[R_1][R_2][R_3]...} \right) \tag{6.4.2}$$

Terms in square brackets denote concentrations, ΔG° is the standard free energy for the reaction (as carried out with 1M concentration of each reactant, at 298K and at 1 atm pressure), R is the gas constant (1.985 cal $K^{-1}mol^{-1}$), and T is the temperature in Kelvin. In a simpler system in which there are just two reactants and two products:

$$a\,\mathrm{A} + b\,\mathrm{B} \rightleftharpoons c\,\mathrm{C} + d\,\mathrm{D} \tag{6.4.3}$$

the equation for free energy change becomes

$$\Delta G = \Delta G° + RT\ln([C]^c[D]^d/[A]^a[B]^b)$$

This is important to us as cell biologists because although cells are not very well suited to regulating chemical reactions by varying the temperature or the pressure of the reaction conditions, they can relatively easily alter the concentrations of substrates and products. In fact, by doing so, it is even possible to drive a non-spontaneous reaction (ΔG > 0) forward spontaneously (ΔG < 0) either by increasing substrate concentration (possibly by transporting them into the cell) or by decreasing product concentration (either secreting them from the cell or by using them up as substrates for a different chemical reaction).

Changes in substrate or product concentration to drive a non-spontaneous reaction are an example of the more general idea of coupling reactions to drive an energetically unfavorable reactions forward. Endergonic reactions can be coupled to exergonic reactions as a series of reactions that ultimately is able to proceed forward. The only requirement is that the overall free energy change must be negative (ΔG < 0). So, assuming standard conditions (ΔG = ΔG°'), if we have a reaction with a free energy change of +5 kcal/mol, it is non-spontaneous. However, if we couple this reaction, to ATP hydrolysis for example, then both reactions will proceed because the standard free energy change of ATP hydrolysis to ADP and phosphate is an exergonic -7.3 kcal/mol. The sum of the two ΔG values is -2.3 kcal/mol, which means the coupled series of reactions is spontaneous.

In fact, ATP is the most common energy "currency" in cells precisely because the -7.3 kcal/mol free energy change from its hydrolysis is enough to be useful to drive many otherwise endergonic reactions by coupling, but it is less costly (energetically) to make than other compounds that could potentially release even more energy (e.g. phosphoenolpyruvate, PEP). Also, much of the -14.8 kcal/mol (ΔG°') from PEP hydrolysis would be wasted because relatively few endergonic reactions are so unfavorable as to need that much free energy.

> Why is ATP different from other small phosphorylated compounds? How is it that the γ-phosphoanhydride bond (the most distal) of ATP can yield so much energy when hydrolysis of glycerol-3-phosphate produces under a third of the free energy? The most obvious is electrostatic repulsion. Though they are held together by the covalent bonds, there are many negative charges in a small space (each phosphate carries approximately 4 negative charges). Removing one of the phosphates significantly reduces the electrostatic repulsion. Keeping in mind that DG is calculated from the equilibrium of both reactants and products, we also see that the products of ATP hydrolysis, ATP and phosphate, are very stable due to resonance (both ADP and Pi have greater resonance stabilization) and stabilization by hydration. The greater stability of the products means a greater free energy change.

Even when a reaction is energetically favorable (ΔG < 0), it may not occur without a little "push", chemically speaking. The "push" is something called *activation energy*, and it overcomes thermodynamic stability. Consider glucose, for instance. This simple sugar is the primary source of energy for all cells and the energy inherent within its bonds is released as it breaks down into carbon dioxide and water. Since this is large molecule being broken down into smaller ones, entropy is increased, thus energy is released from reaction, and it is technically a spontaneous reaction. However, if we consider a some glucose in a dish on the lab bench, it clearly is not going to spontaneously break down unless we add heat. Once we add sufficient heat energy, we can remove the energy source, but the sugar will continue to break down by oxidation (burn) to CO_2 and H_2O.

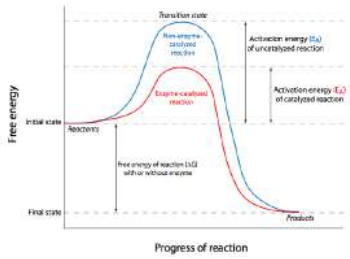


Figure 6.4.2. Catalysts lower the activation energy barrier to chemical reactions without altering the free energy change for that reaction.

Put another way, the reactant(s) must be brought to an unstable energy state, known as the transition state (as shown at the peak of the graphs in Figure 6.4.2). This energy requirement barrier to the occurrence of a spontaneous thermodynamically favored reaction is called the activation energy. In cells, the activation energy requirement means that most chemical reactions would occur too slowly/infrequently to allow for all the processes that keep cells alive because the required energy would probably come from the chance that two reactants slam into one another with sufficient energy, usually meaning they must be heated up. Again, cells are not generally able to turn on some microscopic Bunsen burner to generate the activation energy needed, there must be another way. In fact, cells overcome the activation energy problem by using catalysts for their chemical reactions. Broadly defined, a catalyst is a chemical substance that increases the rate of a reaction, may transiently interact with the reactants, but is not permanently altered by them. The catalyst can be re-used because it is the same before the reaction starts, and after the reaction completes. From a thermodynamic standpoint, it lowers the activation energy of the reaction, but it does not change the ΔG. Thus it cannot make a non-spontaneous reaction proceed; it can only make an already spontaneous reaction occur more quickly or more often.

This page titled 6.4: The Laws of Thermodynamics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by E. V. Wong via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: The Laws of Thermodynamics** by E. V. Wong is licensed CC BY-NC-SA 3.0. Original source: https://www.axopub.com/wp01/2012/02/26/download-cells-molecules-and-mechanisms/.

6.5: Thermodynamics

Thermodynamics

MCAT Connection

Thermodynamics is concerned with describing the changes in systems before and after a change. This usually involves a discussion about energy transfers and its dispersion within the system and its surroundings. In nearly all practical cases, these analyses require that the system and its surroundings be completely described. For instance, when discussing the heating of a pot of water on the stove, the system may include the stove, the pot, and the water and the environment or surroundings may include everything else. Biological organisms are what are called open systems; energy is transferred between them and their surroundings.

The First Law of Thermodynamics

The first law of thermodynamics deals with the total amount of energy in the universe. It states that this total amount of energy is constant. In other words, there always has been, and always will be, exactly the same amount of energy in the universe.

According to the first law of thermodynamics, energy may be transferred from place to place, but it cannot be created or destroyed. Energy transfers take place around us all the time. Light bulbs transfer energy from electrical power stations into heat and photons of light. Gas stoves transfer energy stored in the bonds of chemical compounds into heat and light. (Heat, by the way, is the amount of energy transferred from one system to another because of a temperature difference.)

Plants perform one of the most biologically useful energy transfers on earth: they transfer energy in the photons of sunlight into the chemical bonds of organic molecules. In every one of these cases, energy is neither made nor destroyed, and we must try to account for *all* the energy when we examine some of these reactions.

The First Law and the Energy Story

The first law of thermodynamics is deceptively simple. Students often understand that energy cannot be created or destroyed. Yet, when describing an energy story of a process they often make the mistake of saying things such as "energy is produced from the transfer of electrons from atom A to atom B." While most of us will understand the point the student is trying to make, the wrong words are being used. Energy is not made or produced; it is simply transferred. To be consistent with the first law, when telling an energy story, make sure you try to track explicitly all the places that ALL the energy in the system at the start of a process goes by the end of a process.

The Second Law of Thermodynamics

An important concept in physical systems is **entropy**. Entropy relates to how energy can be distributed or dispersed within the particles of a system. The Second Law of Thermodynamics states that entropy is always increasing in a system *and* its surroundings (that is, everything inside and outside the system combined).

This idea helps explain the directionality of natural phenomena. The notion is that the directionality comes from the tendency for energy in a system to move towards a state of maximal dispersion. The Second Law, therefore, implies that in any transformation, we should look for an overall increase in entropy (or dispersion of energy), somewhere. As dispersion of energy in a system or its surroundings increases, the ability of the energy to be directed towards work decreases.

Keep in mind: you will find many examples in which the entropy of a system decreases *locally*. However, according to the Second Law, the entropy of the entire universe can *never* decrease. This must mean that there is an equal or greater increase in entropy somewhere else in the surroundings (most likely in a closely connected system) that compensates for the local decrease.

We associate the four scenarios below with increasing entropy of the system. Try to think of specific examples for when:

a. the system gains energy;
b. a change of state occurs from solid to liquid to gas;

c. a mixing of substances occurs;
d. the number of particles increases during a reaction.

Possible NB Discussion Point

Justify or refute the following statement: *"Biological systems are an exception to the Second Law of Thermodynamics, since cells are known to arrange themselves into very highly organized structures (think: tissues, organs, etc.) rather than into a more disordered state."* Make sure to check out what your peers are saying -- do you agree or disagree with their position and/or their rationale?

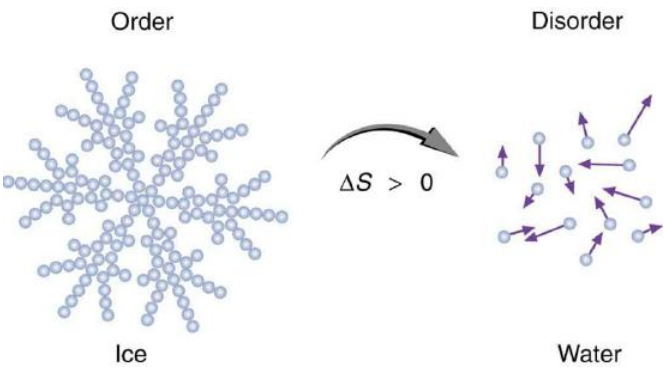


Figure 1. *An increase in disorder can happen in different ways. An ice cube melting on a hot sidewalk is one example. Here, ice is displayed as a snowflake, with organized, structured water molecules forming the snowflake. Over time, the snowflake will melt into a pool of disorganized, freely moving water molecules. It is common to describe entropy as a measure of order as a way to simplify the more concrete description relating entropy to the number of states in which energy can be dispersed in a system. While the idea of measuring order to define entropy has some flaws, it is sometimes a useful, if imperfect, proxy. (Source)*

If we consider the first and second laws together, we come to a useful conclusion. Whenever energy is transferred or redistributed within a system, entropy must increase. This increase in entropy is related to how "useful" the energy is to do work. Recall again that this energy becomes less and less available as entropy increases.

Keep in mind: you will find many examples in which the entropy of a system decreases *locally*. However, according to the Second Law, the entropy of the entire universe can *never* decrease. This must mean that there is an equal or greater increase in entropy somewhere else in the surroundings (most likely in a closely connected system) that compensates for the local decrease.

We conclude that while all the energy must be conserved, if the required change increases entropy, it means that some energy will become distributed in a way that makes it less useful for work. Most times, particularly in biology, some increase in entropy can be chalked up to a transfer of energy to heat in the environment.

This page titled 6.5: Thermodynamics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Marc Facciotti.

- **6.2: Thermodynamics** by Marc Facciotti has no license indicated.

6.6: Energy in Metabolism

Source: BiochemFFA_5_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Living organisms are made up of cells, and cells contain a horde of biochemical components. Living cells, though, are not random collections of these molecules. They are extraordinarily organized or "ordered". By contrast, in the nonliving world, there is a universal tendency to increasing disorder. Maintaining and creating order in cells takes the input of energy. Without energy, life is not possible.

Oxidative Energy

The primary mechanism used by non-photosynthetic organisms to obtain energy is oxidation and carbon is the most commonly oxidized energy source. The energy released during the oxidative steps is “captured” in ATP and can be used later for energy coupling. The more reduced a carbon atom is, the more energy can be realized from its oxidation. Fatty acids are highly reduced, whereas carbohydrates are moderately so. Complete oxidation of both leads to carbon dioxide, which has the lowest energy state. Conversely, the more oxidized a carbon atom is, the more energy it takes to reduce it.

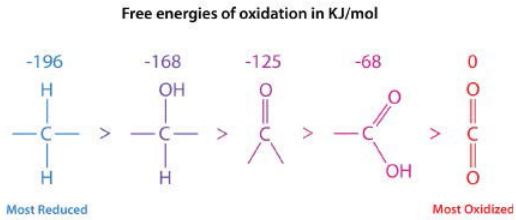


Figure 6.6.1: Five oxidation states of carbon. Image by Pehr Jacobson

In the series shown in Figure 6.6.1, the most reduced form of carbon is on the left. The energy of oxidation of each form is shown above it. The reduction states of fatty acids and carbohydrates can be seen by their formulas.

- Palmitic acid: $C_{16}H_{34}O_2$
- Glucose: $C_6H_{12}O_6$

Palmitic acid only contains two oxygens per sixteen carbons, whereas glucose has six oxygen atoms per six carbons. Consequently, when palmitic acid is fully oxidized, it generates more ATP per carbon (128/16) than glucose (38/6). It is because of this that we use fat (contains fatty acids) as our primary energy storage material.

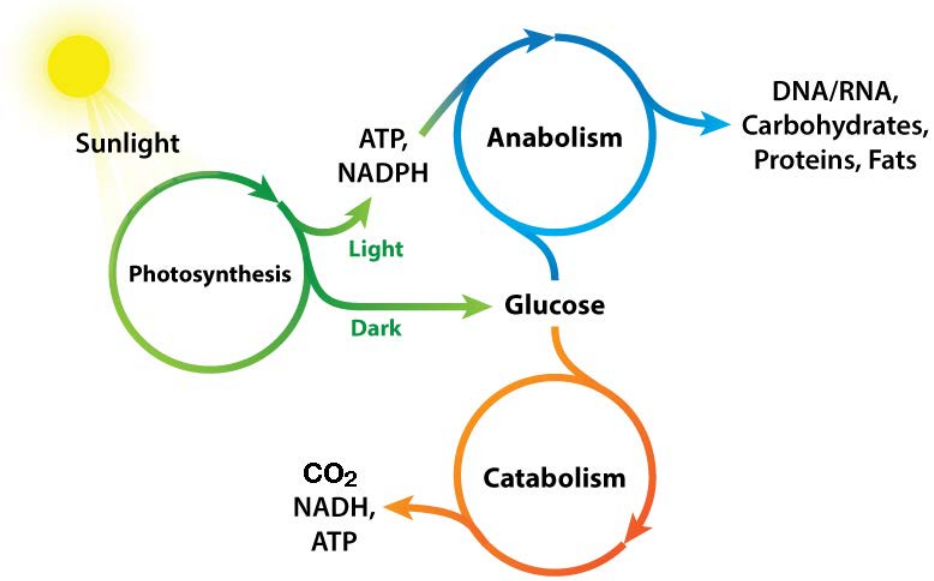


***Figure** 6.6.2: Photosynthesis: The primary source of biological energy. Image by Aleia Kim*

Oxidation vs. Reduction in Metabolism

Biochemical processes that break things down from larger to smaller are called *catabolic processes*. Catabolic processes are often oxidative in nature and energy releasing. Some, but not all, of that energy is captured as ATP. If not all of the energy is captured as ATP, what happens to the rest of it? The answer is simple. It is released as heat and it is for this reason we get hot when we exercise.

By contrast, synthesizing large molecules from smaller ones (for example, making proteins from amino acids) is referred to as anabolism. Anabolic processes are often reductive in nature (Figures 5.3 & 5.4) and require energy input. By themselves, they would not occur, as they are reversing oxidation and decreasing entropy (making many small things into a larger one). To overcome this energy barrier, cells must expend energy. For example, if one wishes to reduce CO_2 to carbohydrate, energy must be used to do so. Plants do this during the dark reactions of photosynthesis (Figure 6.6.3). The energy source for the reduction is ultimately the sun. The electrons for the reduction come from water, and the CO_2 is removed from the atmosphere and gets incorporated into a sugar.

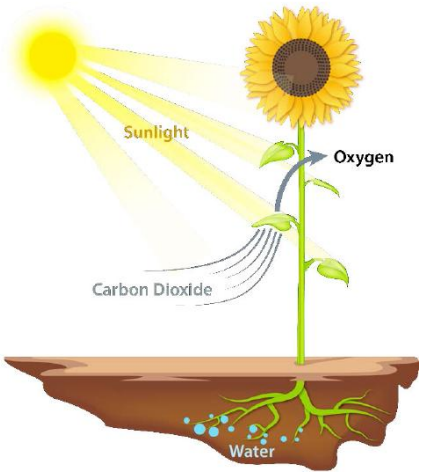


Figure 6.6.3: Movement of biological energy. Image by Aleia Kim

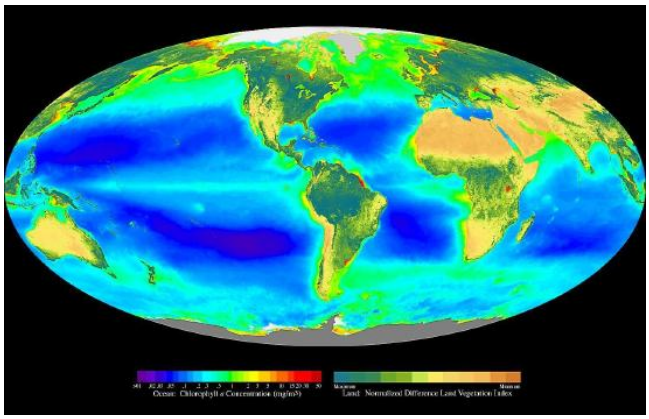


Figure 6.6.4: Photosynthesis as measured by chlorophyll concentration

Energy Coupling

The synthesis of the many molecules needed by cells needs the input of energy to occur. Cells overcome this energy obstacle by using ATP to “drive” the reaction (Figure 6.6.6). The energy needed to drive reactions is harvested in very controlled conditions in enzymes. This involves a process called ‘coupling’. Coupled reactions rely on linking an energetically favorable reaction (i.e., one with a negative ΔG°’) with the reaction requiring an energy input, which has a positive ΔG°’. As long as the overall ΔG°’ of the two reactions together is negative, the reaction can proceed. Hydrolysis of ATP is a very energetically favorable reaction that is commonly linked to many energy requiring reactions in cells. Without the hydrolysis of ATP (or GTP, in some cases), the reaction would not be feasible.

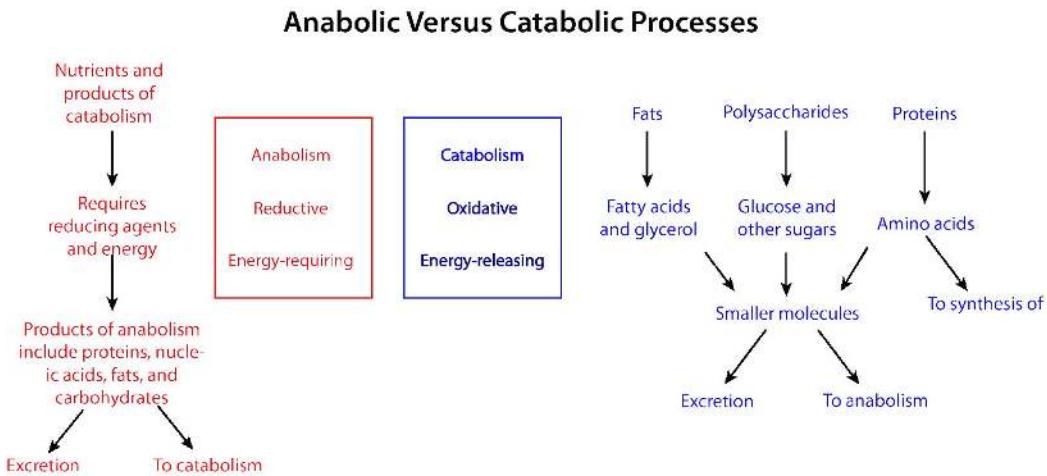


Figure 6.6.5: Synthesis and breakdown pathways in metabolism. Image by Pehr Jacobson

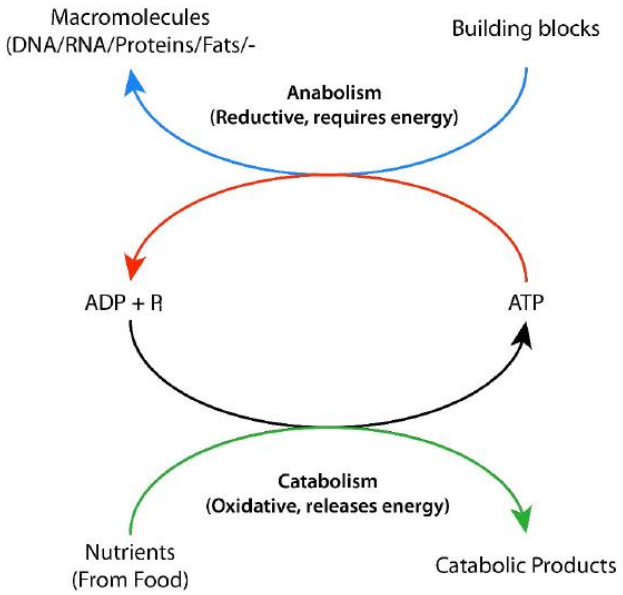


Figure 6.6.6: Cycling of biological energy via ADP and ATP. Image by Pehr Jacobson

Entropy and energy

Most students who have had some chemistry know about the Second Law of Thermodynamics with respect to increasing disorder of a system. Cells are very organized or ordered structures, leading some to mistakenly conclude that life somehow violates the second law. In fact, that notion is incorrect. The second law doesn't say that entropy always increases, just that, left alone, it tends to do so, in an isolated system. Cells are not isolated systems, though, in that they obtain energy, either from the sun, if they are autotrophic, or food, if they are heterotrophic.

To counter the universal tendency towards disorder on a local scale requires energy. As an example, take a fresh deck of cards which is neatly aligned with Ace-King-Queen . . . . 4,3,2 for each suit. Throw the deck into the air, letting the cards scatter. When you pick them up, they will be more disordered than when they started. However, if you spend a few minutes (and expend a bit of energy), you can reorganize the same deck back to its previous, organized state. If entropy always increased everywhere, you could not do this. However, with the input of energy, you overcame the disorder. This illustrates an important concept: the cost of fighting disorder is energy.

Biological energy

There are, of course, other reasons that organisms need energy. Muscular contraction, synthesis of molecules, neurotransmission, signaling, thermoregulation, and subcellular movements are examples. Where does this energy come from? The currencies of energy are generally high-energy phosphate-containing molecules. ATP is the best known and most abundant, but GTP is also an important energy source (energy source for protein synthesis). CTP is involved in synthesis of glycerophospholipids and UTP is used for synthesis of glycogen and other sugar compounds. In each of these cases, the energy is in the form of potential chemical energy stored in the multi-phosphate bonds. Hydrolyzing those bonds releases the energy in them.

Of the triphosphates, ATP is the primary energy source, acting to facilitate the synthesis of the others by action of the enzyme NDPK. ATP is made by three distinct types of phosphorylation – oxidative phosphorylation (in mitochondria), photophosphorylation (in chloroplasts of plants), and substrate level phosphorylation (in enzymatically catalyzed reactions).

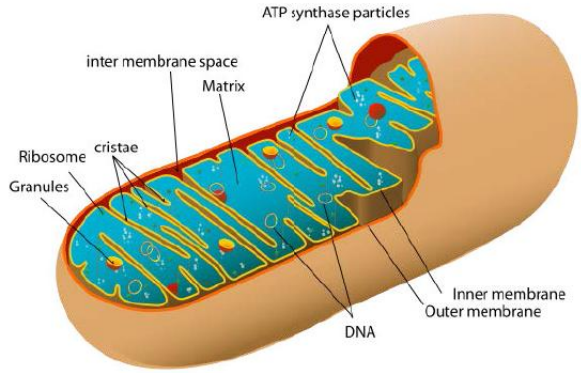


Figure 6.6.7: Mitochondrion

Gibbs free energy in Biology

ATP is generally considered the "storage battery" of cells (See also 'Molecular Battery Backups for Muscles HERE). In order to understand how energy is captured, we must first understand Gibbs free energy and in doing so, we begin to see the role of energy in determining the directions chemical reactions take.

Gibbs free energy may be thought of as the energy available to do work in a thermodynamic system at constant temperature and pressure. Mathematically, the Gibbs free energy is given as:

$$G = H – TS \tag{6.6.1}$$

where H is the enthalpy, T is the temperature in Kelvin, and S is the entropy. At standard temperature and pressure, every system seeks to achieve a minimum of free energy. Thus, increasing entropy, S, will reduce Gibbs free energy. Similarly, if excess heat is available (reducing the enthalpy, H), the free energy can also be reduced.

Cells must work within the laws of thermodynamics, as noted, so all of their biochemical reactions, too, are ruled by these laws. Now we shall consider energy in the cell. The change in Gibbs free energy (ΔG) for a reaction is crucial, for it, and it alone, determines whether or not a reaction goes forward.

$$\Delta G = \Delta H – T\Delta S. \tag{6.6.2}$$

There are three cases

- ΔG < 0: the reaction proceeds as written
- ΔG = 0: the reaction is at equilibrium
- ΔG > 0: the reaction runs in reverse

For a reaction

$$a\,\mathrm{A} \rightleftharpoons b\,\mathrm{B} \tag{6.6.3}$$

(where 'a' and 'b' are integers and A and B are molecules) at pH 7, ΔG can be determined by the following equation,

$$\Delta G = \Delta G^{o\prime} + RT\ln\left(\frac{[B]^b}{[A]^a}\right) \tag{6.6.4}$$

For multiple substrate reactions, such as

$$a\,\mathrm{A} + c\,\mathrm{C} \rightleftharpoons b\,\mathrm{B} + d\,\mathrm{D} \tag{6.6.5}$$

$$\Delta G = \Delta G^{o\prime} + RT\ln\left(\frac{[B]^b[D]^d}{[A]^a[C]^c}\right) \tag{6.6.6}$$

The ΔG°' term is called the change in Standard Gibbs Free energy, which is the change in energy that occurs when all of the products and reactants are at standard conditions and the pH is 7.0. It is a constant for a given reaction.

In simple terms, we can collect all of the terms of the numerator together and call them {Products} and all of the terms of the denominator together and call them {Reactants},

$$\Delta G = \Delta G^{\circ\prime} + RT\ln(\frac{\{\text{Products}\}}{\{\text{Reactants}\}}) \tag{6.6.7}$$

For most biological systems, the temperature, T, is a constant for a given reaction. Since ΔG°' is also a constant for a given reaction, the ΔG is changed almost exclusively as the ratio of {Products}/{Reactants} changes.

Importance of ΔG°'

If one starts out at standard conditions, where everything except protons is at 1M, the RTln({Products}/{Reactants}) term is zero, so the ΔG°' term equals the ΔG, and the ΔG°' determines the direction the reaction will take (only under those conditions). This is why people say that a negative ΔG°' indicates an energetically favorable reaction, whereas a positive ΔG°' corresponds to an unfavorable one.

Increasing the ratio of {Products}/{Reactants} causes the value of the natural log (ln) term to become more positive (less negative), thus making the value of ΔG more positive. Conversely, as the ratio of {Products}/{Reactants} decreases, the value of the natural log term becomes less positive (more negative), thus making the value of ΔG more negative.

System response to stress

Intuitively, this makes sense and is consistent with Le Chatelier's Principle – a system responds to stress by acting to alleviate the stress. If we examine the ΔG for a reaction in a closed system, we see that it will always move to a value of zero (equilibrium), no matter whether it starts with a positive or negative value.

Another type of free energy available to cells is that generated by electrical potential. For example, mitochondria and chloroplasts partly use Coulombic energy (based on charge) from a proton gradient across their membranes to provide the necessary energy for the synthesis of ATP. Similar energies drive the transmission of nerve signals (sodium and potassium gradients) and the movement of some molecules in secondary active transport processes across membranes (e.g., H+ differential driving the movement of lactose). From the Gibbs free energy change equation,

$$\Delta G = \Delta H – T\Delta S \tag{6.6.8}$$

it should be noted that an increase in entropy will help contribute to a decrease in ΔG. This happens, for example when a large molecule is being broken into smaller pieces or when the rearrangement of a molecule increases the disorder of molecules around it. The latter situation arises in the hydrophobic effect, which helps drive the folding of proteins.

Chemical and electrical potential

It is said that absence makes the heart grow fonder. We won't tackle that philosophical issue here, but we will say that separation provides potential energy that cells can and do harvest. The lipid bilayer of cell and (in eukaryotic cells) organelle membranes provide the necessary barrier for separation.

Impermeable to most ions and polar compounds, biological membranes are essential for processes that generate cellular energy. Consider Figure 5.8. A lipid bilayer separates two solutions with different concentrations of a solute. There is a greater concentration of negative ions in the bottom and a greater concentration of positive ions on the top.

Whenever there is a difference in concentration of molecules across a membrane, there is said to be a concentration gradient across it. A difference in concentration of ions across a membrane also creates a charge (or electrical) gradient. Because there is a difference in both the chemical concentration of the ions and in the charge on the two sides of the membrane, this is described as an electrochemical gradient (Figures 5.8 -5.10).

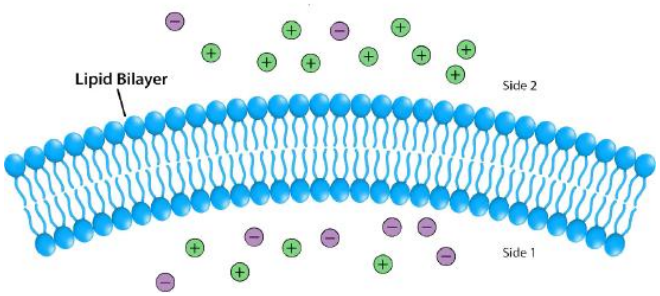


Figure 5.8: Differences in ion concentration across a membrane give rise to chemical and electrical gradients. Image by Pehr Jacobson

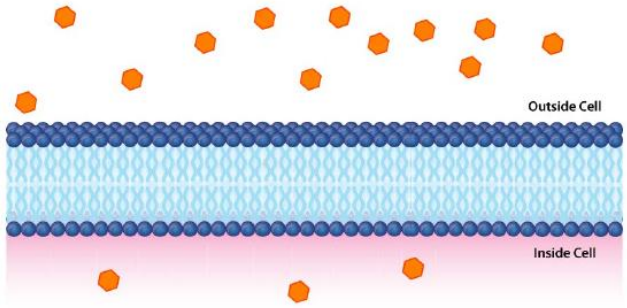


Figure **6.6.9**: A chemical gradient. Image by Aleia Kim

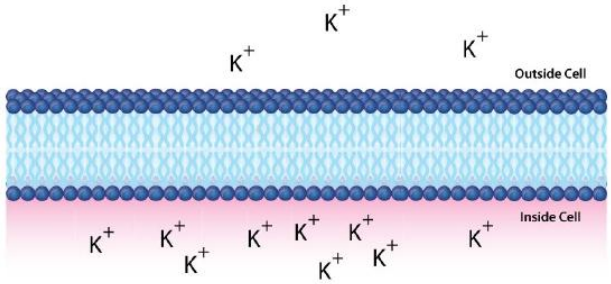
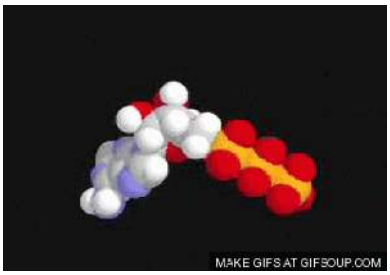


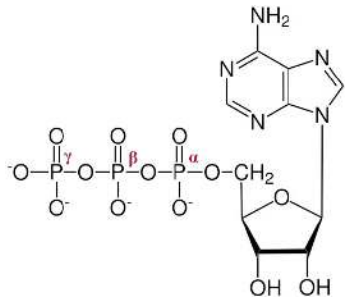
Figure **6.6.10**: An electrochemical gradient of potassium ions. Image by Aleia Kim

Energy Storage in Triphosphates



Movie 5.1: ATP: The fuel of the cell

Formation of triphosphates, like ATP, is essential to meeting the cell's energy needs for synthesis, motion, and signaling. In a given day, an average human body makes and breaks down more than its weight in triphosphates. This is especially remarkable considering that there is only about 250 g of the molecule present in the body at any given time. Energy in ATP is released by hydrolysis of a phosphate from the molecule.


NH2
N
N
N
N
γ
β
α
O
O
O

cell additional, needed energy.

It is important to note that enzymes cannot make reactions happen that are energetically unfavorable. Enzymes speed reactions, but do not change their direction. Cells are thus bound by the rules of Gibbs free energy. So, how do energetically unfavorable reactions happen in a cell?

Reaction coupling

Reactions that are energetically unfavorable, can be made favorable by coupling them with the hydrolysis of ATP, a very energetically favorable reaction. There are numerous parallels in the "real world." Movement of automobiles is energetically unfavorable, but coupling movement of the automobile to oxidation of gasoline makes an unfavorable process favorable. Another approach to making an unfavorable reaction favorable is to manipulate the concentration of reactants and products. Consider the reaction below, which occurs in pyrimidine nucleotide metabolism:

$$\text{orotate} + \text{PRPP} \rightleftharpoons \text{OMP} + \text{PP}_i$$

The ΔG°' for this reaction is -0.8 kJ/mol, meaning that if one starts with equal concentrations of reactants and products, at equilibrium, there will be a small excess of products. In the cell, however, this reaction moves strongly to the right (ΔG = very negative). Given that the ΔG°' is very close to zero, a very negative ΔG can only occur if the concentrations of reactants and products are altered, since

$$\Delta G = \Delta G^{\circ\prime} + RT\ln\left(\frac{[\text{OMP}][\text{PP}_i]}{[\text{Orotate}][\text{PRPP}]}\right) \tag{6.6.9}$$

Manipulation is exactly what happens here. The key item whose concentration is adjusted in this reaction is the pyrophosphate (PPi). This is possible because cells contain an enzyme called pyrophosphorylase that catalyzes the following reaction

$$\text{PP}_i + \text{H}_2\text{O} \rightleftharpoons 2\ \text{P}_i$$

Hydrolysis of pyrophosphate is very energetically favored, causing the PPi produced in the reaction to be quickly hydrolyzed. As a result, the concentration of PPi in the cell is kept very low. A low concentration of a product (PPi) causes the natural log (ln) term of the orotate equation to become more negative, driving the ΔG term for the overall reaction to become much more negative.

Pushing and pulling

Reactions that yield pyrophosphate as a product are produced in the synthesis of DNA and RNA, as well as many other molecules. As shown in the previous example, this pyrophosphate is rapidly hydrolyzed, causing the overall reaction to move in the direction of pyrophosphate production. When reactants are removed/reduced in a metabolic reaction to decrease the concentration of a product, we say that the reaction is "pulled", to represent the increase in the forward reaction as a result of product depletion.

Pushing happens when reactants in a reaction are added/increased. This too has the effect of reducing the ΔG of a reaction and making it more favorable because the ratio of [Products]/[Reactants] is decreased with increasing [Reactants]. Pushing and pulling of reactions are additional tools for cells to overcome energy barriers, just like coupling energetically favorable processes to energetically unfavorable ones.

This page titled 6.6: Energy in Metabolism is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **6.3: Energy in Metabolism** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

CHAPTER OVERVIEW

7: Enzyme Kinetics

7.1: Enzyme Kinetics
7.2: Enzyme Activity
7.2.1: How Enzymes Work
7.2.2: Kinetics without Enzymes
7.2.3: Kinetics with Enzymes
7.2.4: Enzyme Inhibition
7.2.5: Enzymatic Reaction Mechanisms
7.2.6: Enzymes and Protein Regulation
7.2.7: Ribozymes - RNA Enzymes
7.2.8: Cofactors and Catalysis - A Little Help From My Friends
7.3: Basic Principles of Catalysis
7.4: Derivation of Michaelis-Menten equation
7.5: Binding - The First Step Towards Protein Function

7: Enzyme Kinetics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7.1: Enzyme Kinetics

Unlike uncatalyzed (but readily occurring) reactions, in which the rate of the reaction is dependent only on the concentration of the reactants, the rate of enzyme-catalyzed reactions is limited by the number of enzyme molecules available. This maximal rate of turnover from substrate to product is a function of the speed of the enzyme and the number of enzyme molecules. V_{max}, this theoretical maximal rate or reaction, is approached when there is such a high concentration of substrate molecules that not only is every available enzyme at a given time occupied, but as soon as an enzyme finishes converting substrate to product, it immediately binds a new substrate. Another term, Km, is related to V_{max} in that K_m (the Michaelis constant) is the concentration of substrate at which half-maximal reaction rate $V_{max}/2$ occurs. These two terms are related in the Michaelis-Menten equation, which describes the reaction rate v with respect to the substrate concentration [S].

$$v = \frac{V_{\max}[\mathrm{S}]}{\mathrm{K_M} + [\mathrm{S}]} \tag{7.1.1}$$

> The Michaelis-Menten equation was derived by Leonor Michaelis and his graduate student Maud Menten in 1913, based on work by Victor Henri, and is applicable only to simple enzyme kinetics in which there is only one substrate that is changed immediately to a product during the reaction without forming any intermediate compound, the enzyme in question shows no allostericity, and the reaction is unidirectional.
>
> It should be noted that the reaction rate *v* is actually the initial reaction rate at a particular substrate concentration, and is sometimes denoted *vo*. Naturally, as the reaction continues, the substrate concentration decreases, along with the reaction rate.

The Michaelis-Menten equation assumes a simple reaction of the form:

$$\mathrm{E} + \mathrm{S} \rightleftharpoons \mathrm{ES} \longrightarrow \mathrm{E} + \mathrm{P} \tag{7.1.2}$$

where E is an enzyme, S is the substrate, and P is the product. Note the formation of the intermediate enzyme-substrate complex, ES, which is a transition state (recall Figure 7.1.2) in which the substrate is unstable and associated with the enzyme. In fact, ES could as easily be considered EP, since this state is essentially the tipping point between the conversion from substrate to product. In this construction, the **Michaelis constant**, K_M, of an enzyme-catalyzed reaction is $(k_2 + k_3)/k_1$. That is the rate of ES dissociation over the rate of ES association. K_M, of course, varies not only depending on the enzyme, but also with respect to the identity of the substrate. Some enzymes can work with multiple substrates, and the K_M of that enzyme for the different substrates is usually different. Because the saturation curve in Figure 7.1.5 can be difficult to work with, linearizations of the Michaelis-Menten equation (Equation 7.1.1) were developed. The most common is the **double reciprocal plot**, better known as the Lineweaver-Burk plot. On this type of graphical representation of enzyme kinetics, the reciprocal of the substrate concentration is plotted against the reciprocal of the reaction velocity. This generates a line in which the x-intercept is then $-1/K_m$, the y-intercept is $1/V_{max}$, and the slope of the line is K_m/V_{max}.

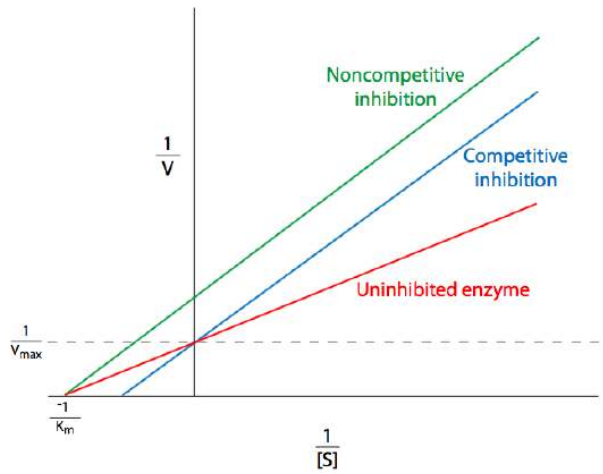


Figure 7.1.7. Lineweaver-Burk plot of enzyme kinetics and the effect of competitive and non-competitive inhibitors at constant concentrations.

The Lineweaver-Burk plot

Obtaining V_{max} and K_m from a direct plot of v against [S] can be difficult because even at very high substrate concentrations, experimental data may still be significantly under the V_{max}. This leads to underestimation of the V_{max}.

The Lineweaver-Burk plot addresses this concern, but has some shortcomings of its own. Because it is easier to obtain data at high concentrations, most of the data points are near 0, and fewer data points are available further out (to the right of the graph). Because these are reciprocals, under these low [S] conditions, small errors in measured values of \(v\) turn into large errors in 1/v, and therefore large errors in K_M and V_{max}. This is evident on examination of the Lineweaver-Burk equation:

$$\frac{1}{v} = \left(\frac{K_M}{V_{\max}}\right)\frac{1}{[S]} + \frac{1}{V_{\max}} \tag{7.1.3}$$

This page titled 7.1: Enzyme Kinetics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by E. V. Wong via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Enzyme Kinetics** by E. V. Wong is licensed CC BY-NC-SA 3.0. Original source: https://www.axopub.com/wp01/2012/02/26/download-cells-molecules-and-mechanisms/.

SECTION OVERVIEW

7.2: Enzyme Activity

7.2.1: How Enzymes Work

7.2.2: Kinetics without Enzymes

7.2.3: Kinetics with Enzymes

7.2.4: Enzyme Inhibition

7.2.5: Enzymatic Reaction Mechanisms

7.2.6: Enzymes and Protein Regulation

7.2.7: Ribozymes - RNA Enzymes

7.2.8: Cofactors and Catalysis - A Little Help From My Friends

This page titled 7.2: Enzyme Activity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

7.2.1: How Enzymes Work

Learning Goals (ChatGPT, 1/30/25)

- **Understanding Reaction Energetics and Transition States**
  - Describe why uncatalyzed reactions (such as ester hydrolysis) are inherently slow due to high activation energies and the formation of high-energy, charge-separated transition states.
  - Analyze energy diagrams to identify how the transition state's energy and structure (often resembling an unstable intermediate) influence reaction rates.
- **Transition State Stabilization and Its Role in Catalysis**
  - Explain how any factor that stabilizes the charges developed in the transition state (e.g., protonation, deprotonation, or metal ion coordination) lowers the activation energy and catalyzes the reaction.
  - Apply transition state theory and Pauling's hypothesis to quantitatively compare the binding affinity of an enzyme for the substrate versus its transition state.
- **Acid-Base Catalysis Strategies**
  - Differentiate between specific acid (or base) catalysis and general acid (or base) catalysis in terms of reaction mechanism and pH dependency.
  - Illustrate, using mechanistic diagrams, how general acid catalysis (via proton donation) and general base catalysis (via proton abstraction) reduce the energy barrier in ester hydrolysis.
- **Metal Ion (Electrostatic) Catalysis**
  - Discuss how metal ions (such as Cu^{2+}, Zn^{2+}, or Fe^{3+}) stabilize negative charges in transition states through coordination, thereby enhancing reaction rates.
  - Evaluate examples (e.g., decarboxylation of β-keto acids and carbonic anhydrase) to understand the role of metal ions in enzyme catalysis.
- **Covalent (Nucleophilic) Catalysis**
  - Describe how a nucleophilic catalyst (e.g., pyridine) can form a covalent intermediate with a substrate, effectively creating a lower-energy pathway for the reaction.
  - Use energy diagrams to compare uncatalyzed versus covalently catalyzed reactions, noting the formation of additional intermediates and transition states.
- **Intramolecular Catalysis and Effective Concentration**
  - Explain how intramolecular reactions (e.g., in aspirin derivatives or phenylsuccinate) benefit from increased effective concentration, leading to dramatic rate enhancements compared to analogous intermolecular reactions.
  - Quantitatively compare intramolecular versus intermolecular rate constants to illustrate the concept of effective molarity.
- **Quantitative Deconstruction of Enzyme Catalysis**
  - Integrate the concepts of transition state binding, intramolecular versus intermolecular reactions, and the roles of various catalytic strategies (acid-base, metal ion, covalent) to analyze enzyme rate enhancements.
  - Derive and interpret relationships between equilibrium constants for substrate and transition state binding, and relate these to overall catalytic efficiency.
- **Design and Use of Transition State Analogs (Abzymes)**
  - Discuss how stable transition state analogs can be used as inhibitors or to generate antibodies (abzymes) that mimic enzyme catalysis.
  - Evaluate case studies where transition state analogs are synthesized to target specific enzyme active sites.
- **Asymmetric Catalysis and Enantioselectivity in Enzyme Reactions**

- Describe how enzymes achieve asymmetric catalysis by selectively binding one enantiomer over another, resulting in the formation of a single stereoisomer of product (e.g., in the reaction catalyzed by triose phosphate isomerase).
- Compare enzyme catalysis to small-molecule organocatalysis (such as L-proline catalyzed aldol reactions) to understand the broader implications of chiral catalysis in biological and synthetic chemistry.

By mastering these goals, students will gain a deep understanding of how enzymes and small molecule catalysts lower activation energies through multiple strategies, the quantitative relationships governing these processes, and the practical applications of these concepts in both natural and synthetic systems.

In this section, we will explore chemical and physical factors that speed up reactions and begin to relate these effects to reactions catalyzed by enzymes. We will see that enzymes employ various chemical strategies to increase the rates of reactions, in addition to physical ones like reactant proximity and the introduction of strain. This can result in reactions 10 million (or more!) times faster than the uncatalyzed reaction. To put this 10 million-fold rate enhancement in perspective, if the catalyzed reaction takes one second, the uncatalyzed one would take nearly four months!

This chapter section has been written by Kristen Procko and Henry Jakubowski.

Reactions in solutions that are not catalyzed are slow. Consider the hydrolysis of an ester in water, illustrated in Figure **7.2.1.1**. The ester is stabilized by resonance and is therefore weakly electrophilic; thus, attack by weakly nucleophilic water is a slow process. Examining the transition state, we can see that charge development and separation occur in the transition state for the uncatalyzed reaction, resulting in an intermediate (**P**) with both positive and negative charges.

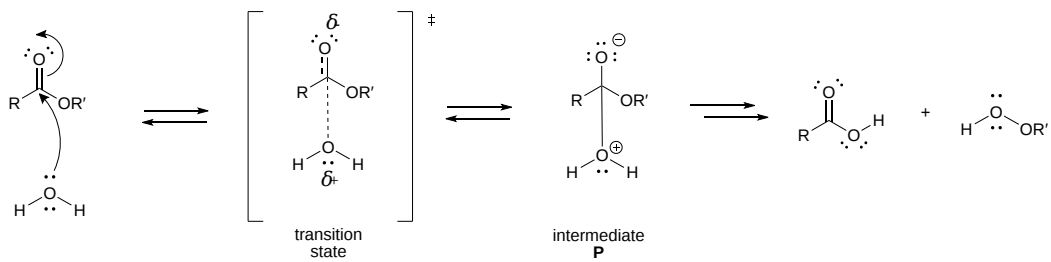


Figure **7.2.1.1**: Charge development in the transition state during ester hydrolysis

When bonds are made or broken, charged intermediates are often formed, which are higher in energy than the reactants. Consider the energy diagram for the first step of a generic endergonic reaction, shown in **7.2.1.2**. The transition state is closer in energy to the intermediate **P** than to the reactant **R.** Therefore, the TS more closely resembles **P** than the starting reactants. Applying this analysis to the ester hydrolysis reaction from Figure **7.2.1.1**, the transition state is closer in energy to the charge-separated intermediate **P** and, therefore, more closely resembles the charge-separated species. In this example, the intermediate is higher in energy than the reactants; thus, the transition state is *even higher* in energy than the intermediate.

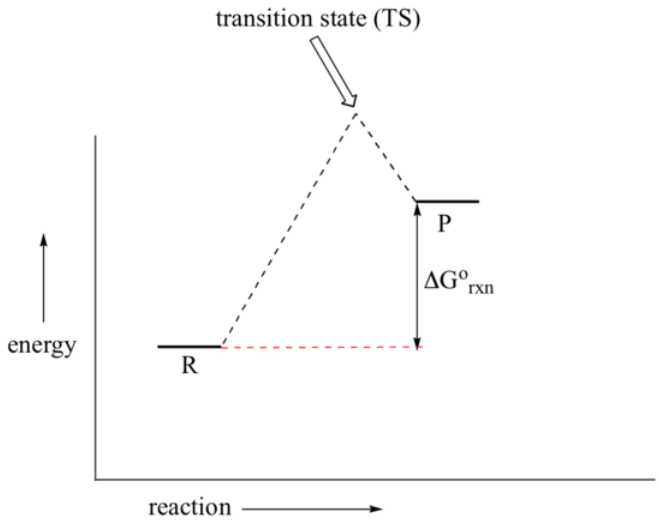


Figure 7.2.1.2: An energy diagram for an endergonic reaction (Image modified from: Chemistry LibreTexts, 6.9)

Anything that can stabilize the charges on the intermediate will also stabilize the developing charges in the transition state. This lowers the energy of the transition state and catalyzes the reaction. This section will investigate the mechanism underlying the catalysis by small molecules of chemical reactions. Presumably, biological macromolecular catalysts (like protein enzymes) will use similar mechanisms in their catalytic effects (which will be discussed in the next section).

Catalysts, including enzymes, can stabilize transition states in at least five ways.

7.2.1.1: Chemical Strategies for Rate Enhancement

7.2.1.1.1: General Acid and Base Catalysis

Considering intermediate **P** in Figure 7.2.1.1, we can envision two strategies to reduce the charge separation: the negative charge on the anionic oxygen could be protonated, or the positive charge on the cationic oxygen could be removed by deprotonation. If the reaction is pH dependent, and the reaction rate solely depends on hydronium ion concentration, $[H_3O^+]$, then **specific acid catalysis** is operative. **Specific acid catalysis** occurs when the hydronium ion concentration is the sole factor determining the reaction rate, and the concentration of any buffer components present in the solution does not influence the rate. In other words, the reaction rate depends *specifically* on the concentration of hydronium ions. **Specific base catalysis** occurs when the reaction rate depends solely on the hydroxide ion concentration and is again independent of any buffer components in the solution.

By contrast, **general acid catalysis** occurs when the reaction is not solely dependent on the $[H_3O^+]$ concentration; that is, the concentration of a buffer component influences the reaction rate. With general acid catalysis, the charge separation in the transition state is decreased by the donation of a proton to the carbonyl from general acids (e.g., acetic acid or a protonated imidazole ring). Proton donation decreases the developing negative charge in the transition state. In Figure 7.2.1.3, the first step of the ester hydrolysis mechanism is shown via specific acid catalysis alongside the general acid catalysis mechanism using the weak acid, acetic acid.

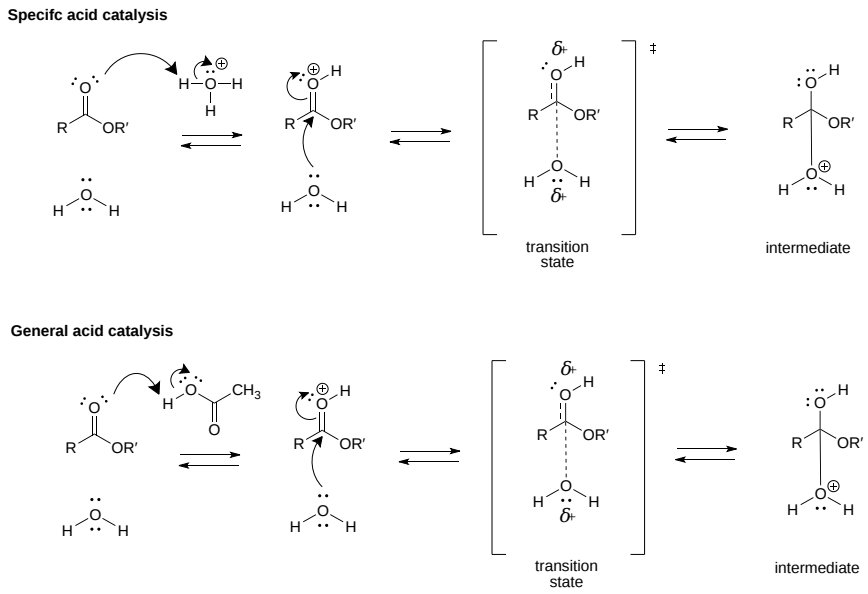


Figure 7.2.1.3: Specific vs. general acid catalysis

Alternatively, the first step of the ester hydrolysis mechanism can be base-catalyzed, increasing the nucleophile's strength. In Figure 7.2.1.1, the attacking water molecule develops a partial positive charge in the transition state as it begins to form a bond with the electrophilic carbon of the carbonyl. In the base-catalyzed mechanism shown in Figure 7.2.1.4, hydroxide becomes the nucleophile in the specific base-catalyzed mechanism. The energy of the transition state can also be lowered by the presence of a general base (e.g., acetate, a deprotonated imidazole ring). Proton abstraction decreases the developing positive charge.

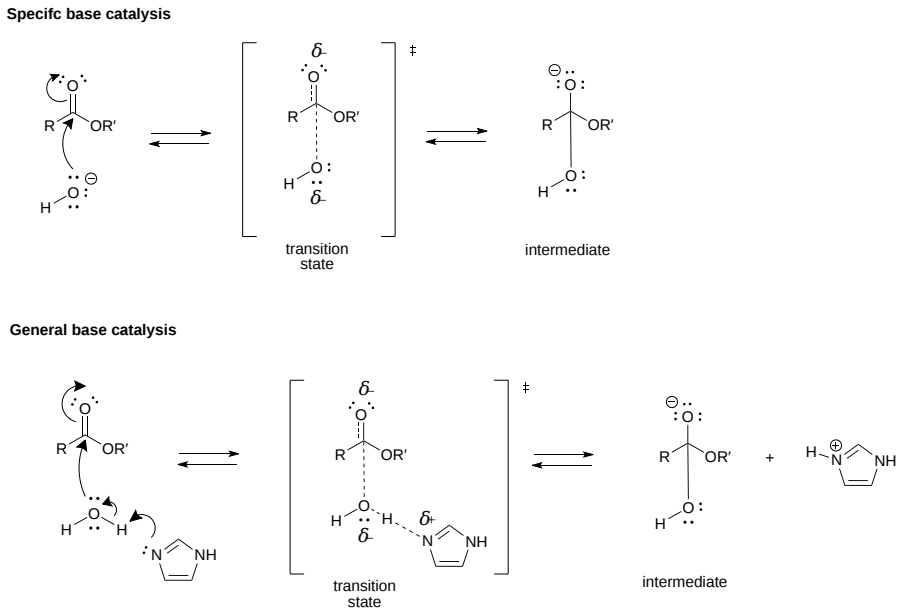


Figure 7.2.1.4: Specific vs. general base catalysis

General acid/base catalysis is common with enzymes because enzymes often use amino acid side chains to promote acid-base reactions within the **active site**, the enzyme region where the chemical reaction occurs. Acetic acid is similar to glutamic and aspartic acid side chains, and the imidazole ring shown in the general base catalysis reaction in Figure 7.2.1.4 is present in the side chain of the amino acid histidine.

7.2.1.1.2: Metal Ion or Electrostatic Catalysis

A metal such as Cu^{2+} or Zn^{2+} can also stabilize the transition state. The metal must be able to bind the charged intermediate and hence the transition state. An oxyanion intermediate formed during the reaction of an electrophilic carbonyl C can interact with a metal, especially when there is an O on an adjacent atom, which can help coordinate the metal ion. This charge stabilization of the developing negative in the transition state and the full negative in the intermediate is often called **electrostatic catalysis.** It is illustrated in the decarboxylation of a β-keto carboxylic acid, Figure 7.2.1.5. Coordination of Cu^{2+} to the β-keto carboxylic acid increases the electrophilicity of the carbonyl, making it a superior electron acceptor, which facilitates decarboxylation. The intermediate enolate formed in the decarboxylation is protonated, giving the more stable ketone as the product.

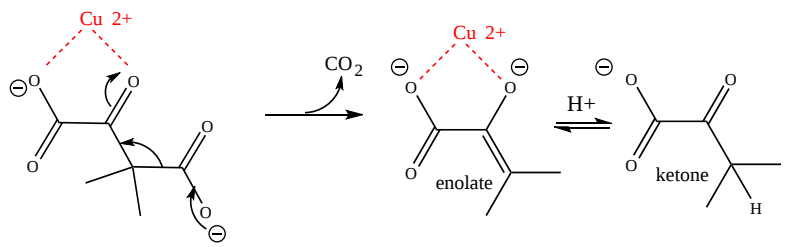


Figure 7.2.1.5: Metal ion catalysis

Electrostatic catalysis is likely to occur in many enzymes since nearly 1/3 of all enzymes require metal ions. A classic example of an enzyme using metal ion catalysis is carboxypeptidase A. Figure 7.2.1.6s shows an interactive iCn3D model of Zn and the inhibitor citric acid bound to carboxypeptidase A (3KGQ). Note the histidine and aspartate amino acid side chains of the active site coordinating to the Zn^{2+} ion, along with the carboxylate group of citrate.

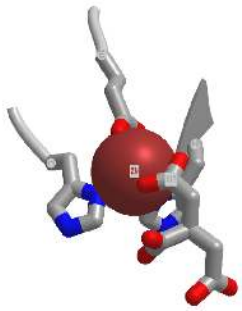


Figure 7.2.1.6: Zn binding to carboxypeptidase A (3KGQ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?pH7tydTANZw5ayh9A

Metals can also act differently. They may coordinate a water molecule and, by further polarizing the H-O bond, increase the acidity of the bound water. For instance, a water molecule in the hexaaquairon(III) ion has a pKa of 9.4, compared to pure water, with a pK_a of 14 (7.2.1.7). The complexed hydroxide is a better nucleophile than bulk water.

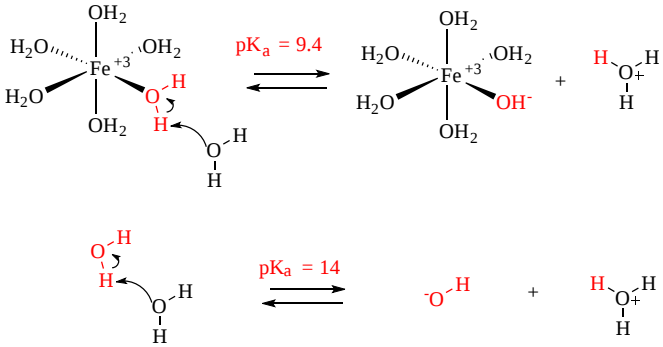


Figure 7.2.1.7: Metal ion decrease of pK_a of coordinated water

Another enzyme that utilizes Zn^{2+} is carbonic anhydrase. It is among the fastest enzymes, with a k_{cat} of 10^6 s^{-1} and a k_{cat}/K_m of 8.3 x 10^7 $M^{-1}s^{-1}$ (reference). It is diffusion-controlled at low substrate (CO_2) concentration and converts one million bound CO_2 per second to HCO_3^-! The Zn^{2+} appears to bind a water molecule and reduce its pK_a such that the bound form is OH^-. This is illustrated in Figure 7.2.1.8, which depicts the local environment of the bound Zn^{2+} (coordinated by three histidine side chains and an OH^-) in the absence (left) and presence (right) of CO_2.

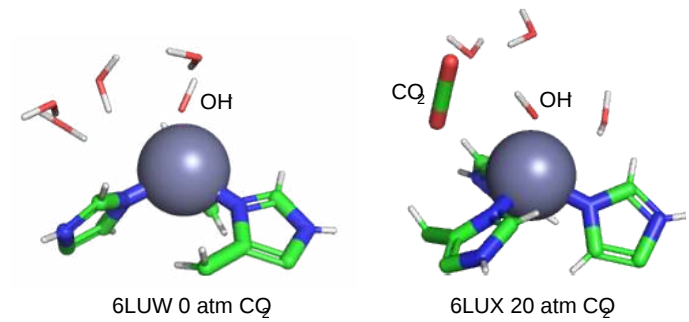


Figure 7.2.1.8: Coordination of OH^- to Zn^{2+} in carbonic anhydrase in the absence (left) and presence (right) of substrate CO_2.

7.2.1.1.3: Covalent or Nucleophilic Catalysis

One way to change the activation energy of the reaction is to change the reaction mechanism in ways that introduce new steps with lower activation energy. As shown in Figure 7.2.1.9, the catalyzed reaction has a new lower energy well, representing the formation of the covalent intermediate, and the activation energy is lowered overall. Formation of the new intermediate results in two transition states, represented by the two high-energy points of the blue line in the plot.

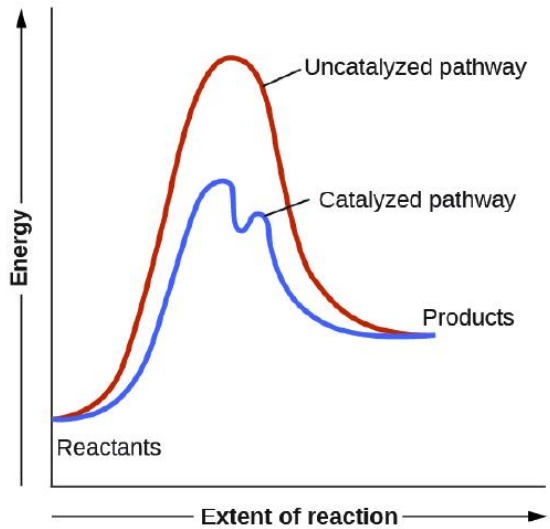


Figure 7.2.1.9: Energy diagram for an uncatalyzed reaction compared to a catalyzed reaction that utilizes covalent catalysis.

A typical way to achieve covalent catalysis is to add a nucleophilic catalyst, which forms a covalent intermediate with the reactant. Figure 7.2.1.8 shows how pyridine (red) acts as a nucleophilic or covalent catalyst in the hydrolysis of an anhydride. The anhydride is very reactive initially, and the charged pyridinium ion intermediate contains a very good leaving group. The desired nucleophile, water, can then interact with the intermediate in a nucleophilic substitution reaction. In these reactions, in general, as long as the nucleophilic catalyst is a better nucleophile than the ultimate nucleophile (usually water), the activation energy is lowered, and the reaction is catalyzed. The nucleophilic catalyst and the original nucleophile usually interact with a carbonyl C in a substitution reaction.

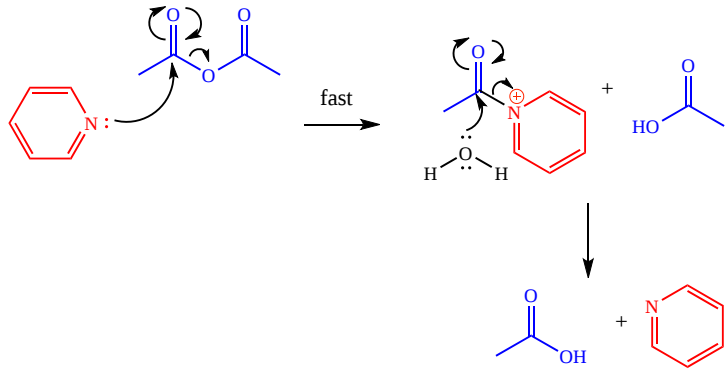


Figure 7.2.1.10: Nucleophilic covalent catalysis by pyridine

7.2.1.1.3.1: Reactions involving iminium ions are a recurring theme in biochemistry

Positively charged nitrogen cations (iminium ions) form as intermediates in many biochemical mechanisms. The iminium ion is a powerful electron acceptor and can promote the cleavage of bonds that would otherwise be difficult to break, such as C–C and C–H bonds. To begin our analysis of how iminium ions promote such cleavage reactions, let's revisit a common carbon-carbon bond cleaving reaction, the decarboxylation of a β-keto acid, which we examined briefly above in Figure 7.2.1.5 with metal ion catalysis.

Under acidic conditions, β-keto acids usually decarboxylate with gentle warming. A cyclic transition state is often invoked, and the presence of the carbonyl of the ketone adjacent to the breaking bond gives the electrons somewhere to go (Figure 7.2.1.11). The decarboxylation product is an enol, which tautomerizes to the more stable ketone. The equilibrium favors the deprotonated carboxylate form under the slightly basic conditions that characterize the medium for most biochemical reactions. The adjacent carbonyl again gives the electrons somewhere to go in the decarboxylation reaction, and under these basic conditions, an enolate is formed. Protonation gives the final product, a ketone.

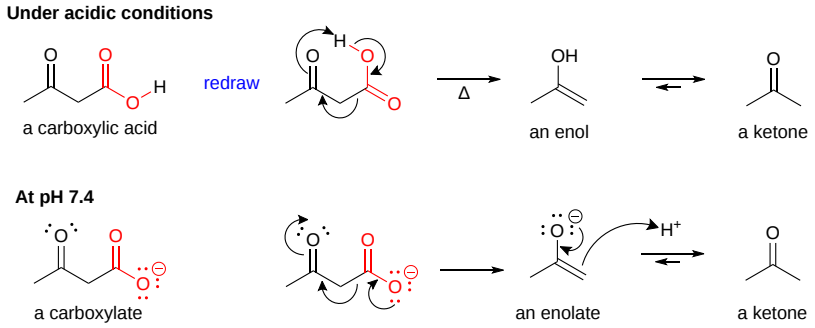


Figure 7.2.1.11: Decarboxylation of a β-keto acid under acidic conditions, and at pH 7.4

In Figure 7.2.1.5, we saw that a metal ion can promote the decarboxylation reaction by interacting with the electron-accepting ketone carbonyl, which makes it even more electrophilic. Another strategy to create a better electron acceptor involves forming a full positive charge on the electron-accepting atom, which can be done by converting the ketone to an iminium ion. Amines react with aldehydes or ketones to form iminium ions. Figure 7.2.1.12 illustrates this strategy involving covalent catalysis. The amine, RNH_2, reacts to form a new intermediate, the iminium ion, with a full positive charge. The protonated nitrogen serves as an excellent electron "sink" for decarboxylation reactions of beta-keto acids.

This **iminium ion** or protonated **Schiff Base** has a pK_a of about 7, so the protonated iminium and deprotonated imine are in equilibrium near pH 7. Figure 7.2.1.12 also illustrates a simple way to view reaction mechanisms. Electrons in chemical reactions can be viewed as flowing from a source (such as a carboxyl group) to a sink (such as an electrophilic carbonyl O or a positively charged N in a Schiff base).

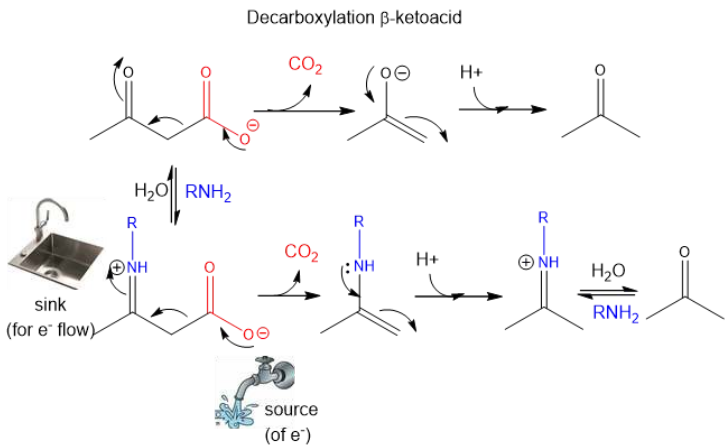


Figure 7.2.1.12: Schiff-base catalyzed decarboxylation of a β-ketoacid - Source/Sink analogy

Acid- and base-catalyzed reaction mechanisms for Schiff base formation are shown in Figure 7.2.1.13. An amine is used as the nucleophilic catalyst, forming the initial addition product, a carbinolamine. The carbinolamine dehydrates, since the free pair of electrons on the N is more likely to be shared with the carbon to form a double bond than electrons from the original carbonyl O, which is more electronegative than the N. If catalyzed by a general acid, an iminium ion and the base-catalyzed reaction form an imine. Near pH 7.4, the imine easily protonated to form a positively charged N at the former carbonyl O center.

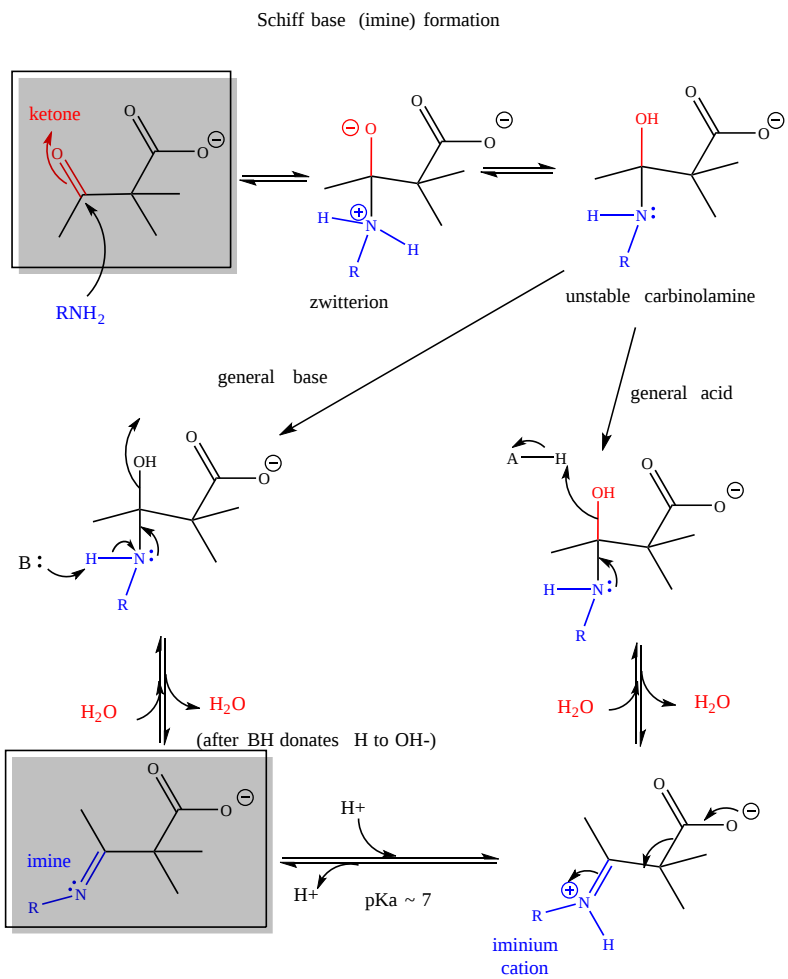


Figure 7.2.1.13: Schiff Base Formation - reaction mechanism

An actual Schiff base intermediate between fructose-1,6-bisphosphate (2FP-400) and Lys 239 from the enzyme fructose bisphosphate aldolase from Leishmania mexicana is shown in Figure 7.2.1.14. Only a single bond between the carbon and nitrogen in the Schiff base is shown.

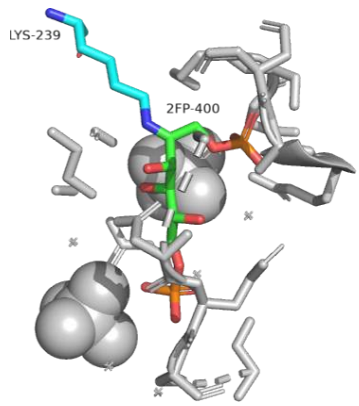


Figure 7.2.1.14: Schiff base intermediate between fructose-1,6-bisphosphate (2FP-400) and Lys 239 from the enzyme fructose bisphosphate aldolase from Leishmania mexicana(2QDG)

7.2.1.1.4: Transition State Stabilization

In the middle of the 20th century, Linus Pauling postulated that the only thing that a catalyst must do is bind the transition state more tightly than the substrate. This can be discerned in Figure 7.2.1.15 and a little math. The diagram shows how the substrate (S) and the transition state (S*) can react with an enzyme (E) to form a complex, which then proceeds to product (following the diagram from "start here" in the right-hand direction), or can go to product in the absence of enzyme (E) (following the diagram from "start here" in the left-hand direction). Note that the diagram is arbitrarily drawn such that the standard Gibbs free energy (G°) of the free product P is higher than that of the free substrate, S.

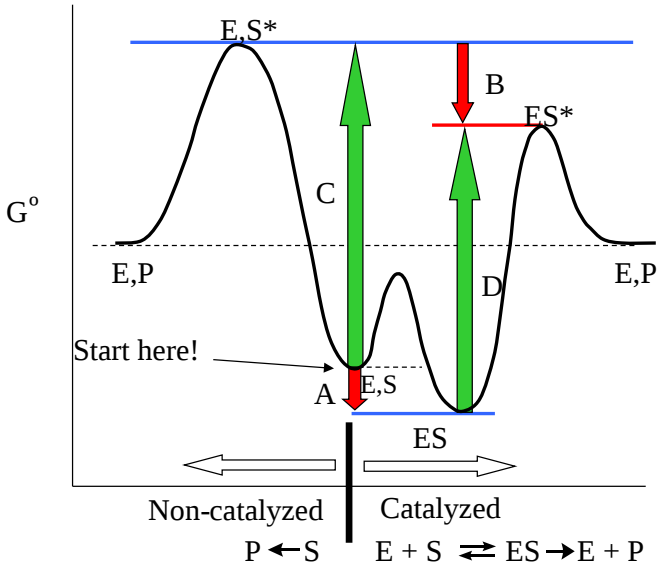


Figure 7.2.1.15: Enzymes bind the transition state more tightly than the substrate

The large colored vertical arrows represent the ΔG° for the transition shown:

- The red arrows A and B represent the ΔG°s for the binding of E + S (arrow A) and E + S* (arrow B), respectively.
- The green arrows C and D represent the ΔG°s for the activation energy of the free substrate (arrow C) and the enzyme-bound substrate (arrow D).

Now consider the two pairs of arrows (A,C and D,B). Add each set like a vector in elementary physics. Since the distance between the two horizontal blue lines is the same for the left-hand process (uncatalyzed) and the right-hand one (catalyzed), it follows that

$$C - A = D - B \tag{7.2.1.1}$$

The negative signs for A and B are used since, in the diagram, both A and B have negative ΔG° values.

Now, for an enzyme to be a catalyst, the activation energy D for the reaction in the presence of the enzyme E must be less positive (i.e., smaller) than the activation energy C in the absence of the enzyme. Therefore, after rearranging the equation:

$$C - D = A - B > 0 \tag{7.2.1.2}$$

and substituting in the ΔG° values for A and B, we can directly compare the free energy of binding the substrate (S) vs. binding the transition state (S*):

$$-RT \ln \mathrm{K_{eqS}} - (-\mathrm{RT} \ln \mathrm{K_{eqS*}}) > 0 \tag{7.2.1.3}$$

Hence, the equilibrium constant for binding the transition state is larger than that for binding free substrate:

$$\mathrm{K_{eq\ S*}} > \mathrm{K_{eq\ S}} \tag{7.2.1.4}$$

Pauling was right. The enzyme just needs to bind the transition state more tightly than the substrate to catalyze the reaction. This is why chemists synthesize stable transition state analogs as potential tight-binding inhibitors of target proteins.

The stability of the transition state also affects the reaction kinetics (which makes sense, given that the activation energy clearly affects the speed of a reaction). As you probably remember from organic chemistry, biomolecular nucleophilic substitution (S_N2) reactions are slow when the central atom where the substitution will occur is surrounded by bulky substituents (sterics once again). We discussed this in the context of nucleophilic substitution on a sp^2 hybridized carbonyl carbon in carboxylic acid derivatives versus on a sp^3 hybridized phosphorus in phosphoesters and diesters. The explanation for this phenomenon has usually been attributed to hindered access to the central atom caused by bulky substituents (intrinsic effects). Is this true? Studies on S_N2 reactions of methylchloroacetonitrile and t-butylchloroacetonitrile (with the reagent labeled with ^{35}Cl) using $^{37}Cl^-$ as the incoming nucleophile in the gas phase showed that the more hindered t-butyl derivative's activation energy was only 1.6 kcal/mol (6.7 kJ/mol) higher than the methyl derivative, but in aqueous solution, the difference is much greater for comparable reactions (Figure 7.2.1.16:).

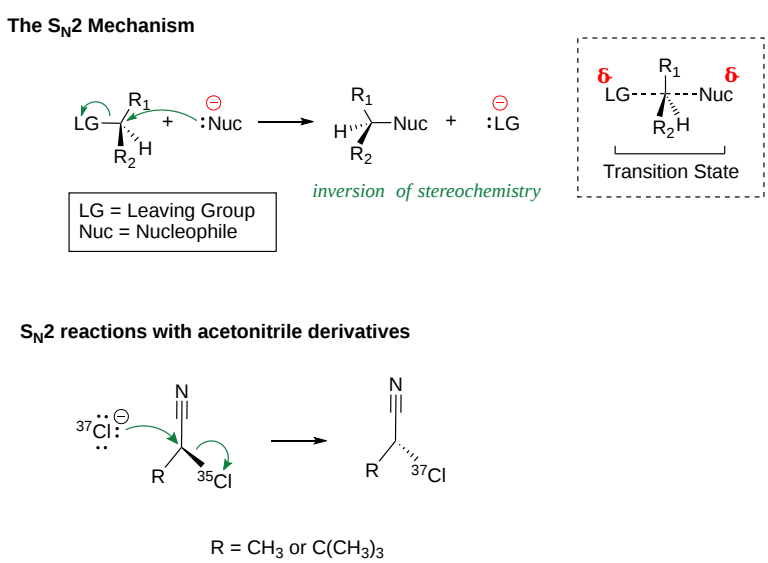


Figure 7.2.1.16: S_N2 reactions are characterized by a pentavalent transition state

They attributed the differences to the solvation effects of the transition state. The bulkier the substituents on the central atom, the more difficult it is to solvate the transition state since water also can't reorient around it as well. In effect, there is steric hindrance for both reactant and solvent.

What does it take for a macromolecule (M) to be a catalyst - an enzyme? It seems the minimum criteria are:

- M binds a reactant
- M binds the transition state more tightly than the substrate

Anything above these is just "icing on the cake." If different functional groups are present in the "active" site of the enzyme that would allow electrostatic, intramolecular, covalent, general acid and/or base catalysis, the better the catalyst.

7.2.1.1.4.1: A transition state analog case study: Abyzmes (Antibody Catalysis)

Recall that antibodies are immune system proteins that bind foreign molecules (see Chapter 5.4). The usual role of an antibody is to initiate an immune response. When the antigen-binding site, located in the variable region of an antibody, binds to an antigen, it triggers the formation of new antibodies (within B-cells) to optimize the immune response to that antigen. These new antibodies are made with mutations in the antigen-binding region. Those that bind better than the original antibody will be selected to form longer-lived memory B-cells, ready for the next time the body encounters the antigen.

Decades ago, Linus Pauling hypothesized that antibodies could be produced with an atypical role—to act as catalysts! If antibody molecules could be made to bind to a compound resembling the transition state of a chemical reaction, they should also presumably catalyze the chemical reaction. In 1987, his prediction was verified. Lerner et al. made a transition-state analog of an ester. When an ester is hydrolyzed, as shown in Figure 7.2.1.17, the sp^2 hybridized carbonyl carbon is converted to an sp^3 hybridized center in the intermediate, with the carbonyl oxygen becoming an oxyanion.

The transition state presumably resembles this unstable intermediate (sp^3, oxyanion). Thus, Lerner synthesized a phosphonate, an ester mimic with a sp^3 hybridized phosphorous replacing the carbonyl C. It also has a negatively charged oxygen, as does the intermediate for the ester. This phosphonate ester is very resistant to hydrolysis. When injected into a mouse (after first being covalently attached to a carrier protein so the small molecule becomes "immunogenic"), the mouse makes a protein antibody that binds to the phosphonate. When the corresponding carboxylic acid ester is added to the antibody, it is cleaved with nominal k_{cat} and K_M values. Site-directed mutagenesis can then be done to make it an even better catalyst! The antibody enzymes have been called **abzymes**. The structure shown in Figure 7.2.1.17 also shows how phosphonamides act as transition state analogs.

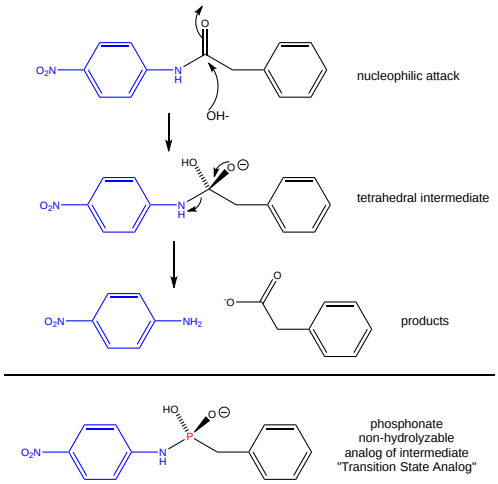


Figure 7.2.1.17: PHOSPHONAMIDES: TRANSITION STATE ANALOGS

Figure 7.2.1.18s shows an interactive iCn3D model of the transition state analog 5-(para-nitrophenyl phosphonate)-pentanoic acid bound to a mouse Fab antibody fragment with esterase activity (1aj7)

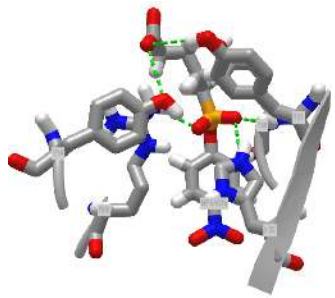


Figure 7.2.1.18: Transition state analog 5-(Para-Nitrophenyl Phosphonate)-Pentanoic Acid (NPE406) bound to Fab antibody fragment with esterase activity (1aj7) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?zSWAhDag8t4LpJaN7

Transition state theory can be used to quantify the relationships described in the graphical analysis above more clearly. This analysis will use the equilibrium constant (in contrast to the last two chapters, which used dissociation constants to characterize macromolecules, receptors, and enzymes binding to ligands). Let's assume that a substrate S is in equilibrium with its transition state $S^{\ddagger}$. Hence Keq = $[S^{\ddagger}]/[S]$. The following reaction can be written: S → $S^{\ddagger}$ → P. Based on our previous kinetic analysis and experience in writing differential equations, dP/dt = k1$[S^{\ddagger}]$. By analogy, enzyme-bound S (ES) can be converted to ($ES^{\ddagger}$) and then on to product as shown in the following chemical equation:

$$\mathrm{E + S \rightleftharpoons ES \longrightarrow ES^{\dagger} \longrightarrow E + P.}$$

For the non-enzyme catalyzed reaction, transition state theory can be used to show that the first order rate constant k_1= kT/h where k is the Boltzmann constant, T is the Kelvin temperature, and h is Planck's constant. Hence, using $K_{eq} = [S^{\dagger}]/[S]$, equation 1 can be derived

$$\frac{dP}{dt} = \frac{kT}{h}\,[S^{\dagger}] = \frac{kT}{h}K^{\dagger}[S] = k_n[S] \tag{7.2.1.5}$$

where k_n (hereafter written as k_N) =(kT/h)$K^{\dagger}$ is the effective first-order rate for the non-catalyzed rate. Now, let's create a more complicated linked equilibrium that shows the same reaction in the presence of an enzyme. Figure 7.2.1.19

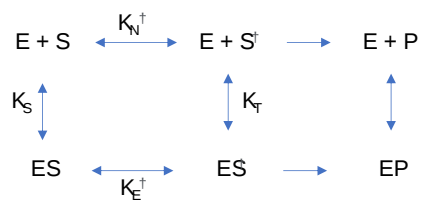


Figure 7.2.1.19: Linked equilibrium for conversion of free $S^{\dagger}$ and $ES^{\dagger}$ to product

Remember that the K values for this analysis are equilibrium constants, not dissociation constants. Note two important equilibrium constants, K_S, the equilibrium constant for the binding of free S to E, and K_T, the equilibrium constant for the binding of free $S^{\dagger}$ to E (assuming that free $S^{\dagger}$ could bind to E before it converted to product). As we have seen for linked equilibrium before, since the K_{eq} values are related to the standard free energy changes, which are state functions, the sum of the standard free energies going from E + S to $ES^{\dagger}$ (by either the top or bottom paths) are path independent, so the products of the K_{eq} for the top path are equal to those for the bottom paths. This gives the following equation:

$$\frac{K_T}{K_S} = \frac{K_{E\dagger}}{K_{N\dagger}} = \frac{k_E}{k_N} \tag{7.2.1.6}$$

The right-hand side is the ratio of the effective first-order rate constant for conversion or $ES^{\dagger}$ → E + P, k_E divided by the rate constant for the conversion of $S^{\dagger}$ → P for the noncatalyzed rate, k_N. The final ratio of rate constants can be derived from the simple relationship that k_x=(kT/h)$K^{\dagger}_x$ where x is either N (non-catalyzed) or E (enzyme-catalyzed). Equation 2 states that the equilibrium constant for the binding of $S^{\dagger}$ to E, K_T, is greater than that for the binding of S to E, K_S (as $k_E > k_N$). K_T/K_R ranges from 10^8 - 10^{14}. Given common values for the equilibrium constant for binding of S to E (10^3 – 10^5 M^{-1}), which is equivalent to the dissociation constant values K_d = 10 uM – 1 mM, the calculated value of $K_T = 10^{15}$ $M^{-1,}$ which gives a dissociation constant for the enzyme and

transition state of $K_d = 10^{-15}$ M (1 femtomolar). This is as tight as one of the highest affinity binding interactions in the biological world, the binding of avidin and biotin. As we noted in Chapter 5.1, assuming that the second order rate constant for avidin/biotin binding and as shown above for $E/S^{\dagger}$ is diffusion-controlled (about 10^8 $M^{-1}s^{-1}$), the off rate for the avidin-biotin or $ES^{\dagger}$ complex is 10^{-7} s^{-1}, equivalent to a half-life of the complex of 80 days. It doesn't get much tighter than that.

Figure 7.2.1.20 represents an image of an enzyme and three molecules, 1–3, that could bind to it. Using the analysis above, which molecule do you think represents the substrate? Transition state? Product?

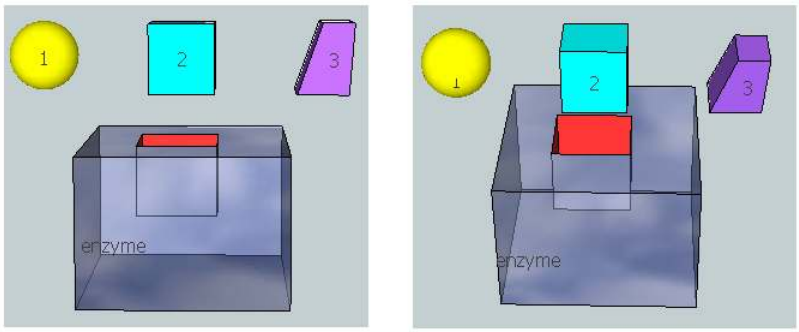


Figure 7.2.1.20: 3D model for binding substrate, transition state and product to an enzyme

7.2.1.2: Physical Strategies for Rate Enhancement

7.2.1.2.1: Intramolecular Catalysis

Consider the hydrolysis of phenylacetate. This reaction, a nucleophilic substitution reaction, could be catalyzed by the addition of the general base acetate to the solution, as described above. Since this reaction rate would double with the doubling of the solution acetate, the reaction is bimolecular (first order in reactant and catalyst). Now consider the same reaction only when the general base part of the catalyst, the carboxyl group, is part of the reactant phenylacetate. Such a case occurs in the acetylated form of salicylic acid—i.e., aspirin. When the carboxy group is ortho compared to the acetylated phenolic OH, it is in the perfect position to accept a proton from water, decreasing the charge development on the O in the transition state. The general base does not have to diffuse to the appropriate site when it is intramolecular with respect to the carbonyl C of the ester link. The rate of this intramolecular base catalysis is about 100-fold greater than that of an intermolecular base catalyst like acetate. It is as if the effective concentration of the intramolecular carboxyl base catalyst is much higher due to its proximity to the reaction site.

Another type of reaction involving a carboxyl group (in addition to simple proton transfer) is when the negatively charged carboxyl O acts as a nucleophile and attacks an electrophilic carbonyl carbon. When the carbonyl is part of an ester, the carboxyl group engages in a nucleophilic substitution reaction, expelling the alcohol part of the ester as a leaving group. The remaining examples below consider the nucleophilic (carboxyl) substitution on phenylesters, with phenolate as the leaving group. The reactions, in effect, transfer an acyl group to the carboxyl group to create an anhydride.

First, consider acyl transfer with aspirin derivatives. As you know, aspirin contains a carboxyl group ortho to an ester substituent. Hence, the carboxyl group can act as a nucleophile and attack the carbonyl carbon of the ester in a nucleophilic substitution reaction. The net effect is to transfer the acetyl group from the phenolic OH to the carboxyl group, converting it to an anhydride. This is an intramolecular reaction. Compare this reaction to a similar bimolecular reaction shown in Figure 7.2.1.21.

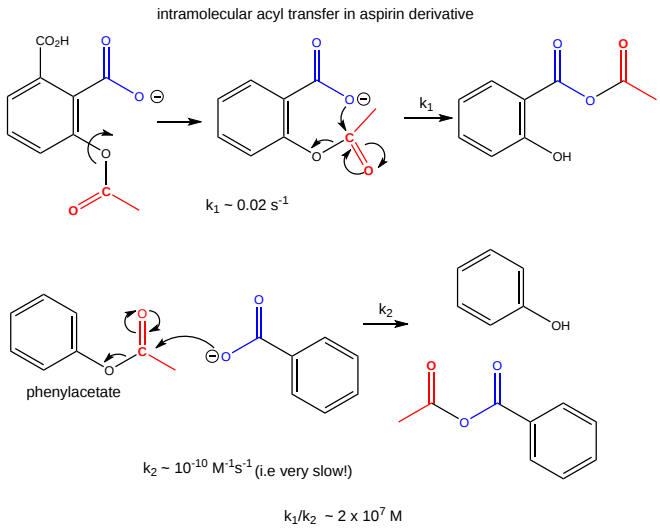


Figure 7.2.1.21: Acyl transfer in aspirin derivatives.

The first-order rate constant of the intramolecular transfer of the acetyl group to the carboxyl group is $k_1 = 0.02\ s^{-1}$. The analogous bimolecular reaction rate constant $k_2 \sim 10^{-10}\ M^{-1}s^{-1}$. Dividing k_1/k_2 gives the relative rate enhancement of the intramolecular over the intermolecular reaction. With units of molarity, this ratio can be interpreted as the relative effective concentration of the intramolecular nucleophile. This makes the effective concentration of the carboxylate in the aspirin derivative 2×10^7 M.

Now consider the cleavage of phenylacetate using acetate as the nucleophile, as shown in Figure 7.2.1.22. The products are acetic anhydride and phenolate. This is a bimolecular reaction (a slow one at that), with a bimolecular rate constant, k_2, which I will arbitrarily set to 1 for comparison to some similar reactions.

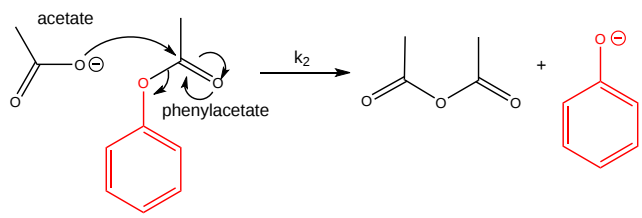


Figure 7.2.1.22: Reaction of acetate with phenylacetate

Now consider a monoester derivative of succinic acid - phenyl succinate - in which the free carboxyl group of the ester attacks the carbonyl carbon of the ester derivative, as shown in Figure 7.2.1.23.

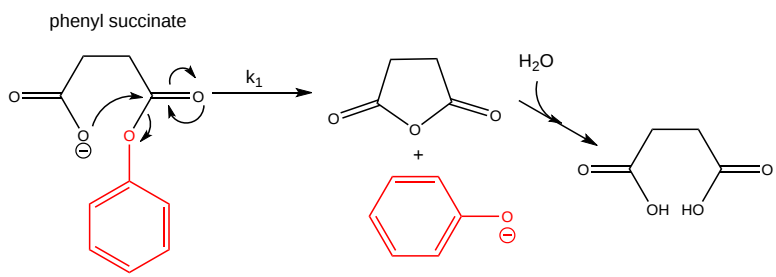


Figure 7.2.1.23: Intramolecular reaction of phenylsuccinate

Suppose you assign a second-order rate constant $k_2 = 1\ M^{-1}s^{-1}$ to the analogous intermolecular reaction of acetate with phenylacetate (as described above). Then, the first-order rate constant for the intramolecular reaction of phenylsuccinate is $10^5\ s^{-1}$. The ratio of rate constants, $k_1/k_2 = 10^5$ M. That is, it would take a 10^5 M concentration of acetate reacting with 1 M phenylacetate in the first bimolecular reaction to get a reaction as fast as the intramolecular reaction of phenylsuccinate. The intramolecular reaction of an even more sterically restricted bicyclic phenylcarboxylate shown in Figure 7.2.1.24 has a $k_1/k_2 = 10^8$ M.

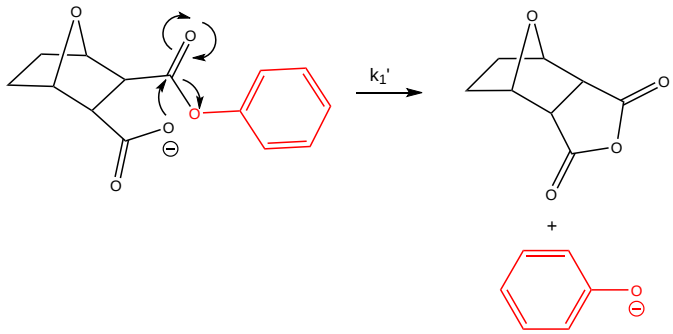


Figure 7.2.1.24: Intramolecular reaction of bicyclic phenylcarboxylate

Another example is anhydride formation between two carboxyl groups. The ΔG^o for such a reaction is positive, suggesting an unfavorable reaction. Consider two acetic acid molecules condensing to form acetic anhydride. For this intermolecular reaction, $K_{eq} = 3x10^{-12}\ M^{-1}$. Now consider the analogous intramolecular reaction of the dicarboxylic acid succinic acid. It condenses in an intramolecular reaction to form succinic anhydride with a $K_{eq} = 8x10^{-7}$ (no units). The ratio $K_{eq\text{-}intra}/K_{eq\ inter} = 3\ x10^5$ M. It is as if the effective concentration of the reacting groups, because they do not have to diffuse together to react, is $3\ x10^5$ M.

How does this apply to enzyme-catalyzed reactions? Enzymes bind substrates in physical steps, which are typically fast. The slow step is often the chemical conversion of the bound substrate, which is effectively intramolecular if the initial binding reaction is fast. These three kinds of reactions, intermolecular, intramolecular, and enzyme-catalyzed, can be broken down into two hypothetical steps: a binding followed by catalysis, as shown in Figure 7.2.1.25.

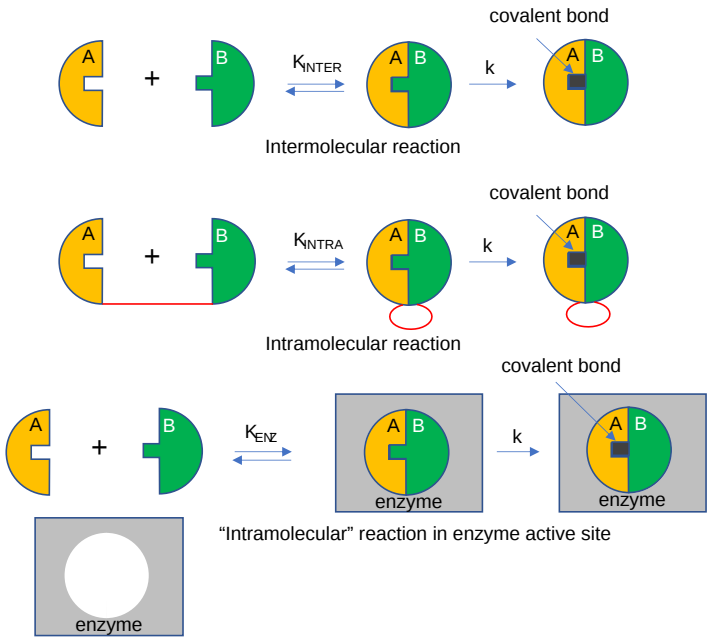


Figure 7.2.1.25: Intermolecular, intramolecular, and enzyme-bound reactions

Suppose the rate constants for the chemical steps are all identical. Then, the advantages of the intramolecular and enzyme-catalyzed reaction over the intermolecular reaction are K_{INTRA}/K_{INTER} and K_{ENZ}/K_{INTER}, respectively.

The advantage of intramolecular reactions can be seen by studying the Ca-EDTA complex. Calcium in solution exists as an octahedrally coordinated complex with water occupying all the coordination sites. EDTA, a multidentate ligand, first interacts through one of its potential six electron donors to Ca in a reaction entropically disfavored from the Ca-EDTA perspective, although one water molecule is released. Once this first intramolecular complex is formed, the rest of the ligands on the EDTA rapidly

coordinate with the Ca and release bound water, as illustrated in Figure **7.2.1.26**. The former is no longer entropically disfavored since it is now an intramolecular process, while the latter is favored by releasing the remaining five water molecules.

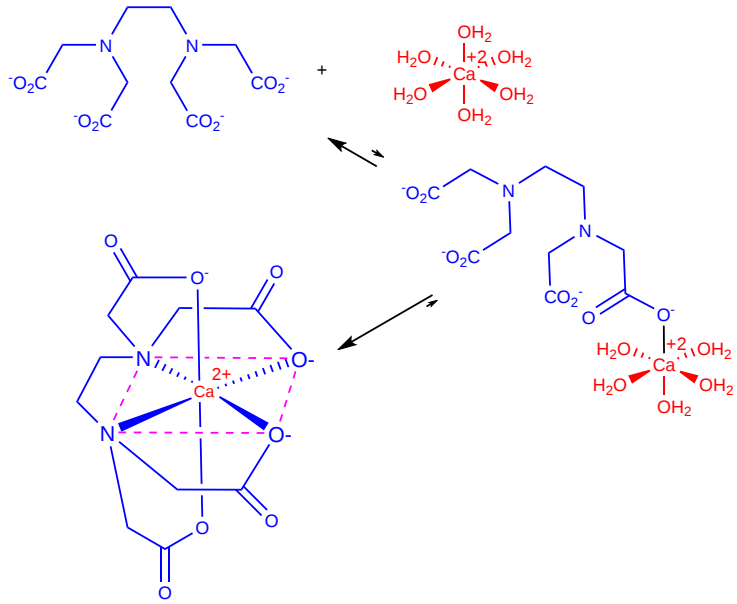


Figure **7.2.1.26**: Binding of Ca^{2+} and EDTA

We've shown above the catalytic advantage offered by intramolecular reaction in terms of a dramatic increase in the effective concentration of reactants, which sometimes reached levels of 10^8 M. Another way is to look at entropy changes associated with dimer formation. The table below shows that an intramolecular reaction is favored over an intermolecular reaction since, in the latter, significant decreases in translational and rotational entropy result.

Translational, Rotational, and Internal Entropies for Dimer Formation: A + B ↔ A-B (cal/K·mol)

System	A	B	A-B	ΔS
		Gas		
S trans	30	30	30	-30
S rot	20	20	20	-20
S int	5	5	20	+10
Gas → Solution	-10	-10	-15	
S sol	45	46	55	-35 (Corresponds to 10^8-10^9 M)

7.2.1.2.2: Strain Distortion

In organic chemistry, you learned that certain structures, such as three-membered and four-membered ring structures, such as epoxides, were highly reactive due to the strain distortion inherent to the unfavored bond angles inherent to the ring. Enzyme active sites can also utilize strain distortion within a bound substrate to increase the molecule's reactivity and favor the formation of the transition state. Many enzymes that function by the induced fit model also utilize strain distortion within their catalytic mechanism. Within the unbound state, they remain in a low catalytic state; however, the interaction with the substrate induces the destabilization of the enzyme active site or may induce strain within the substrate, causing the initiation of the catalytic activity of the enzyme.

Quantitative Deconstruction of Enzyme Catalysis

New ideas have emerged that quantitatively partition the overall catalytic enhancement by enzymes into discrete features and account for the full rate enhancement conferred by enzymes. We will discuss this in Chapter 6.5, using a class of enzymes called serine proteases as an example.

7.2.1.3: A Note on Asymmetric Catalysis/Organocatalysis

In a subsequent section, we will discuss how protein enzymes use the catalytic strategies described above. An intriguing question arises: how much of the structure of a large protein is really needed for catalysis? Much work has been directed to developing small-molecule catalysis mimetics of large protein enzymes. How small can you reduce the size of a protein and still get catalysis?

One important feature of enzyme catalysis is that it catalyzes reactions in which only one enantiomer is produced. That is, the synthesis is asymmetric. This is typically a consequence of the asymmetric enzyme (itself chiral) binding only one enantiomer as a reactant and/or the imposition of steric restrictions on the possible reactions of the bound substrate. L-Pro alone can act as an asymmetric catalyst in an aldol condensation reaction. Figure **7.2.1.27**:

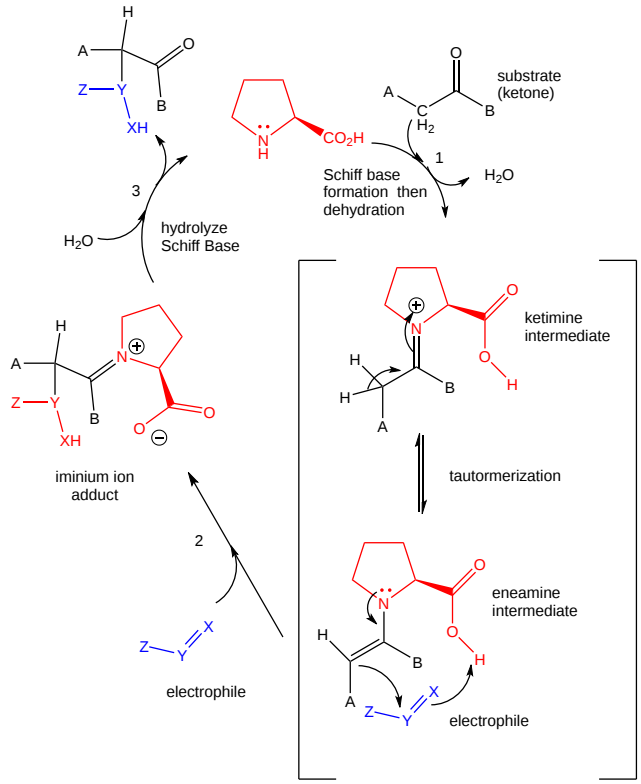


Figure **7.2.1.27**: L-PRO CATALYSIS OF AN ALDOL CONDENSATION: POSSIBLE MECHANISM

Catalysts are vital in biological settings and the laboratory synthesis of molecules that sustain our culture and economy. Transition metal and, increasingly, protein enzymes have been used as industrial catalysts. New **asymmetric catalysts have now joined them** (a subset of **organocatalysts**). The work of Benjamin List and David MacMillan, who were instrumental in developing the ideas of asymmetric catalysts, has been recognized by the Nobel Commission, which awarded them the 2021 Nobel Prize in Chemistry.

The enzyme triose phosphate isomerase catalyzes an asymmetric reaction in which only one enantiomer of glyceraldehyde-3-phosphate is produced from the achiral dihydroxyacetone phosphate. Figure **7.2.1.28** shows this enantiospecific reaction.

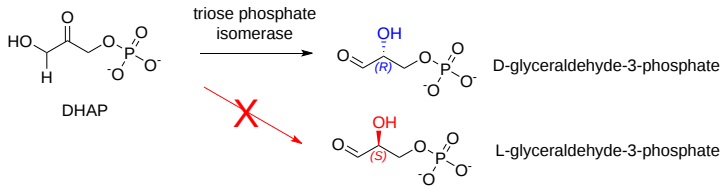


Figure 7.2.1.28: Reaction catalyzed by triose phosphate isomerase

Figure 7.2.1.29s contains an interactive iCn3D model of a triose phosphate isomerase from Trypanosoma brucei brucei (1KV5), which shows a conserved active site Pro 168 (spacefill) and amino acid side chains within 4 Å (stick) within the context of one monomer (cartoon) of the dimeric protein.

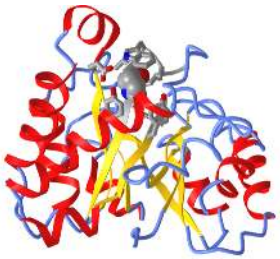


Figure 7.2.1.29: Active site proline in triose phosphate isomerase from Trypanosoma brucei brucei (1KV5). (Copyright; author via source). Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/icn3d/share.html?1Hnu1xmU5j76Jwrx9

7.2.1.4: Summary

This chapter provides an in-depth analysis of why uncatalyzed reactions in solution are slow and how catalysts, particularly enzymes, overcome these barriers by lowering the activation energy. The discussion begins with the example of ester hydrolysis, illustrating that the high activation energy of the uncatalyzed reaction stems from the formation of a charge-separated, high-energy transition state that closely resembles an unstable intermediate.

Key Concepts Covered:

1. **Energy Profiles and Transition States:**
   - The chapter explains how the energy diagram of an endergonic reaction features a transition state that is higher in energy than a charge-separated intermediate. This high-energy state, which arises from significant charge development during bond making and breaking, limits the reaction rate in the absence of a catalyst.
2. **Strategies for Transition State Stabilization:**
   - **Acid-Base Catalysis:**
     - *Specific Acid/Base Catalysis:* The reaction rate depends solely on the concentration of hydronium or hydroxide ions.
     - *General Acid/Base Catalysis:* Buffer components (or functional groups within an enzyme) donate or abstract protons to reduce charge separation in the transition state, thereby lowering the activation energy.
   - **Metal Ion (Electrostatic) Catalysis:**
     - Metal ions (e.g., Zn^{2+}, Cu^{2+}) coordinate with reaction intermediates to stabilize developing charges, enhance electrophilicity, and sometimes lower the pKa of bound water molecules, which is critical in enzymes like carbonic anhydrase.
   - **Covalent (Nucleophilic) Catalysis:**
     - Nucleophilic catalysts such as pyridine form covalent intermediates with the substrate. This alternative reaction pathway features a lower activation energy due to the formation of a new, more stable intermediate.
   - **Intramolecular Catalysis:**
     - When a catalytic group is part of the same molecule as the reactive site, as in aspirin derivatives or phenylsuccinate, the effective concentration is dramatically increased. This proximity effect accelerates the reaction by minimizing the

entropic cost of bringing reactants together.

3. **Quantitative Aspects of Enzyme Catalysis:**
   - Transition state theory is used to show that catalysts, including enzymes, enhance reaction rates by binding the transition state more tightly than the substrate. This principle, first proposed by Linus Pauling, is quantified by comparing equilibrium constants for substrate and transition state binding.
   - The chapter presents mathematical derivations linking free energy differences, rate constants, and equilibrium constants (K_eq) for both catalyzed and uncatalyzed reactions, demonstrating that enzymes can achieve binding affinities for transition states comparable to some of the strongest noncovalent interactions known in biology.
4. **Applications and Implications:**
   - **Transition State Analogs and Abzymes:**
     - By synthesizing stable molecules that mimic the transition state, researchers can design potent inhibitors and even generate catalytic antibodies (abzymes) that accelerate chemical reactions.
   - **Asymmetric Catalysis:**
     - The inherent chirality of enzymes ensures that only one enantiomer of the product is formed, a principle that is also applied in the development of small-molecule organocatalysts.
   - **Effective Concentration and Entropy Considerations:**
     - The chapter discusses how intramolecular reactions overcome entropic barriers that limit bimolecular reactions, with quantitative comparisons illustrating rate enhancements that can approach effective concentrations of up to 10^8 M.

In summary, the chapter unifies various catalytic strategies—acid-base, metal ion, covalent, and intramolecular—under the common theme of transition state stabilization. It highlights how enzymes, through precise structural features and dynamic interactions, achieve extraordinary rate enhancements. This understanding not only elucidates fundamental principles of chemical reactivity but also informs the design of synthetic catalysts and therapeutic inhibitors in modern biochemistry.

7.2.1.5: References

Amyes TL, Richard JP. Biochemistry. 2013, 52(12), 2021-35. doi: 10.1021/bi301491r

Ferst and Kirby, J. Am. Chem. Soc. 1967, 89, 19, 4857–4863. https://doi-org.ezproxy.csbsju.edu/10.1021/ja00995a007

Komiyama et al., PNAS, **1977,** 74 (7) 2634-2638. https://doi.org/10.1073/pnas.74.7.263

Lerner, R. A., & Tramontano, A. (1987). Antibodies as enzymes. *Trends in Biochemical Sciences*, *12*, 427-430. https://doi.org/10.1016/0968-0004(87)90208-8

OpenStax College. (2022). *Chemistry*. OpenStax. https://openstax.org/details/books/chemistry-2e

Regan, C. K., Craig, S. L., & Brauman, J. I. *Science.* 2002 *295*(5563), 2245-2247. DOI: 10.1126/science.1068849

This page titled 7.2.1: How Enzymes Work is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.1: How Enzymes Work** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.2.2: Kinetics without Enzymes

Learning Goals (ChatGPT o1, 1/30/25)

- **Fundamentals of Reaction Kinetics and Calculus in Chemistry**
  - Understand the use of calculus in describing reaction kinetics, including the meaning of the derivative dX/dt as the instantaneous rate of change in concentration.
  - Translate between the differential (dX/dt) and integrated (X as a function of time) forms of rate equations.
- **Analyzing First-Order Reactions**
  - Derive and solve the first-order differential equation for an irreversible reaction (e.g., A → P) and explain the significance of the solution, $A = A_0e^{-k_1t}$.
  - Interpret kinetic graphs, including exponential decay curves for A versus time and the linear relationship in ln[A] versus time plots, and relate these to reaction velocity.
- **Understanding Second-Order and Pseudo First-Order Kinetics**
  - Derive and solve the integrated rate equation for second-order reactions (e.g., A + A → P or A + B → P) and explain how the reaction order influences the time course of reactant consumption.
  - Recognize conditions under which a bimolecular reaction can be treated as pseudo first-order and calculate the effective rate constant.
- **Initial Rate vs. Integrated Rate Approaches**
  - Compare and contrast initial rate methods (measuring the slope at t ≈ 0) with progress curve analysis (integrated rate equations) and discuss the advantages and limitations of each in kinetic studies.
- **Multi-Step and Reversible Reaction Kinetics**
  - Write differential equations for simple reversible reactions (A ↔ P) and multi-step processes (e.g., consecutive irreversible reactions A → B → C and reversible networks A ↔ B ↔ C).
  - Interpret progress curves and steady-state behavior in reversible systems and understand how changes in rate constants affect the approach to equilibrium.
- **Modeling and Data Analysis in Kinetics**
  - Utilize computational tools (e.g., VCell) to solve ordinary differential equations (ODEs) for complex reaction systems and generate progress curves that illustrate concentration changes over time.
  - Apply statistical methods and curve fitting to experimental kinetic data, linking theoretical models with observed reaction behavior.
- **Developing Chemical and Biological Intuition**
  - Integrate mathematical concepts with chemical intuition by relating the rate equations and progress curves to molecular behavior and reaction mechanisms.
  - Understand how the kinetic analysis of simple chemical reactions provides the foundation for exploring more complex enzyme-catalyzed reactions in biological systems.

By achieving these learning goals, students will be equipped to analyze, model, and interpret the kinetics of single-step and multi-step reactions, forming a critical basis for understanding enzyme-catalyzed processes and broader biochemical reaction mechanisms.

7.2.2.1: Single Step Reactions

First, we will explore the kinetics of non-catalyzed reactions, which is needed to understand the kinetics of the more complicated enzyme-catalyzed reactions.

Calculus: Derivatives and Integrals

In this book, we will refer to the change of concentration of a chemical species X as a function of time as dX/dt instead of ΔX/Δt, where dX/dt is the derivative of X with respect to time t. This is the language of calculus, which most readers would have studied. The use of calculus will be mainly limited to writing equations of the form dX/dt = f(t), a type of differential equation. We will also use a few integrals, but a working knowledge of calculus is not required. For readers who have not studied calculus, replace dX/dt with ΔX/Δt in your mind, and you will derive the same meaning.

You studied two types of kinetic equations in introductory chemistry to analyze kinetic data:

- Initial Rates: In this method, the initial rate, v_0, is measured as a function of the concentration of reactants. The initial velocity, v_0, is the initial slope of a graph of the concentration of reactants or products as a function of time, taken over a range of times such that only a small fraction of A has reacted. Under this condition, [A] over this short time range is approximately constant and equal to Ao. Initial rate graphs are often based on the measurement of product increase with time, ΔP/Δt, so v_0 vs. A plots have positive slopes. The velocity at time t along the A vs. t curve, dA/dt, constantly changes as [A] decreases since the velocity depends on [A]. To reiterate, the initial velocity of the reaction is the slope of the initial linear part of the decay curve when the rate is essentially linear over a narrow range of [A].
- Integrated Rates and Progress Curves: In this method, a differential equation that gives the change of A or P with time (dA/dt or dP/dt) is integrated to give an equation that shows the concentration of A or P as a function of time. For any given reaction, it is essential to be able to write the integrated rate equation, often called an ordinary differential equation (ODE). In almost every case, we provide the solutions to the ODEs in either a mathematical equation or a fitted graph of A vs. t. As the reactions get more complicated, we let computers solve them numerically and show the output. The ODEs are also called **progress curves** as they show how the concentration of reactants and products changes with time.

In either case, the mathematical equations describing the reaction are used to fit experimental data taken in the lab. Statistics are used to see how closely the experimental data fit the rate equations. The best-fit equation gives the most likely chemical reaction equations for the interconversion of reactants and products.

Most biochemistry textbooks focus on initial velocities when explaining enzyme kinetics. However, in many ways, this method is less intuitive than exploring how the concentration of a molecule changes with time. An analogy is the concept of density (mass/volume), which is more complicated than understanding mass or volume separately. We use both methods to develop both chemical and biological intuitions of kinetic properties.

We will first explore simple irreversible reactions (1st and 2nd order), then make them reversible, and then couple them together to form more complex reaction schemes, much as we are compelled to do for enzyme-catalyzed reactions.

7.2.2.1.1: First Order Reaction

$$\mathrm{A} \xrightarrow{k_1} \mathrm{P}$$

where k_1 is the first-order rate constant. For these reactions, the velocity of the reaction, v, is directly proportional to [A], or

$$v = -\frac{dA}{dt} = +\frac{dP}{dt} = k_1 A \tag{7.2.2.1}$$

The negative sign in -d[A]/dt indicates that the concentration of A decreases. The equation could also be written as:$v = \frac{dA}{dt} = -k_1 A$

For the rest of the reactions in this book, we follow the convention of writing all velocities expressed as d[x]/dt as positive numbers. A negative sign for a term on the right-hand side of the differential equation will indicate that the concentration dependency of that term will lead to a decrease in [x] with time. Likewise, a positive sign for the term on the right-hand side of the equation will indicate that the concentration dependency of that term will lead to an increase in [x] with time.

Examples: dA/dt = **-k**A shows that A will **decrease** with time. dA/dt = **+k**A shows that A will **increase** with time.

Here is the solution to the differential Equation 6.3.A.1 for [A] as a function of t.

$$A = A_0 e^{-k_1 t} \tag{7.2.2.2}$$

? A derivation of the first order rate equation

Here it is!

Derivation

$$\begin{gathered}\int_{A_O}^{A} \frac{dA}{A} = -k_1 \int_0^t dt \\ \ln A - \ln A_0 = -k_1 t \\ \ln A = \ln A_0 - k_1 t \\ A = A_0 e^{-k_1 t}\end{gathered} \tag{7.2.2.3}$$

Equation 6.3.A.2 is an example of an integrated rate equation. The following graphs show plots of A vs t and lnA vs. t for a first-order process. Note that the derivative of the graph of A vs. t (dA/dt) is the velocity of the reaction. The graph of ln[A] vs. t is linear with a slope of $-k_1$. The velocity of the reaction (slope of the A vs t curve) decreases with decreasing A, which is consistent with equation 1. Again, the initial velocity is determined from data taken in the first part of the decay curve when the rate is linear and little A has reacted. That is, [A] is approximately equal to $[A_0]$.

Figure **7.2.2.1** shows two ways to plot 1st order reaction data. The left graph shows the exponential decay of A with time and the corresponding rise in P when A_0 is 0 and k_1=2. The other shows the linear fall of ln[A] vs time.

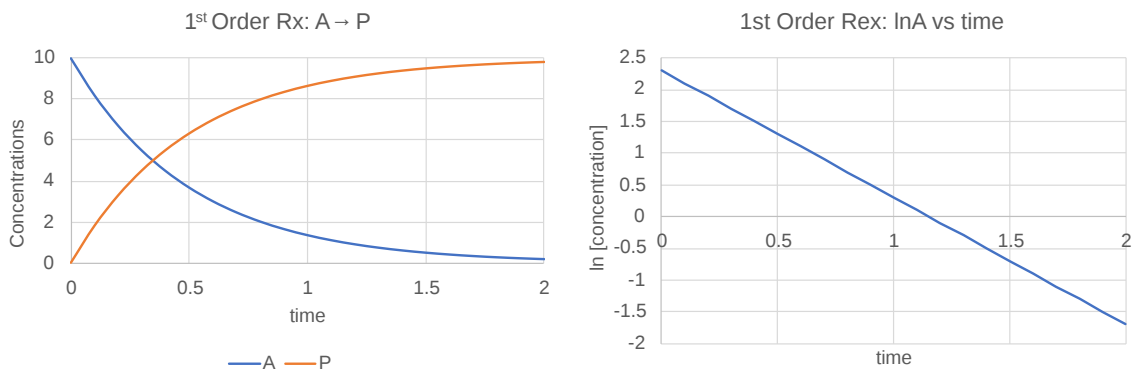


Figure **7.2.2.1**: Two ways to plot 1st order reaction data

Once again, for complete clarification, another way of analyzing the kinetics of a reaction, in addition to following the concentration of a reactant or product as a function of time and fitting the data to an integrated rate equation, is to plot the initial velocity, v_o, of the reaction as a function of the concentration of reactants. The initial velocity is the initial slope of a graph of the concentration of reactants or products as a function of time, taken over a range of times such that only a small fraction of A has reacted, so [A] is approximately constant = A_o. From the first-order graph of A vs. t above, the slope approaches 0 with increasing time as [A] approaches 0, which indicates that the reaction velocity depends on A. For this first-order process, two equivalent equations can be written showing the

- disappearance of A as v = -d[A]/dt = k_1[A], while
- appearance of A as v = d[A]/dt = $-k_1$[A],

Both equations show that v is directly proportional to A. As [A] is doubled, the initial velocity is doubled.

Velocity graphs used by biochemists often show the initial velocity of product formation (not reactant decrease) as a function of reactant concentration. Hence, as product concentration increases, the initial velocity is positive. A graph of v (= dP/dt) vs [A] for a first-order process would have a positive slope and be interpreted as showing that the rate of appearance of P depends linearly on [A].

7.2.2.1.2: Second-Order/Pseudo First-Order Reactions

$$A + B \xrightarrow{k_2} P \text{ or } A + A \xrightarrow{k_2} P \tag{7.2.2.4}$$

where k_2 is the second-order rate constant. For the first of these irreversible reactions, the velocity of the reaction, v, is directly proportional to [A] and [B], or

$$v = \frac{dA}{dt} = -k_2[A][B] \tag{7.2.2.5}$$

We will consider two special cases of this reaction type:

1. [B] >> [A]. Under these conditions, [B] never changes, so Equation 5 becomes

$$v = -\left(k_2[B]\right)[A] = -k_1'[A] \tag{7.2.2.6}$$

where $\mathbf{k}_1$' is the pseudo first order rate constant (= k2[B]) for the reaction. The reaction appears to be first order, depending only on [A].

2. As illustrated in the second reaction above, the only reactant is A, which must collide with another A to form P.

The following differential equation can be written and solved to find [A] as a function of t.

$$v = \frac{dA}{dt} = 2\frac{dP}{dt} = -k_2A^2 \tag{7.2.2.7}$$

Solving the differential equation for A gives the following:

$$\frac{1}{A} = \frac{1}{A_0} + k_2t \tag{7.2.2.8}$$

? A derivation of the second order rate equation

Here it is!

Derivation

$$\begin{gathered}\frac{dA}{dt} = -k_2A^2 \\ \int_{A_0}^{A}\frac{dA}{A^2} = \int_{A_0}^{A}A^{-2}dA = \int_{A_0}^{A}A^{-2}dA = -k_2\int_0^t dt \\ \left.\frac{A^{n+1}}{n+1}\right]_{A_0}^{A} = \left.\frac{A^{-1}}{-1}\right]_{A_0}^{A} = \left.-k_2t\right]_{A_0}^{A} \\ -\frac{1}{A} - \left(\frac{1}{A_0}\right) = -k_2t\end{gathered} \tag{7.2.2.9}$$

Figure **7.2.2.2** shows plots of A vs. t and 1/A vs. t for a second-order process when A_0 is 0 and k_2=1. The right graph shows the linear rise of 1/[A] with time.

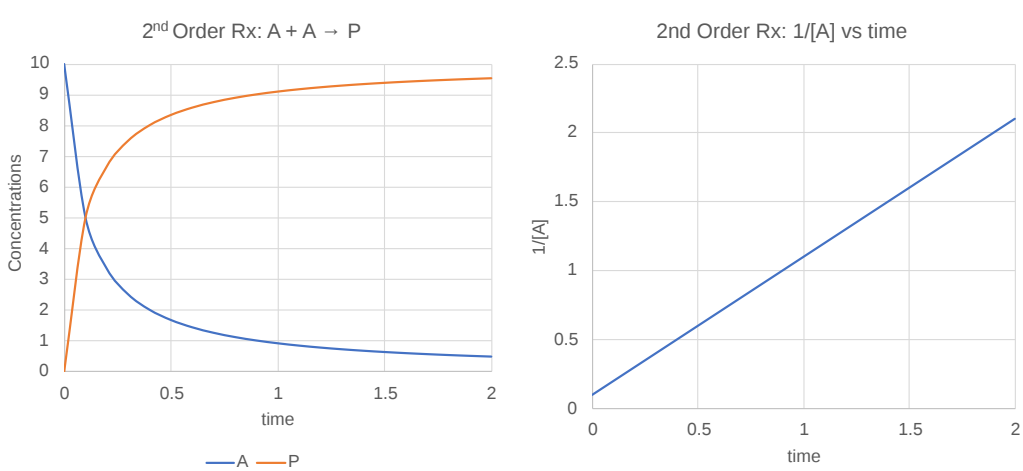
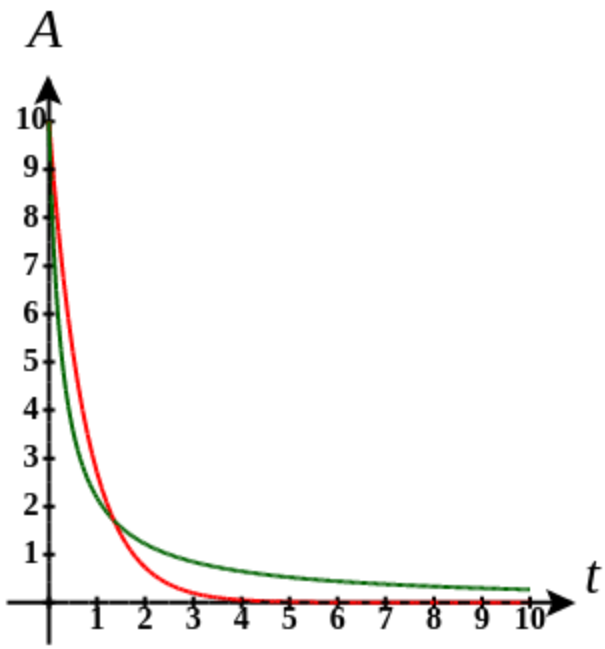


Figure **7.2.2.2**: Plots of A vs t and 1/A vs t for a second order process when A_0 is 0 and k_2=1

Note that just from a plot of A vs. t, it would be difficult to distinguish a first from a second-order reaction. If the plots were superimposed, you would observe that at the same concentration of A (10, for example), the v_o of a first-order reaction would be proportional to 10, but for a second-order reaction, to 10^2 or 100. Therefore, the second-order reaction is faster (assuming similarity in the relative magnitude of the rate constants), as indicated by the steeper negative slope of the curve. However, at low A (0.1 example), the v_o of a first-order reaction would be proportional to 0.1, but v_o of a second-order reaction is proportional to 0.1^2 or 0.01. Therefore, at low A, the second-order reaction is slower.

The interactive graphs below show the first and second-order conversion of reactant A to product. Change the sliders to see how the curves are different.



By comparing these curves, you should see how difficult it is to differentiate between a 1st and 2nd order process unless the reaction progresses to almost completion.

7.2.2.2: Multi-Step Reactions

7.2.2.2.1: Reversible First Order Reactions

$$\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{P} \tag{7.2.2.10}$$

Here is the differential equation for dA/dt..

$$\mathrm{v} = \frac{\mathrm{dA}}{\mathrm{dt}} = -\mathrm{k}_1\ \mathrm{A} + \mathrm{k}_2\mathrm{P} \tag{7.2.2.11}$$

Here are the solutions for both A and P as a function of time

$$A = \frac{A_0\left(k_2 + k_1\left[e^{-(k_1+k_2)t}\right]\right.}{k_1 + k_2}$$
$$P = A_0 - \left(\frac{A_0\left(k_2 + k_1\left[e^{-(k_1+k_2)t}\right]\right.}{k_1 + k_2}\right) \tag{7.2.2.12}$$

Figure 7.2.2.3 shows graphs of A and P vs t for this reaction at two different sets of values of k_1 and k_2.

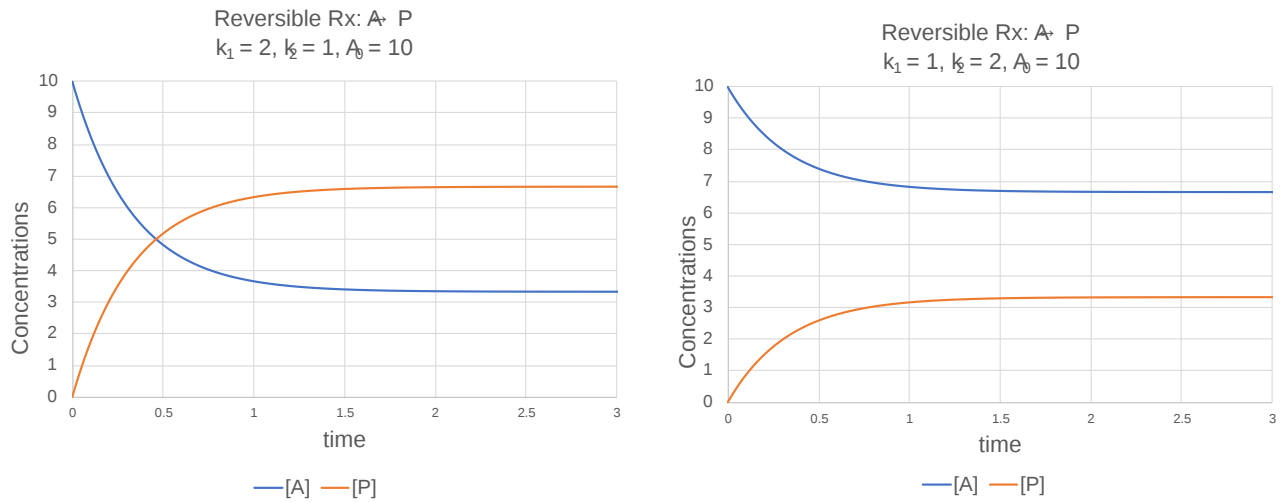
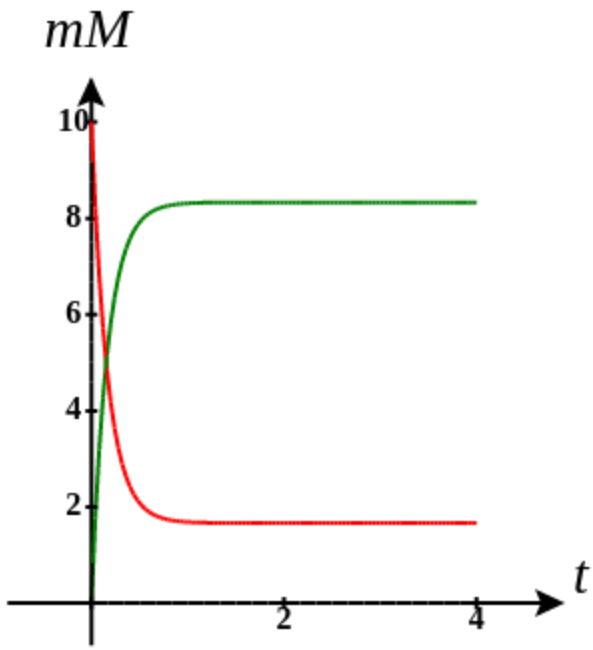


Figure 7.2.2.3: Graphs of A and P vs t for the reversible reaction A ↔ P at two different sets of values of k_1 and k_2

Change the sliders on the interactive graph below of a reversible reaction to see how changing the relative values of the forward and reverse rate constants affects the concentrations at which the concentration plateaus are reached.

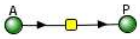


We all grew up on mathematical graphs that give you valuable insight into textual descriptions and data tables from which the graphs were made. These graphs are enhanced when you can use sliders to change constants, as for the reversible reaction A ↔ P above. Even then, you might not infer that when the reaction has reached equilibrium, product is still being made from reactant, and reactant from product, since the equilibrium is dynamic. To add insight into simple and complex reactions, animations showing the continual disappearance of reactants and products are valuable.

This book will incorporate many animations to show the changes in the reactant and product concentrations visually. Hui Liu and Shraddha Nayak (Animation Lab, University of Utah) made all the animations in this book using mathematical solutions to the progress curves for the reactions. Multiple modes of presentation are useful as learners with different backgrounds and preferred ways of learning attempt to understand complex materials.

It is relatively simple to write the differential equations (differential) to show how the rate of disappearance of a reactant A (for example), dA/dt, depends on the concentration of its immediate participants in the reaction. It is not so easy to solve the equations (as we did above) for the progress curve, which shows how [A] changes with time t (i.e., [A] = f(t)). Luckily, many programs have been developed that produce numerical solutions to the differential equations and give progress curve graphs like [A] = f(t). Two interrelated, freely available programs, Copasi and Virtual Cell (Vcell), can solve all the equations for hundreds of cellular reactions simultaneously. They use a format called **Systems Biology Markup Language** (**SBML**) for describing and storing computational models. We will use Vcell models in this book as they are straightforward to create. All the coding to describe the reactions is built into Vcell and this book, and hidden from you. All you will see are the output results. You can change the progress curves by moving sliders to change constants and see the resulting changes in graphical outputs.

The VCell models use a reaction diagram that shows all of the interconnected species. The first Vcell model we will run is for the reversible conversion of A to P (A ↔P), which we just discussed and displayed in the graph above. Here is the Vcell reaction diagram and a description of its components.



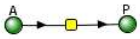
- The reactant A and product P are called species and are shown as green spheres.
- The yellow square indicates a reaction node connecting A to P.
- Lines connect the species that participate in the reaction. The arrows appear unidirectional, **BUT** the equations describing the concentrations of A and P are derived assuming a reversible reaction with rate constants k_f (forward reaction) and k_r (reverse reaction).

The program calculates A and P as a function of time (i.e., it solves the differential equations for both species). The graphs of concentration vs time are called progress curves. It can also calculate **fluxes (J)** (velocities) for each species. The flux at any given time is the slope of the concentration vs time curve at any given time. When we get to metabolism, we will talk about fluxes of metabolites through pathways. Also, fluxes are used to describe the rate of movement of solute through membranes. Here is the result of the simulation run in Vcell, exported as an SBML file, and displayed in the book using a program called MiniSideWinder.

VCell MODEL

Reversible reaction A ↔ P

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



$$A \underset{k_r}{\overset{k_f}{\rightleftharpoons}} P$$

Initial parameter values: $k_f = 2$, $k_r = 4$

Select Load [model name] below

Load AtoP_Rev_Uncatalyzed

Select **Start** to begin the simulation.

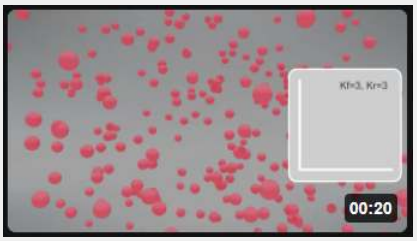
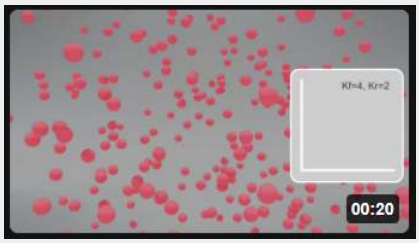
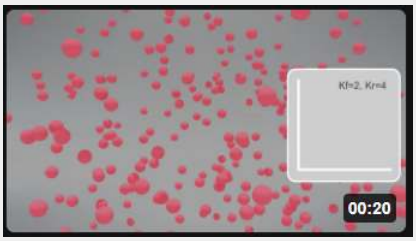
Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Animations

The video animations show particles representing A (red) and P (cyan) interconverting in a reversible process with embedded progress curves showing A (red) and P (cyan) vs. time.

Reversible Rx A (red) ↔ B (cyan): $k_F = k_R = 3$	Reversible Rx A (red) ↔ B (cyan): $k_F=4$, $k_R=2$	Reversible Rx A (red) ↔ B (cyan): $k_F=2$, $k_R=4$
		

7.2.2.2.2: Consecutive Irreversible First-Order Reactions

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

The following differential equations can be written for these reactions:

$$\begin{aligned} \frac{dA}{dt} &= -k_1\ A \\ \frac{dB}{dt} &= k_1\ A - k_2\ B \\ \frac{dC}{dt} &= k_2\ B \end{aligned} \tag{7.2.2.13}$$

Here are the solutions to the differential equations:

$$\begin{aligned} A &= A_0 e^{-k_1 t} \\ B &= \frac{k_1 A_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right) \\ C &= A_0 - A - B = A_0\left[1 + \frac{1}{k_1 - k_2}\left(k_2 e^{-k_1 t} - k_1 e^{-k_2 t}\right]\right. \end{aligned} \tag{7.2.2.14}$$

Figure **7.2.2.4** shows graphs of A, B, and C vs. t for these reactions for a fixed value of k_1 and k_2.

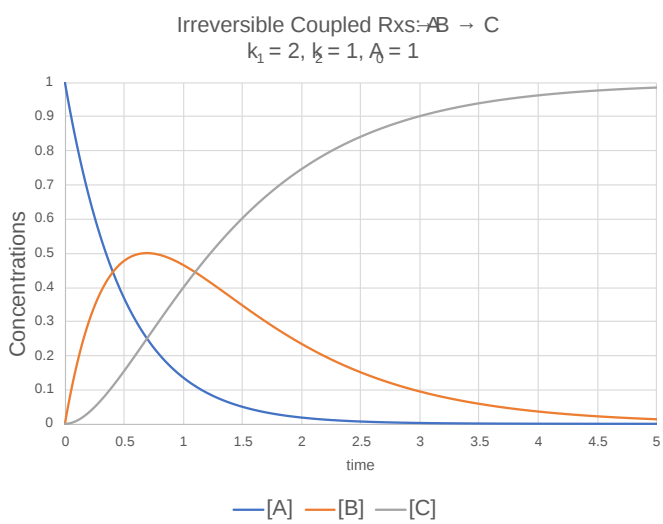
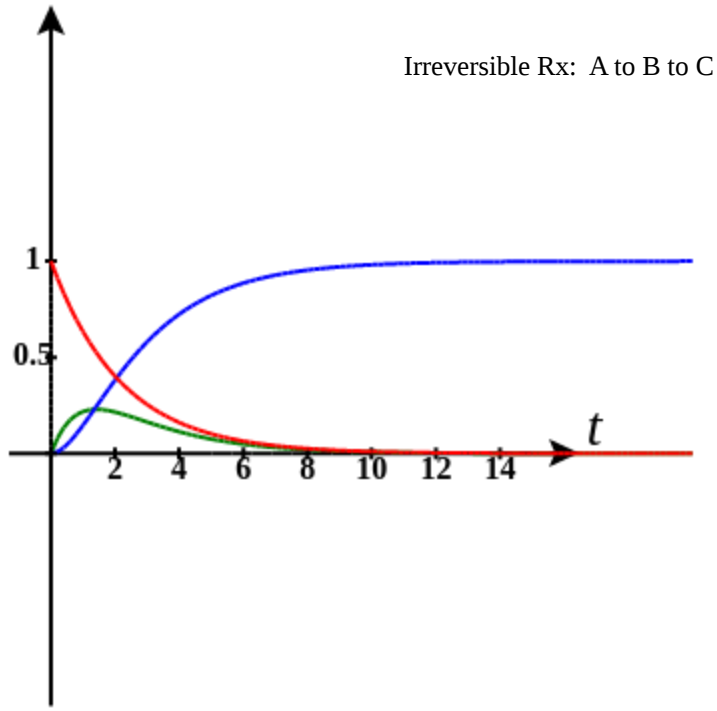
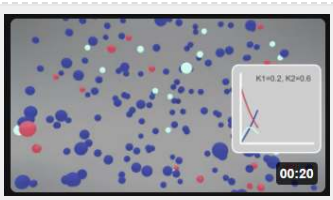
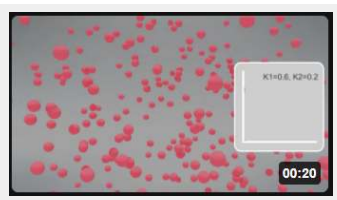


Figure **7.2.2.4**: Graphs of A, B, and C vs. t for the irreversible reactions A → B → C for a fixed value of k_1 and k_2.

Change the sliders on the interactive graph below to see how changing the forward and reverse rate constants affects the curves.



Here are two different animations for the **irreversible** reaction using different rate constants.

Irreversible Rx A (red) → B (cyan) → C (blue) k_1 = 0.2, k_2 = 0.6	Irreversible Rx A (red) → B (cyan)→ C (blue) k_1= 0.6, k_2 = 0.2
	

7.2.2.2.3: Consecutive Reversible First-Order Reactions

You can imagine that solving the equations for the completely reversible reactions of A ↔ B ↔ C would be very difficult. However, writing the differential equations for each step is straightforward and can be done easily in Vcell by choosing the built-in equations for each separate reaction based on mass action. The program can then solve the equations numerically to produce progress curve graphs.

Now let's look at the simulation for the fully **reversible** reactions A ↔ B ↔ C. Again, the model was built and solved in VCell, and then exported in the system's biology markup language (SBML) format. The interactive graphs are made using a program called miniSideWinder.

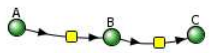
A note: Arrows in VCell Diagrams - In the reaction diagram for the reversible reaction A ↔ B ↔ C below, the arrows go in only one direction, left to right, which simply shows that the species are connected. However, in the Vcell program, the equations for the

reversible reaction were used to produce the graphs below. To run the simulation of the irreversible reaction, the rate constants for the reverse reaction would be set to 0.

VCell MODEL

Reversible reaction A ↔ B ↔ C.

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model)



$$A \underset{k_{1r}}{\overset{k_{1f}}{\rightleftharpoons}} B \underset{k_{2r}}{\overset{k_{2f}}{\rightleftharpoons}} C$$

Initial parameter values: $k_{1f} = 0.2$, $k_{1r} = 0.1$, $k_{2f} = 0.6$, k2r = 0.3 $A_0 = 1$

Select Load [model name] below

Load A to B to C_RevNoEnz

Select **Start** to begin the simulation.

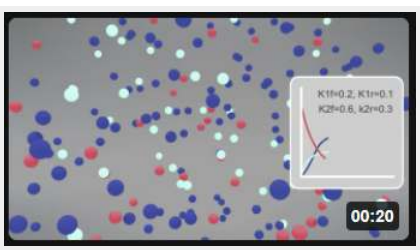
Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Here is the corresponding animation for the fully reversible reaction A ↔ B ↔ C.

Reversible Rx A (red) ↔ B (cyan) ↔ C Blue)
$k_{1f} = 0.2$, $k_{1r} = 0.1$; $k_{2f} = 0.6$, $k_{2r} = 0.3$



7.2.2.3: Summary

This chapter introduces the fundamental principles of chemical kinetics using single-step and multi-step reactions as models. The focus is on understanding how reaction rates are described mathematically, both for simple, uncatalyzed reactions and for more complex reversible and consecutive processes. Key topics include:

1. **Calculus in Kinetics**
   The chapter begins by establishing the language of calculus for describing reaction rates. Students learn to interpret the derivative (dX/dt) as the instantaneous rate of change in concentration and to use integrals to derive expressions that relate

concentration to time. This mathematical framework provides the basis for writing and solving differential equations that describe reaction progress.

2. **First-Order Reactions**
   Using the simple irreversible reaction A → P as a model, the integrated rate equation

   $[A]=[A]_0e-k_1t$

   is derived and explained. Graphical representations, such as exponential decay plots of [A] versus time and linear plots of ln[A] versus time, are used to illustrate how the reaction rate decreases as the concentration of A diminishes.

3. **Second-Order and Pseudo First-Order Reactions**
   The chapter then extends the discussion to second-order reactions, exemplified by reactions involving two molecules (A + B → P) or a self-reaction (A + A → P). The corresponding integrated rate equation,

   $1[A]=1[A]_0+k_2t$

   is derived. Special cases, such as pseudo first-order kinetics (when one reactant is in large excess), are introduced to show how the complexity of bimolecular reactions can be simplified.

4. **Initial Rate vs. Integrated Rate Methods**
   Two approaches for kinetic analysis are compared. The initial rate method involves measuring the slope of the concentration versus time curve at the very beginning of the reaction, while progress curve analysis relies on integrated rate equations to describe the entire time course of the reaction. Each method provides unique insights into the reaction mechanism and rate constants.

5. **Multi-Step and Reversible Reactions**
   The concepts are further developed by considering reversible reactions (A ↔ P) and consecutive reactions (A → B → C). Differential equations for these systems are formulated and solved to yield progress curves that depict how reactant and product concentrations change over time. The dynamic nature of equilibrium in reversible reactions is emphasized, illustrating that even when concentrations plateau, the system remains dynamic.

6. **Computational Modeling and Visualization**
   Finally, the chapter highlights the role of computational tools such as VCell in solving complex kinetic equations. Interactive simulations and animations help visualize progress curves and the impact of varying rate constants on reaction dynamics, thereby linking mathematical models with experimental observations.

In summary, this chapter equips students with a strong quantitative foundation in reaction kinetics, bridging theoretical mathematical concepts with practical applications in chemical and enzymatic processes. Understanding these principles is essential for analyzing more complex biochemical reactions and for developing an intuition about how molecular processes unfold over time.

fully reversible reaction A ↔ B ↔ C

This page titled 7.2.2: Kinetics without Enzymes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.2: Kinetics without Enzymes** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.2.3: Kinetics with Enzymes

Learning Goals (ChatGPT o1, 1/30/25)

Below is a series of learning goals designed to help junior and senior biochemistry majors master the principles of enzyme-catalyzed reactions, kinetic derivations, and data analysis:

- **Understanding Enzyme Catalysis and Reaction Thermodynamics**
  - Explain how enzymes accelerate the rate at which equilibrium is achieved without altering the overall equilibrium constant (Keq) of a reaction.
  - Describe the concept that enzymes lower the activation energy by stabilizing bound transition states and, in doing so, alter the reaction mechanism.
- **Mechanistic Foundations of Enzyme-Catalyzed Reactions**
  - Analyze the basic enzyme-catalyzed reaction scheme, $E+S \leftrightarrow ES \rightarrow E+P$, and discuss the significance of each step in the pathway.
  - Understand the physical meaning of binding substrate to an enzyme and how this enables intramolecular conversion of substrate to product before release.
- **Rapid Equilibrium vs. Steady State Assumptions**
  - Differentiate between the rapid equilibrium and steady state assumptions in enzyme kinetics and identify conditions under which each is applicable.
  - Derive the expression for the enzyme-substrate complex concentration [ES][ES] using the rapid equilibrium assumption and mass balance for enzyme, leading to $[ES]=E_0[S]/(K_S+[S])$.
- **Derivation and Interpretation of the Michaelis-Menten Equation**
  - Derive the Michaelis-Menten equation ($v_0=V_{max}[S]/(K_M+[S])$ under both rapid equilibrium and steady state conditions, understanding the assumptions and limitations of each derivation.
  - Define and interpret the kinetic parameters K_M (Michaelis constant), V_{max} (maximum velocity), and k_{cat} (turnover number), including their units and physical significance.
- **Graphical Analysis and Data Interpretation**
  - Explain how initial rate (v_0) measurements are obtained from progress curves and how these data are used to generate hyperbolic vv vs. [S] plots.
  - Describe methods for linearizing Michaelis-Menten kinetics (e.g., Lineweaver-Burk, Eadie-Hofstee, and Scatchard plots) and discuss their advantages and potential pitfalls.
- **Integration of Reaction Progress Curves**
  - Differentiate between analyses based on initial velocities and those based on integrated rate equations (progress curves) for enzyme-catalyzed reactions.
  - Understand how integrated rate equations for enzyme kinetics can be derived and used to extract kinetic parameters from time-course data.
- **Modeling Reversible and Complex Enzyme-Catalyzed Reactions**
  - Analyze reversible enzyme-catalyzed reactions and understand how both forward and reverse reaction rates contribute to overall kinetics.
  - Describe kinetic mechanisms that involve covalent intermediates (such as acyl-enzyme complexes) and incorporate these into an extended Michaelis-Menten framework.
- **Application of Computational Tools in Kinetic Modeling**
  - Utilize computer modeling platforms (e.g., VCell) to simulate enzyme-catalyzed reactions and visualize progress curves under varying conditions.

- Interpret simulation outputs to understand the effects of changing kinetic constants (e.g., k_1, k_2, k_3) on reaction profiles and the dynamic equilibrium between substrate, enzyme–substrate complex, and product.

By achieving these learning goals, students will develop a comprehensive understanding of the quantitative and mechanistic aspects of enzyme kinetics—from deriving fundamental equations to applying them in experimental and computational contexts.

An enzyme alters the pathways for converting a reactant to a product by binding to the reactant and facilitating the intramolecular conversion of bound substrate to bound product before it releases the product. Enzymes do not affect the thermodynamics of reactions. For reversible reactions (as an example), the equilibrium constant, $K_{eq,}$ is unchanged. What is charged is the rate at which equilibrium is achieved. Enzymes lower the activation energy for bound transition states and change the reaction mechanism.

Figure 7.2.3.1 shows the simplest chemical reaction that can be written to show how an enzyme catalyzes a reaction.

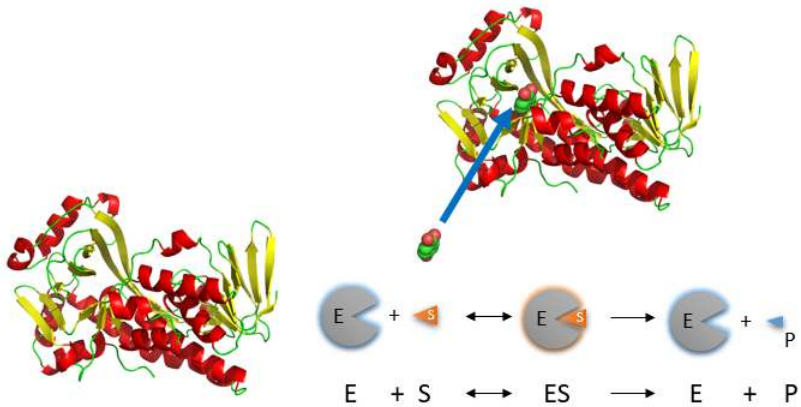


Figure 7.2.3.1: Chemical equation for a simple enzyme-catalyzed reaction

7.2.3.1: Rapid Equilibrium Enzyme-Catalyzed Reactions

We have previously derived equations for the reversible binding of a ligand to a macromolecule. Next, we derived equations for the receptor-mediated facilitated transport of a molecule through a semipermeable membrane. This latter case extended the former case by adding a physical transport step. Now, in what hopefully will seem like *deja vu*, we will derive almost identical equations for the chemical transformation of a ligand, commonly referred to as a substrate, into a product by an enzyme. We will study two scenarios based on two different assumptions, each enabling a straightforward mathematical derivation of kinetic equations:

1. **Rapid Equilibrium Assumption** - enzyme E (macromolecule) and substrate S (ligand) concentrations can be determined using the dissociation constant since E, S, and ES are in rapid equilibrium, as we previously used in deriving the equations for facilitated transport. Sorry about the switch from A to S in the designation of the substrate. Biochemists use S to represent the substrate (ligand) and A, B, P, and Q to represent reactants and products in the case of multi-substrate and multi-product reactions.
2. **Steady State Assumption** (more general) - enzyme and substrate concentrations are not those determined using the dissociation constant.

Enzyme kinetics experiments, as we will see in the following chapters, must be used to determine the detailed mechanism of the catalyzed reaction. Using kinetic analysis, you can determine the order of binding/dissociation of substrates and products, the rate constants for individual steps, and clues to the mechanism used by the enzyme in catalysis.

Consider the following reaction mechanism for the enzyme-catalyzed conversion of substrate S into product P. (We will assume that the catalyzed rate is much greater than the noncatalyzed rate.)

$$\mathbf{E + S \overset{K_s}{\rightleftharpoons} ES \overset{k_3}{\rightarrow} E + P}$$

As we did for the derivation of the equations for the facilitated transport reactions under rapid equilibrium conditions, this derivation is based on the assumption that the relative concentrations of S, E, and ES can be determined by the dissociation constant, K_S, for the interactions and the concentrations of each species during the early part of the reaction (i.e., under initial rate conditions). Assume also the S >> E_0. Remember that under these conditions, S does not change much with time. Is this a valid assumption? Examine the mechanism shown above. S binds to E with a second-order rate constant k_1. ES has two fates. It can dissociate with a first-order rate constant k_2 to S + E, or it can be converted to the product with a first-order rate constant of k_3 to give P + E. If we assume that k_2 >> k_3 (i.e., that the complex falls apart much more quickly than S is converted to P), then the relative ratios of S, E, and ES can be described by Ks. Alternatively, you can think about it this way. If S binds to E, most of S will dissociate, and a small amount will be converted to P. If it does, then E is now free and will quickly bind S and reequilibrate since the most likely fate of bound S is to dissociate, not be converted to P (since $k_3 \ll k_2$). This also makes sense if you consider that the physical step, characterized by k_2, is likely to be quicker than the chemical step, characterized by k_3. Hence, the following assumptions have been used:

- $S \gg E_0$
- $P_0 = 0$
- k_3 is rate limiting (i.e., the slow step)

We will derive equations showing the initial velocity v as a function of the initial substrate concentration, S_0, assuming that P is negligible over the time period used to measure the initial velocity. Also assume that $v_{catalyzed} \gg v_{noncatalyzed}$. In contrast to the first-order reaction of S to P in the absence of E, v is not proportional to S_0 but to S_{bound}. Therefore, $v \propto [ES]$, or

$$v = const[ES] = k_3[ES] \tag{7.2.3.1}$$

where v is the velocity (i.e., reaction rate).

Now, let's get ES from the dissociation constant K_S (assuming rapid equilibrium of E, S, and ES) and mass balance for E (E_0 = E + ES, so E = E_0 - ES). We will use mass balance for S when we derive the equation of the steady state. In many cases, S is approximately equal to S_0 or the total amount of substrate. This makes sense if you consider that the enzyme is a catalyst that acts repeatedly to produce the product, so you don't need significant amounts of the enzyme. The resulting equation of [ES] is shown below.

$$(ES) = \frac{(E_0)\,(S)}{K_S + (S)} \tag{7.2.3.2}$$

? A derivation of [ES] under rapid equilibrium conditions

Here it is!

Derivation

$$\begin{gathered} K_S = \frac{[E][S]}{[ES]} = \left(\frac{([E_0] - [ES])\,[S])}{[ES]} \\ (ES)K_S = (E_0)\,(S) - (ES)(S) \\ (ES)K_S + (ES)(S) = (E_0)\,(S) \\ (ES)\,(K_S + (S)) = (E_0)\,(S) \end{gathered} \tag{7.2.3.3}$$

This derivation assumes that we know S (equal to S_0) and E_{tot} (which is E_0).

Let us assume that S is much greater than E, as is the likely biological case. We can calculate ES using the following equations and the same procedure we used for the derivation of the binding equation, which gives the equation below:

$$[ES] = \frac{[E_0]\,[S]}{K_S + S} \tag{7.2.3.4}$$

which is analogous to

$$[ML] = \frac{[M_0]\,[S]}{K_D + L} \tag{7.2.3.5}$$

This leads to

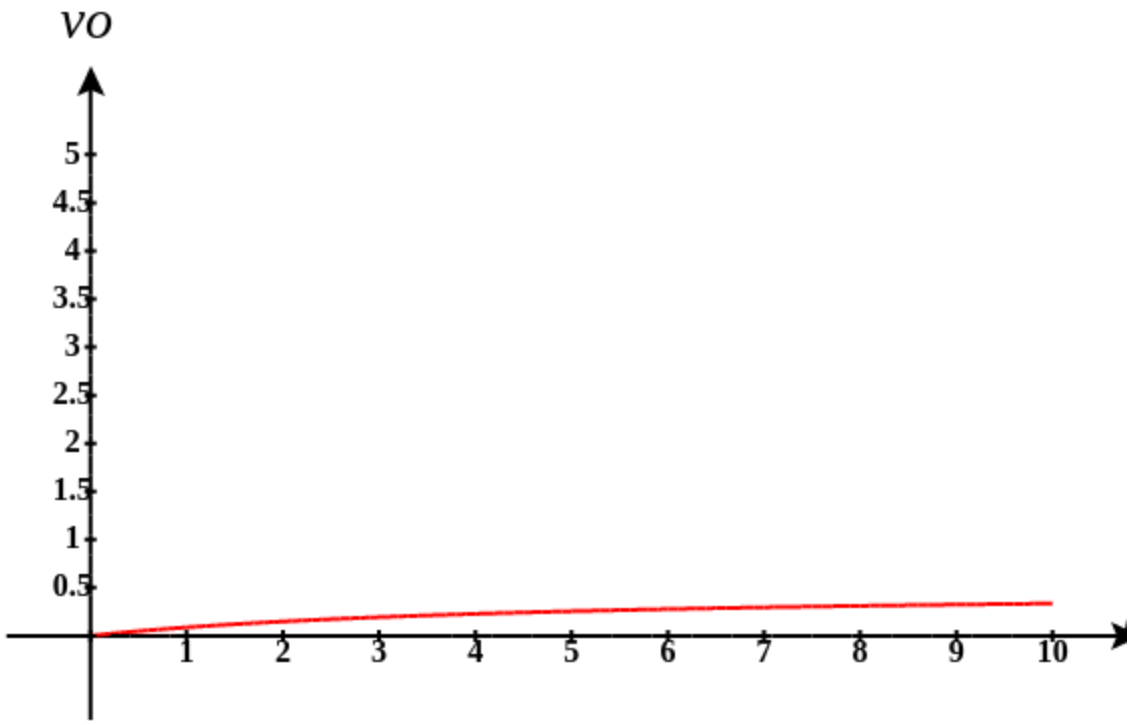
$$v_0 = \frac{(k_3)(E_0)(S)}{K_S + (S)} = \frac{(V_M)(S)}{K_S + (S)} \tag{7.2.3.6}$$

where

$$V_M = k_3 E_0 \tag{7.2.3.7}$$

This is the world-famous Henri-Michaelis-Menten Equation. It is a hyperbola just like the graph for the binding of a ligand to a macromolecule with a given dissociation constant, K_D.

Move the sliders in the interactive graph below to see how changing Km and Vm alters the graph. Note that this graph is identical to the graph for M + L ↔ ML.



Just as in the case with noncatalyzed first-order decay, it is easiest to measure the initial velocity of the reaction when [S] does not change much with time, and the velocity is constant (i.e., the slope of the dP/dt curve is constant). A plot of [P] vs t (called a **progress curve**) is made for each studied substrate concentration. From these curves, the initial rates at each [S] is determined.

Alternatively, one reaction time that gives a linear rise in [P] with time is determined for all the different substrate concentrations. At that specified time, the reaction can be stopped (quenched) with a reagent that does not cause any change in S or P. Then, initial rates can be easily calculated for each [S] from a single data point.

Under these conditions:

- a plot of v vs. S is hyperbolic

- v = 0 when S = 0
- v is a linear function of S when S<<Ks.
- v = V_{max} (or V_M) when S is much greater than Ks
- S = K_S when v = $V_M/2$.

These are the same conditions we detailed for our understanding of the binding equation

$$(ML) = \frac{(M_0)\,L}{K_D + L} \tag{7.2.3.8}$$

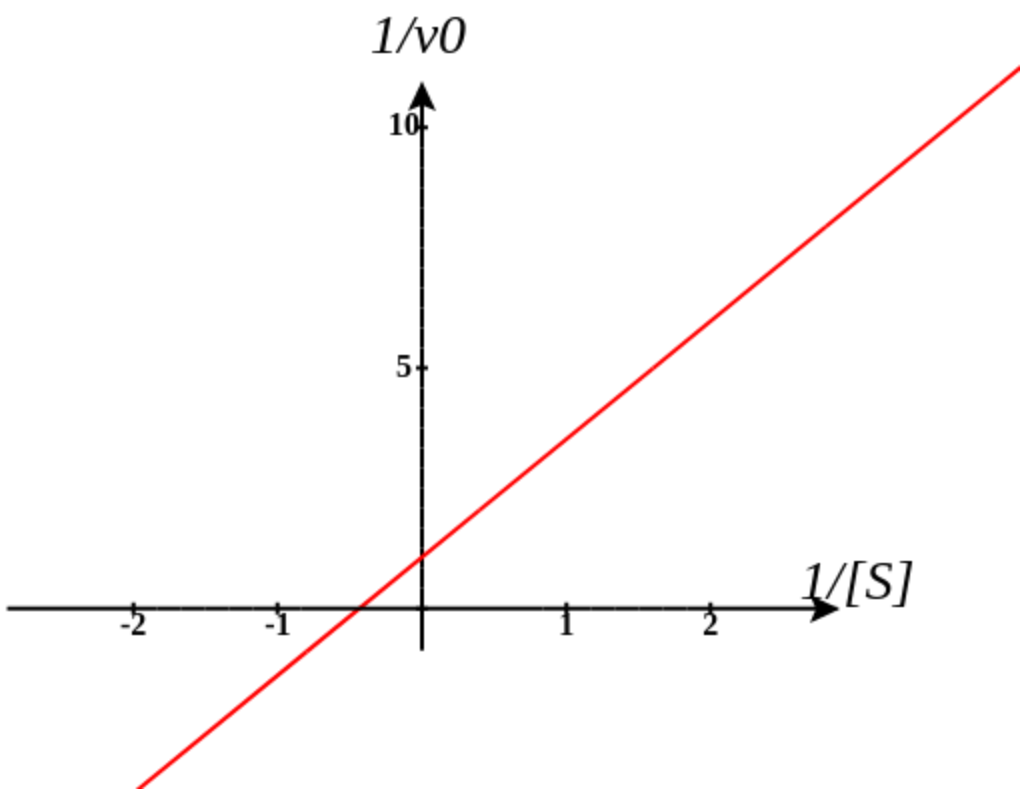
Note that when S is not >> K_S, the graph does not reach saturation or look hyperbolic. It should be apparent from the graph that only if S >> K_S (or when S is approximately 100 x > K_s) will saturation be achieved.

The K_S constant is usually called the **Michaelis constant, K_M**. We will see in a bit that the K_M for most enzyme-catalyzed reactions is **not** equal to the dissociation constant for ES, which we call K_S.

These graphs are often transformed into **double reciprocal or Lineweaver-Burk** plots, as shown below.

$$\frac{1}{v_0} = \frac{K_M + S}{V_M S} = \left(\frac{K_M}{V_M}\right)\frac{1}{S} + \frac{1}{V_M} \tag{7.2.3.9}$$

These plots are used to estimate V_M from the 1/v intercept ($1/V_M$) and K_M from the 1/S axis ($-1/K_M$). These values should be used as "seed" values for a nonlinear fit to the hyperbola that models the actual v vs. S curve. An interactive graph of $1/v_0$ vs 1/[S] is shown below. Change the sliders for K_M and V_M and note the change in the slope and intercepts of the plot

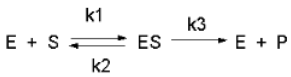


Help

As we saw in the graph of A or P vs t for a noncatalyzed, first-order reaction, the velocity of the reaction, given as the slope of those curves, is always changing. Which velocity should we use in Equation 5? The answer invariably is the initial velocity, v_0,

measured in the early part of the reaction when little substrate is depleted. Hence, v vs S curves for enzyme-catalyzed reactions invariably are really v_0 vs [S] curves.

7.2.3.2: Steady State Enzyme-Catalyzed Reactions



In this derivation, we will consider the following equations and all the rate constants, and will not arbitrarily assume that $k_2 >> k_3$. We will still assume that $S >> E_0$ and that $P_0 = 0$. An added assumption, however, is that d[ES]/dt is approximately 0. Look at this assumption this way. When an excess of S is added to E, ES is formed. In the rapid equilibrium assumption, we assumed that it would fall back to E + S (a physical step) faster than it would go onto the product (a chemical step). In the steady state case, we will assume that ES might go on to product either less or more quickly than it will fall back to E + S. In either case, a steady state concentration of ES arises within a few milliseconds, and its concentration does not change significantly during the initial part of the reaction under which the initial rates are measured. Therefore, d[ES]/dt is about 0. For the rapid equilibrium derivation, $v = k_3[ES]$. We then solved for ES using K_S and the mass balance of E. In the steady state assumption, the equation $v= k_3[ES]$ still holds, but now we will solve for [ES] using the steady state assumption that d[ES]/dt =0.

$$\frac{d[ES]}{\mathrm{dt}} = \mathrm{k_1}[E][S] - \mathrm{k_2}[ES] - \mathrm{k_3}[ES] = 0 \tag{7.2.3.10}$$

We can solve this and obtain the Michaelis-Menten equation for the reaction.

$$\mathrm{v} = \mathrm{k_3}[\mathrm{ES}] = \frac{\mathrm{k_3}\,[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2+k_3}}{\mathrm{k_1}} + \mathrm{S}} = \frac{\mathrm{V_M}[\mathrm{S}]}{\mathrm{K_M} + \mathrm{S}} \tag{7.2.3.11}$$

? Derivation of the Michaelis-Menten Equation for the steady state

To see how to derive this, click below.

Answer

Applying mass balance for E (i.e $E = E_0$-ES) in the appropriate term below gives

$$\begin{aligned}
&\mathrm{k_1}[E][S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0} - \mathrm{ES}]\,[S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] - \mathrm{k_1}[\mathrm{ES}][S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] + \mathrm{k_1}[\mathrm{ES}][S] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] = [ES]\,(\mathrm{k_2} + \mathrm{k_3} + \mathrm{k_1S})
\end{aligned} \tag{7.2.3.12}$$

Solving for [ES] gives

$$[\mathrm{ES}] = \left(\frac{[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2+k_3}}{\mathrm{k_1}} + \mathrm{S}}\right) \tag{7.2.3.13}$$

Substituting into the Henri Michaelis-Menten equation gives

$$\mathrm{v} = \mathrm{k_3}[\mathrm{ES}] = \frac{\mathrm{k_3}\,[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2+k_3}}{\mathrm{k_1}} + \mathrm{S}} = \frac{\mathrm{V_M}[\mathrm{S}]}{\mathrm{K_M} + \mathrm{S}} \tag{7.2.3.14}$$

Note that

$$V_M = k_3 E_0 \tag{7.2.3.15}$$

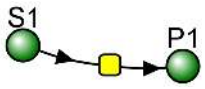
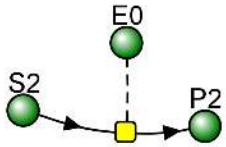
and

$$K_M = \frac{k_2 + k_3}{k_1} \tag{7.2.3.16}$$

Now let's look at a progress curve simulation that compares the Michaelis-Menten equation derived from rapid equilibrium assumptions (when $k_2 >> k_3$) and in which K_M is the actual dissociation constant to the steady state approximation when k_2 is not $>> k_3$. Reaction 1 is the rapid equilibrium, and Reaction 2 is the steady state.

VCell MODEL

A comparison of the rapid equilibrium (left) and steady state (right)Michaels-Menten reaction

	
$$v_0 = \frac{k_3\,[E_O]\,[S]}{K_M + S} = \frac{V_M[S]}{K_M + S} \quad (7.2.3.17)$$	$$v_0 = \frac{k_3\,[E_O]\,[S]}{\frac{k_2+k_3}{k_1} + S} = \frac{V_M[S]}{K_M + S} \quad (7.2.3.18)$$
Initial Condition: $V_M = 10$; $K_M = 10$; $S = 50$	Initial Conditions: $k_1 = 10$; $k_2 = 90$; $k_3 = 10$; $E_0 = 1$ uM; $S = 50$ $V_M = k_3E_0 = 10$; $K_M = (k_2 + k_3)/k_2 = 10$)

Select Load [model name] below

Load MMvsSteadyState

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The initial conditions in the graph are set so the graphs of the rapid equilibrium and steady state are identical.

Recommendations:

- Scale the y-axis to 50
- change the slider k_{2_r2} for the steady state graph to other values. Watch the curves separate. Although these plots are only for one substrate concentration (50 uM), the effects of changing k_2 for the steady state plot are very dramatic. This should convince you that, in general, unless $k_3 << k_2$, the calculated value of K_M is **not** equal to the thermodynamic dissociation constant, K_D.

7.2.3.3: Analysis of the General Michaelis-Menten Equation

This equation can be simplified and studied under different conditions. First, notice that $(k_2 + k_3)/k_1$ is a constant, which is a function of relevant rate constants. This term is usually replaced by K_M, which is called the **Michaelis constant**. Likewise, when S approaches infinity (i.e., $S >> K_M$, equation 5 becomes $v = k_3(E_0)$, which is also a constant called V_M for maximal velocity. Substituting V_M and K_M into equation 5 gives the simplified equation:

$$\mathrm{v} = \frac{\mathrm{V_M[S]}}{\mathrm{K_M + S}} \quad (7.2.3.19)$$

It is extremely important to note that K_M in the general equation does **n**ot equal the K_S, the dissociation constant used in the rapid equilibrium assumption! K_M and K_S have the same units of molarity, however. A closer examination of K_M shows that under the

limiting case when $k_2 >> k_3$ (the rapid equilibrium assumption), then,

$$K_M = \frac{k_2 + k_3}{k_1} = \frac{k_2}{k_1} = K_D = K_S \tag{7.2.3.20}$$

If we examine these equations under several different scenarios, we can better understand the equation and the kinetic parameters:

- when S = 0, v = 0.
- when $S >> K_M$, $v = V_M = k_3E_0$. (i.e., v is zero order with respect to S and first order in E. Remember, k_3 has units of $s^{-1>}$ since it is a first-order rate constant. k_3 is often called the **turnover number,** because it describes how many molecules of S "turn over" to product per second.
- $v = V_M2$, when $S = K_M$.
- when $S << K_M$, $v = V_MS/K_M = k_3E_0S/K_M$ (i.e., the reaction is bimolecular, dependent on both S and E. k_3/K_M has units of $M^{-1}s^{-1}$, the same as a second-order rate constant.

7.2.3.4: More Complicated Enzyme-catalyzed Reactions

7.2.3.4.1: A reversibly-catalyzed reaction

You have learned previously that enzymes don't change the equilibrium constant for a reaction but rather decrease the activation energy barrier to move from reactants to products. This implies that the activation energy to move in the reverse direction, from products to reactants, is also lowered. Hence, the enzyme speeds up both the forward and reverse reactions. We haven't accounted for that yet in our kinetic equations. Many reactions in metabolic reactions that do not have large negative values of ΔG (ie, they are not significantly favored) are reversible, allowing the enzyme to be used in the reverse direction. Take, for example, the pathway to break down glucose to pyruvate (glycolysis). It has nine steps, of which five are reversible, allowing them to be used in the reverse pathway to take pyruvate to glucose (gluconeogenesis).

Let's set up the equations for the reversible reaction of substrate S to product P catalyzed by enzyme E. Assume that the K_M for the forward reaction is K_{MS} (or K_S) and for the reverse reaction is K_{MP} (or K_P) as shown in the reaction scheme in Figure 7.2.3.2. The rate constant k_2 is the k_{cat} (forward rate constant) for conversion of ES to EP, and k_{-2} is the k_{cat} (reverse rate constant) for conversion of EP to ES

$$E + S \underset{K_S}{\overset{K_{MS}}{\leftrightarrow}} ES \underset{k_{-2}}{\overset{k_2}{\leftrightarrow}} EP \underset{K_P}{\overset{K_{MP}}{\leftrightarrow}} E + P$$

Figure 7.2.3.2: Reaction scheme for a simple enzyme-catalyzed reaction

The following simple Michaelis-Menten equations can be written for just the forward reaction and for the reverse reaction:

$$\begin{aligned} v_f &= \frac{V_f S}{K_{MS} + S} = \frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} \\ v_r &= \frac{V_r P}{K_{MP} + P} = \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \end{aligned} \tag{7.2.3.21}$$

You might think that simply subtracting the two would give the net velocity in the forward direction, but that is **NOT** the case.

$$v \neq \left[\frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} - \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \right] \tag{7.2.3.22}$$

The reason is that in the derivation, the equations for both the forward and reverse rates must have terms for the reverse and forward reactions, respectively.

A simple derivation shows that this is the equation for the reversible conversion of substrate to product.

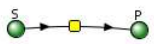
$$v = k_2[ES] - k_{-2}[EP] \tag{7.2.3.23}$$
$$= \frac{V_f \frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} - \frac{V_r \frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} = \frac{V_f \frac{[S]}{K_S} - V_r \frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]}$$

This reversible form of the Michaelis-Menten equation and other equations, written in the format shown in equation 6.24, are commonly used in programs such as VCell and Copasi to model the kinetics of whole pathways of biological interactions and reactions. The figures below show a reaction diagram, graphical results showing S and P vs time for the selected K_M and V_M values, and animations for the reaction. The chemicals (S and P) are shown as green spheres connected by a line. The red dot again represents the enzyme (shown as a node of connection between S and P). Equation 6.24 was used to model the reversible reaction (even though the arrows shown between S and P are unidirectional.

VCell MODEL

Reversible Enzyme-Catalyzed Reaction: E + S ↔ ES ↔ EP ↔ E + P .

Vcell reaction diagram (1-way arrows defined as reversible in the actual mathematical model) and chemical equation



J	reaction rate	$\frac{\left(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev}\right)}{\left(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev}\right)}$
KmFwd	Km forward	10.0
VmaxFwd	max forward rate	10.0
KmRev	Km reverese	20.0
VmaxRev	max reverse rate	5.0
Size	size	1.0
initConc	initial concentration for S	5.0
initConc	initial concentration for P	0.0

Initial parameter values: as shown in above

Select Load [Enz Rev] below

Load Enz Rev

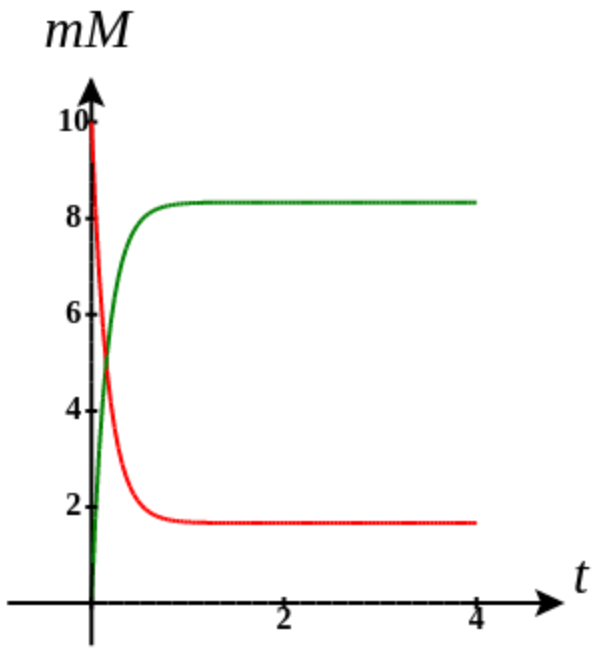
Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

If you reflect on it, this reaction is very similar to the reversible reaction of A ↔ P in the **absence** of an enzyme, which we explored in Chapter section 6.2. Just for comparison, the graph for that reaction is shown below. Change the sliders to produce a curve similar to the enzyme-catalyzed reaction shown in the figure above.



7.2.3.4.2: Reaction with intermediates

Not all reactions can be characterized simply as a substrate interacting with an enzyme to form an ES complex, which then turns over to form the product. Sometimes, intermediates form. For example, a substrate S might interact with E to form a complex, which then is cleaved to products P and Q. Q is released from the enzyme, but P might stay covalently attached. This often happens in the hydrolytic cleavage of a peptide bond by a protease when an activated nucleophile, like serine, reacts with the sessile peptide bond in a nucleophilic substitution reaction, releasing the amine end of the former peptide bond as the leaving group. The carboxy end of the peptide bond remains bonded to the Ser as a Ser-acyl intermediate. Water then enters and cleaves the acyl intermediate, freeing the carboxyl end of the original peptide bond. This is shown in the written reaction in Figure **7.2.3.3**:

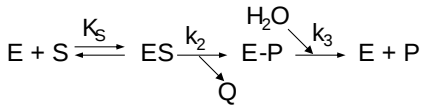


Figure **7.2.3.3**: Reaction scheme for enzyme-catalyzed reaction with an intermediate

Even for this seemingly complicated reaction, you get the standard Michaelis-Menten equation.

To simplify the derivation of the kinetic equation, let's assume that E, S, and ES are in rapid equilibrium defined by the dissociation constant, Ks. Assume Q has a visible absorbance, so it is easy to monitor. Assume from the steady state assumption that:

$$\frac{d[E-P]}{dt} = k_2[ES] - k_3[E-P] = 0 \tag{7.2.3.24}$$

assuming that k_3 is a pseudo first-order rate constant and that $[H_2O]$ doesn't change.

The velocity depends on which step is rate-limiting. If $k_3 << k_2$, then the k_3 step is rate-limiting. Then

$$v = k_3[E-P] \tag{7.2.3.25}$$

If $k_2 << k_3$, then the k_2 step is rate-limiting. Then

$$v = k_2[ES] \tag{7.2.3.26}$$

The following kinetic equation for this reaction can be derived, assuming $v = k_2[ES]$.

$$v = \frac{\frac{k_2k_3}{k_2+k_3}[E_0][S]}{\left(\frac{k_3}{k_2+k_3}\right)K_S + S} = \frac{k_{cat}[E_0][S]}{K_M + S} = \frac{V_M S}{K_M + S} \tag{7.2.3.27}$$

You can verify that you get the same equation if you assume that $v = k_3[E-P]$.

This equation looks quite complicated, especially if you substitute for Ks, k_{-1}/k_1. All the kinetic constants can be expressed as functions of the individual rate constants. However, this equation can be simplified by realizing the following:

- When $S >> \frac{K_S k_3}{k_2+k_3}, v = \left(\frac{k_2k_3}{k_2+k_3}\right)E_0 = V_M$
- $\frac{K_S k_3}{k_2+k_3} = \text{constant} = K_M$

Substituting these into equation 7 gives:

$$v = \frac{V_M S}{K_M + S}$$

This, again, is the general form of the Michaelis-Menten equation

The expression for V_M in the first bulleted expression above is more complicated than our earlier definition of $V_M = k_3E_0$. They are similar in that the term E_0 is multiplied by a constant which is itself a function of rate constant(s). The rate constants are generally lumped together into a generic constant called k_{cat}.

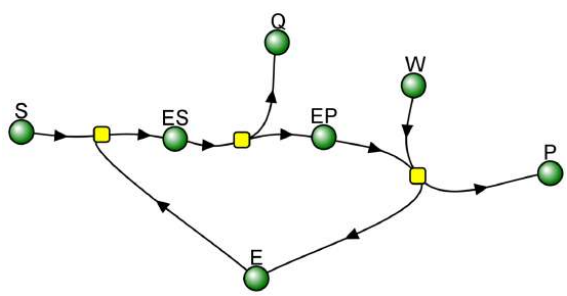
- For the simple reaction $\mathbf{k_{cat} = k3}$
- For the more complicated reaction above with a covalent intermediate, $k_{cat} = \frac{k_2k_3}{k_2+k_3}$
- For all reactions, $V_M = k_{cat}E_0$.

Recent Updates (2/7/24): Here is a VCell computational model for a bisubstrate and biproduct (BiBi) reaction with a covalent intermediate (ping pong reaction).

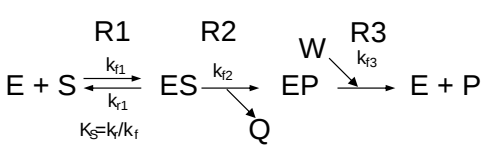
VCell MODEL

BiBi-Ping Pong_Covalent Intermediate Irreversible reaction

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



Yellow dots: Reaction Nodes (R1, R2 and R3 left to right)



Reaction made irreversible since $k_{r2} = 0$, $k_{r3} = 0$.

Initial parameter values:

- S_0 = 100, E_0 = 1, W (water in a hydrolysis reaction) = 50 and fixed throughout
- k_{1f} = 5, k_{1r} = 1, k_{2f} = 0.6,
- k_{2f} = 50, k_{2r} = 0
- k_{3f} = 0.05, k_{3f} = 0

Select Load [model name] below

Load BiBiPingPongChymo

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

To see the burst phase for reaction, change the time and parameters to these values:

- set Run time to 0.3
- Select **Plot** then Update Y axis max to 2
- Click Edit Plot Species and check just P and Q
- reset

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

When k_2 >> k_3, Q forms quickly in a **burst phase** due to a high k_2. This phase is followed by the slower conversion of the accumulated EP to E + P in a **steady state phase**, as shown below in Figure **7.2.3.4**. Make sure to change the parameters in the model above to see the burst phase for the production of product Q. This reaction is an example of a BiBi (two substrates, two products) Ping Pong reaction, since one reactant binds, followed by one product departing before the second reactant binds and the second product departs.

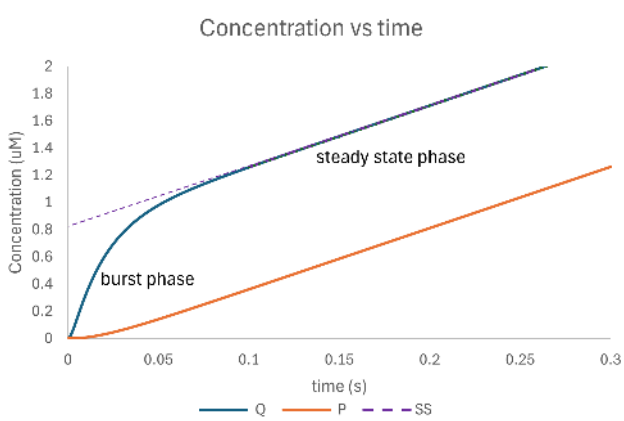


Figure **7.2.3.4**: Burst and steady-state phase in an enzyme-catalyzed reaction with a covalent intermediate and k_2 >> k_3.

Reverse the rate constants in the model to k_2 = 5 and k_3 = 0.5 so that k_2 is **not** >> k_3, and note that the burst phase for Q disappears!

When plotting v_0 vs S Michaelis-Menten plots of this type of reaction, the state state v_0 should be used and **NOT** $(v_0)_{t=0}$ = $(dQ/dt)_{t=0}$, since that rate changes very quickly after the initial burst formation of Q.

Summary to this point

Figure 7.2.3.5 compares the Michaelis-Menten kinetic equations for the rapid equilibrium, steady state assumptions, and covalent intermediate cases.

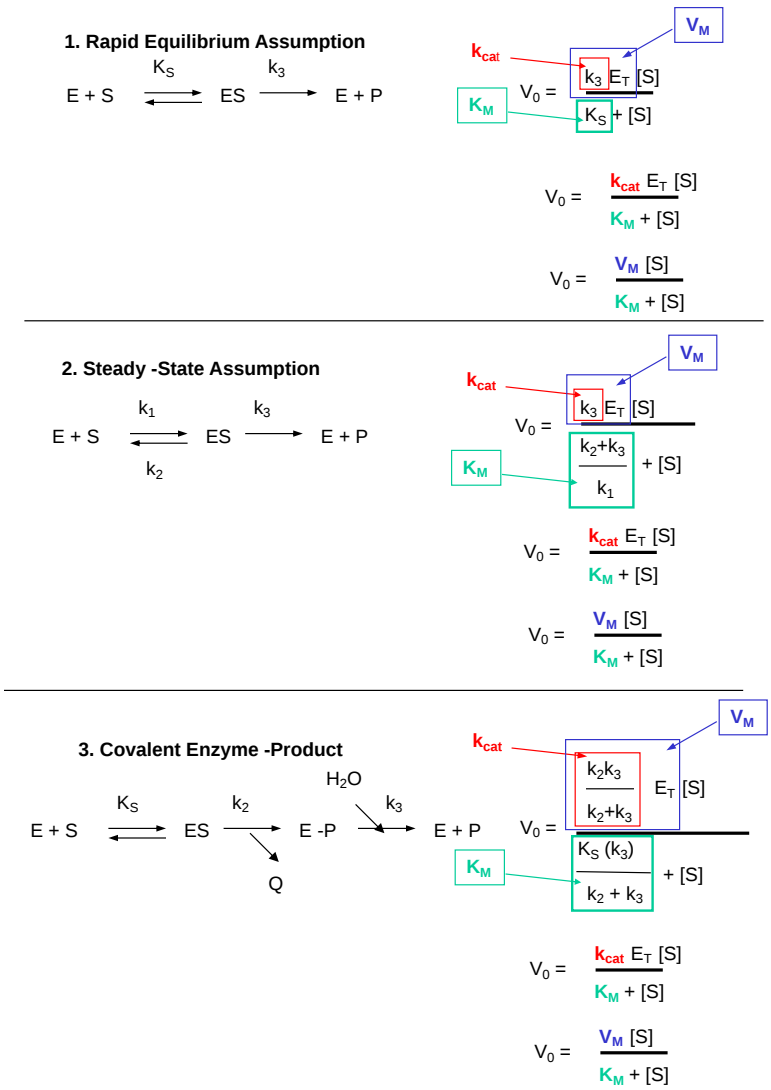


Figure 7.2.3.5: Michaelis-Menten kinetic equations for the rapid equilibrium, steady state assumptions, and covalent intermediate cases

7.2.3.5: Meaning of Kinetic Constants

Getting a "gut-level" understanding of the significance of the rate constants is important. Here they are:

- K_M: The Michaelis constant, with units of molarity (M), is operationally defined as the substrate concentration at which the initial velocity is half of V_M. It is equal to the dissociation constant of E and S only if E, S, and ES are in rapid equilibrium. It can be thought of as an "**effective**" (but not actual) K_D in other cases.
- k_{cat}: The catalytic rate constant, with units of s^{-1}, is often called the **turnover number**. It is a measure of how many bound substrate molecules "turnover" or form product in 1 second. This is evident from equation $v_0 = k_{cat[}ES]$
- k_{cat}/K_M: Under condition when $[S] << K_M$, the Michaelis-Menten equation becomes $v_0 = (k_{cat}/K_M)[E_0][S]$. This really describes a biomolecular rate constant (k_{cat}/K_M), with units of $M^{-1}s^{-1}$, for the conversion of free substrate to product. Some enzymes have k_{cat}/Km values around 10^8, indicating they are diffusion-controlled. That implies that the reaction is essentially done as soon as the enzyme and substrate collide. The constant k_{cat}/Km is also referred to as the **specificity constan**t in that it describes how

well an enzyme can differentiate between two different and competing substrates. (We will show this mathematically in the next chapter.)

Table 7.2.3.1 below shows K_M and k_{cat} values for various enzymes

K_M values		
enzyme	substrate	Km (mM)
catalase	H_2O_2	25
hexokinase (brain)	ATP	0.4
	D-Glucose	0.05
	D-Fructose	1.5
carbonic anhydrase	HCO_3^-	9
chymotrypsin	glycyltyrosinylglycine	108
	N-benzoyltyrosinamide	2.5
b-galactosidase	D-lactose	4.0
threonine dehydratase	L-Thr	5.0
k_{cat} values		
enzyme	substrate	k_{cat} (s^{-1})
catalase	H_2O_2	40,000,000
carbonic anhydrase	HCO_3^-	400,000
acetylcholinesterase	acetylcholine	140,000
b-lactamase	benzylpenicillin	2,000
fumarase	fumarate	800
RecA protein (ATPase)	ATP	0.4

Table 7.2.3.1: K_M and k_{cat} values for various enzymes

Table 7.2.3.2 below shows k_{cat}, K_M, and k_{cat}/K_M values for diffusion-controlled enzymes

Enzymes with k_{cat}/K_M values close to diffusion-controlled (10^8 - 10^9 $M^{-1}s^{-1}$)					
enzyme	substrate		kcat (s^{-1})	Km (M)	kcat/Km ($M^{-1}s^{-1}$)
acetylcholinesterase	acetylcholine		1.4 x 10^4	9 x 10^{-5}	1.6 x 10^8
carbonic anhydrase	CO_2		1 x 10^6	1.2 x 10^{-2}	8.3 x 10^7
	HCO_3^-		4 x 10^5	2.6 x 10-2	1.5 x 10^7
catalase	H_2O_2		4 x 10^7	1.1	4 x 10^7
crotonase	crotonyl-CoA		5.7 x 10^3	2 x 10^{-5}	2.8 x 10^8
fumarase	fumarate		8 x 10^2	5 x 10^{-6}	1.6 x 10^8
	malate		9 x 10^2	2.5 x 10^{-5}	3.6 x 10^7
triose phosphate isomerase	glyceraldehyde-3-P		4.3 x 10^3	4.7 x 10^{-4}	2.4 x 10^8
b-lactamase	benzylpenicillin		2.0 x 10^3	2 x 10^{-4}	1 x 10^8

Table 7.2.3.2 below shows k_{cat}, K_M and k_{cat}/K_M values for diffusion-controlled enzymes

7.2.3.6: Experimental Determination of V_M and K_M

How can V_M and K_M be determined from experimental data?

7.2.3.6.1: From the initial rate data

The most common way to determine V_M and K_M is through initial rates, v_0, obtained from P or S vs time curves. Hyperbolic graphs of v_0 vs [S] can be fitted or transformed as we explored the different mathematical transformations of the hyperbolic binding equation to determine K_D. These included:

- Michaelis-Menten plot: nonlinear hyperbolic fit
- Lineweaver-Burk double reciprocal plot
- Scatchard plot
- Eadie-Hofstee plot

We discussed all of these plots, except for the **Eadie-Hofstee** plot, in the chapter on binding. The Eadie-Hofstee plot is another linearized version of the Michaelis-Menten equation

Here is a derivation of that equation, which starts with each side of the double-reciprocal plot being multiplied by v_0V_M.

$$\begin{aligned}(v_0V_M)\frac{1}{v_0} &= (v_0V_M)\left(\frac{K_M}{V_M}\right)\frac{1}{S} + (v_0V_M)\frac{1}{V_M} \\ V_M &= (v_0)\,K_M\frac{1}{S} + (v_0) \\ v_0 &= -K_M\left(\frac{v_0}{S}\right) + V_M\end{aligned} \tag{7.2.3.28}$$

Note that a graph of v_0 vs v_0/S is linear, so slopes and intercepts can be used to obtain values for V_M and K_M.

The double-reciprocal plot is commonly used to analyze initial velocity vs substrate concentration data. When used for such purposes, the graphs are referred to as **Lineweaver-Burk** plots, where plots of 1/v vs 1/S are straight lines with slope m = K_M/V_M, and y-intercept b = $1/V_M$. Figure 7.2.3.6 common graphs used to display initial rate enzyme kinetic data.

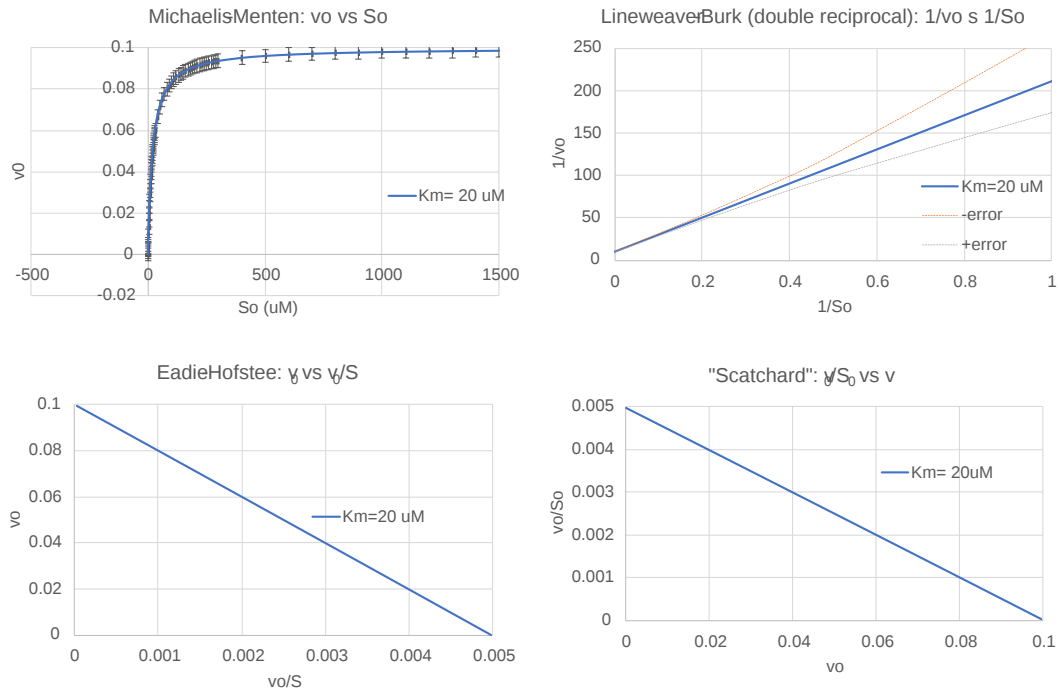


Figure **7.2.3.6**: Common graphs used to display initial rate enzyme kinetic data

The straight-line plots shown above should **not** be analyzed using linear regression since simple linear regression assumes constant error in v_0 values. A weighted linear regression or, even better, a nonlinear fit to a hyperbolic equation should be used. (Common Error in Biochemistry Textbooks: The Shape of the Hyperbola). A rearrangement of the corresponding Scatchard equations in the Eadie-Hofstee plot is also commonly used, especially to visualize enzyme inhibition data, as we will see in the next chapter.

7.2.3.7: An Extension: k_{cat} and V_M from integrated rate equations

K_M and V_M could be theoretically extracted from progress curves of A or P as a function of t at one single A concentration by deriving an integrated rate equation for A or P as a function of t, as we did in equation 2 (the integrated rate equation for the conversion of A → P in the absence of enzyme). In principle, this method would be better than the initial rate method. Why? It is not easy to be certain you are measuring the initial rate for each and every [S], which should vary over a wide range. It's also time-intensive. In addition, think about how much data is discarded if you take an entire progress curve at each substrate concentration, especially if you quench the reaction at a given time point, which effectively limits the data to one time point per substrate.

In practice, the mathematics is complicated, and it is impossible to get a simple, explicit function of [P] or [S] as a function of time. A slight variant of a progress curve can be derived. Let us consider the simple case of a single substrate S (or A) being converted to product P in an enzyme-catalyzed reaction. The analogous equations for first-order, noncatalyzed rates were $A=A_0e^{-k_1t}$ or $P = A_0(1-e^{-k_1t})$.

We can derive the equation for the enzyme-catalyzed reaction shown below.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

Here it is!

$$\frac{P}{t} = \frac{K_M \ln\left(\frac{S_0-P}{S_0}\right)}{t} + V_M \qquad (7.2.3.29)$$

? Derivation: k_{cat} and V_M from integrated rate equation

Click below to see the derivation

Derivation

$$\begin{gathered} v = -\frac{dS}{dt} = +\frac{dP}{dt} = \frac{V_M S}{K_M + S} \\ \int_{S_0}^{S} \frac{K_M + S}{V_M S} dS = -\int_0^t t \end{gathered} \tag{7.2.3.30}$$

$$-t = \frac{S + K_M \ln S - S_0 - K_M \ln S_0}{V_M} \tag{7.2.3.31}$$

On rearrangement, this gives:

$$S_0 - S + K_M \ln \frac{S_0}{S} = V_M t \tag{7.2.3.32}$$

This equation is an **implicit** equation, not an explicit one, as it does NOT give S(t) explicitly as a function of t.

Equation yy can be written with respect to product P as follows:

$$\begin{gathered} P = S_0 - S'' \text{ or } ''S = S_0 - P \\ S_0 - (S_0 - P) + K_M \ln \frac{S_0}{S_0 - P} = V_M t \\ P - K_M \ln \left(\frac{S_0 - P}{S_0} \right) = V_M t \end{gathered} \tag{7.2.3.33}$$

Rearranging this gives

$$\frac{P}{t} = \frac{K_M \ln \left(\frac{S_0 - P}{S_0} \right)}{t} + V_M \tag{7.2.3.34}$$

S(t) explicitly as a function of t.

This equation does not give P(t) explicitly as a function of time. Rather, one can get a graph of P/t vs [ln (1-P/S_0)]/t (shown below) from the derived equation, which *does* give a straight line with a slope of Km and a y-intercept of Vm. Note that the calculated values of V_M and K_M are derived from only one substrate concentration, and the values may be affected by product inhibition.

Figure **7.2.3.7** compares a first-order noncatalyzed conversion of A → P to the enzyme-catalyzed rate. The V_M for the enzyme-catalyzed reaction was chosen to be small to make the two graphs comparable.

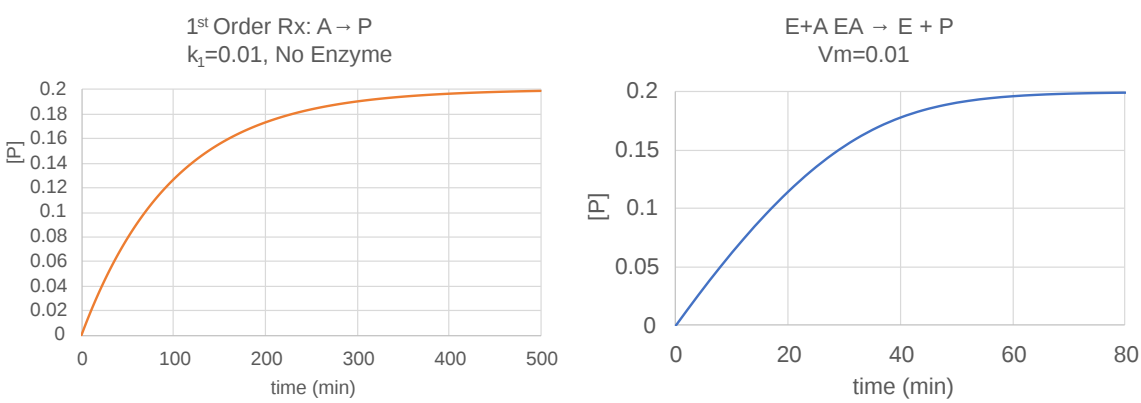


Figure **7.2.3.7**: Comparison of a first order noncatalyzed conversion of A → P to the enzyme catalyzed rate

Note that the curves are similar but not identical. If you didn't know an enzyme was present, you could fit the data to a first-order rise in [P] with time, but it would not be the optimal fit. The progress curves are a lot more complicated to analyze if the product, which shares structural similarities with the substrate, binds the enzyme tightly and inhibits it (called product inhibition).

7.2.3.7.1: Comparison of Progress Curves for Uncatalyzed and Catalyzed Reactions.

Let's explore progress curves for enzyme-catalyzed reactions a bit further. Students usually see v_0 vs. [S] Michaelis-Menten plots in textbooks. These plots are, in some ways, less intuitive than seeing P vs. t curves, which are more in line with how we might contemplate how a reaction proceeds. Hence, it would be illuminating to compare progress curve graphs of A → P (irreversible) for the uncatalyzed and S → P for the enzyme-catalyzed reactions. What might you expect? We saw one example in Figure **7.2.3.6**.

In each case, P should increase with time. In the uncatalyzed reaction, S exponentially decreases to 0, and P concomitantly rises to a value of $P = S_0$. You also get a rise in P vs t for the enzyme-catalyzed rate, but you would think it would be a faster rise with time since the enzyme catalyzes the reaction. Figure **7.2.3.87** shows a comparison of the progress curves for the uncatalyzed first-order reaction of $A \rightarrow P_1$ (red) and $S \rightarrow P_2$ (blue) for enzyme- catalyzed reaction (blue, right) for these conditions: uncatalyzed reaction $A \rightarrow P_1$, $k_1 = 0.1$; Catalyzed reaction: $S \rightarrow P_2$, V_M=10, K_M=5. Note that the rate at which bound S (i.e, ES) goes to P for the catalyzed rate is 100x faster than the rate constant for the catalyzed rate.

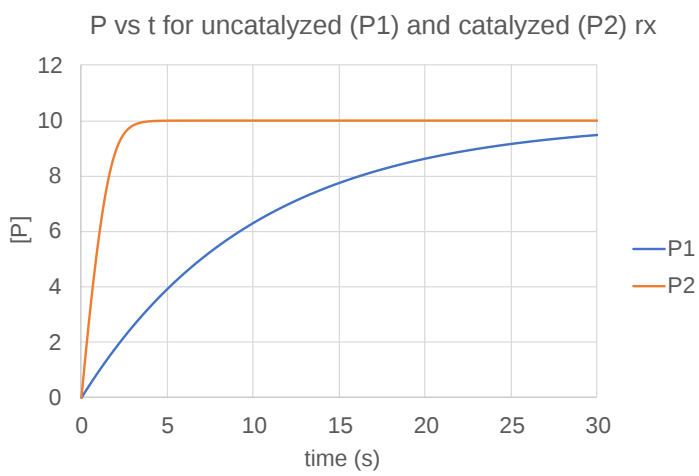


Figure **7.2.3.8**: Another comparison of progress curves (P vs t) for a catalyzed and uncatalyzed conversion of A → P.

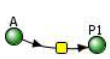
Note that the curves are somewhat similar in shape but also clearly different in comparison to the one shown in Figure **7.2.3.6**.

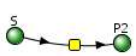
Now, let's use Vcell to compare the reactions for different values of the kinetic constants for the uncatalyzed and the enzyme-catalyzed reactions. Change the constants and try to find a set of conditions so that the catalyzed and uncatalyzed rates for the conversion of reactions to products are superimposable. How can that be?

VCell MODEL

Irreversible reactions: A → P1 and E + S ↔ ES → E + P

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation





Initial conditions

J	reaction rate	$(Kf \cdot A - Kr \cdot P1)$
Kf	forward rate constant	0.1
Kr	reverse rate constant	0.0
J	reaction rate	$\frac{Vmax \cdot S}{(Km + S)}$
Km	Km (1/2 max)	5.0
Vmax	max reaction rate	10.0
Size	size	1.0
initConc	initial concentration for A	10.0
initConc	initial concentration for P1	0.0
initConc	initial concentration for S	10.0
initConc	initial concentration for P2	0.0

Select Load [model name] below

Load Compare Irrev Rx No EnzW ithEnz

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Animations

Now let's look at an animation of the same irreversible reactions in which the enzyme-catalyzed reaction is no faster than the noncatalyzed rate (a worthless enzyme!). Here are the reactions:

$A \rightarrow P$, k_1 =0.1; $S \rightarrow P$, V_M=?, K_M = ?.

The animations show just the accumulation of products. Animations are by Shraddha Nakak and Hui Liu.

Exercise 7.2.3.1

Use the Vcell model above to find a set of values for K_M and V_M that would make the two graphs superimposable - i.e., when the graphs in the absence and presence of E are identical. (Hint: that would be a really bad enzyme if it didn't increase the reaction over the uncatalyzed rate!)

Answer

$K_M = 96$, $V_M = 10$

7.2.3.8: Summary

This chapter delves into the fundamental principles of enzyme catalysis and the quantitative analysis of reaction kinetics. It explains how enzymes accelerate chemical reactions by binding substrates, stabilizing transition states, and facilitating intramolecular conversions, all without changing the overall thermodynamics or equilibrium constant (Keq) of the reaction. Instead, enzymes dramatically increase the rate at which equilibrium is reached by lowering the activation energy.

Key topics covered include:

1. **Mechanistic Overview of Enzyme Catalysis**
   - **Reaction Pathway:** The enzyme binds to the substrate to form an enzyme–substrate complex (ES), converts the bound substrate to a bound product via a lower-energy transition state, and finally releases the product.
   - **Thermodynamics vs. Kinetics:** Although enzymes do not alter the equilibrium constant, they modify the reaction pathway by lowering activation barriers and changing the mechanism.
2. **Derivation of the Michaelis-Menten Equation**
   - **Rapid Equilibrium Assumption:** Under conditions where the substrate concentration far exceeds that of the enzyme, the rapid equilibrium between enzyme, substrate, and the ES complex allows the derivation of $[ES]=E_0[S]/(K_S+[S])$.
   - **Steady State Assumption:** A more general approach assumes that the formation and breakdown of ES reach a quasi-steady state (i.e., $d[ES]/dt\approx0$, leading to the same functional form of the Michaelis-Menten equation: $v_0=V_{max}[S]/(K_M+[S])$ where $V_{max}=k_{cat}E_0$ and $K_M=(k_2+k_3)/k_1$.
3. **Graphical and Mathematical Analysis of Kinetic Data**
   - **Progress Curves and Initial Rates:** The chapter explains how to measure reaction velocities from progress curves (concentration vs. time plots) and emphasizes the importance of using initial rate data when substrate depletion is minimal.
   - **Linearization Techniques:** Various plots (e.g., Lineweaver-Burk, Eadie-Hofstee) are discussed as tools to extract KMK_M and VmaxV_{max} from experimental data, while noting the potential pitfalls and the value of nonlinear curve fitting.
4. **Complex Enzyme-Catalyzed Reactions**
   - **Reversible Reactions:** The extension of the Michaelis-Menten framework to reversible reactions is explored, showing how both forward and reverse reactions are incorporated into kinetic models.
   - **Reactions with Intermediates:** The chapter addresses mechanisms involving covalent intermediates (e.g., acyl-enzyme intermediates in proteases) and discusses how these multi-step processes still conform to a generalized Michaelis-Menten equation under appropriate assumptions.
5. **Computational Modeling and Simulation**
   - The use of computational tools, such as VCell, is highlighted to simulate reaction progress curves and model complex kinetic systems. Interactive graphs and animations illustrate how changes in kinetic parameters (e.g., rate constants, substrate concentration) influence the reaction trajectory, thereby reinforcing the concepts developed through the analytical derivations.

In summary, this chapter provides a comprehensive treatment of enzyme kinetics, from basic principles and derivations to experimental data analysis and computational modeling. The insights gained here lay the groundwork for understanding more complex biochemical pathways and for the rational design of experiments to probe enzyme mechanisms.

This page titled 7.2.3: Kinetics with Enzymes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.3: Kinetics with Enzymes** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.2.4: Enzyme Inhibition

Learning Goals (ChatGPT o1, 1/30/25)

- **Understanding Irreversible Covalent Inhibition:**
  - Explain how irreversible inhibitors modify key amino acid side chains (e.g., cysteine modification by iodoacetamide) to permanently inactivate an enzyme.
  - Recognize the difference between enzyme denaturation (via extremes of pH and temperature) and irreversible covalent modification.
- **Mechanisms of Reversible Enzyme Inhibition:**
  - Define competitive inhibition and describe how substrate and inhibitor bind mutually exclusively to the active site.
  - Interpret the competitive inhibition kinetic equation, $v_0=V_MS/(K_M(1+I/K_{is})+S)$, and explain how this alters the apparent $K_{M\ aoo}$ without affecting V_M.
- **Graphical Analysis of Competitive Inhibition:**
  - Interpret Lineweaver-Burk plots for competitive inhibition, identifying that the y-intercept ($1/V_M$) remains constant while the x-intercept ($–1/K_M$) shifts as inhibitor concentration increases.
  - Explain how plotting v_0 vs. log [S] can reveal the sigmoidal shifts associated with competitive inhibition.
- **Exploring Uncompetitive Inhibition:**
  - Describe uncompetitive inhibition where the inhibitor binds only to the enzyme–substrate (ES) complex, leading to a decrease in both V_M and K_M.
  - Understand that Lineweaver-Burk plots for uncompetitive inhibition display parallel lines due to proportional decreases in both V_M and K_M.
- **Noncompetitive and Mixed Inhibition:**
  - Distinguish between noncompetitive inhibition (where the inhibitor binds to both free enzyme and ES with equal affinity) and mixed inhibition (with differing affinities).
  - Interpret kinetic expressions for noncompetitive/mixed inhibition and explain how they affect V_M and K_M (e.g., K_M remains unchanged in pure noncompetitive inhibition, while both parameters change in mixed inhibition).
  - Analyze interactive graphs and Lineweaver-Burk plots to visualize these effects.
- **Product Inhibition and Competing Substrate Effects:**
  - Explain how product inhibition occurs when the reaction product, due to its structural similarity to the substrate, binds to the enzyme and decreases the overall reaction rate.
  - Understand the concept of specificity constants (k_{cat}/K_M) and how competing substrates can effectively act as inhibitors.
- **In Vivo vs. In Vitro Inhibition:**
  - Contrast the experimental conditions for enzyme inhibition studies in vitro (fixed enzyme, varying substrate) versus in vivo (dynamic flux through metabolic pathways).
  - Interpret how metabolic flux and substrate accumulation can influence the effectiveness of competitive versus uncompetitive inhibitors in a cellular context.
- **Environmental Effects on Enzyme Activity:**
  - Describe how temperature and pH changes affect enzyme activity, including the role of pH in altering the protonation states of key catalytic residues.
  - Predict how variations in temperature or pH might lead to reversible changes in enzyme kinetics or irreversible enzyme denaturation.

By achieving these learning goals, students will gain a thorough understanding of both the mechanistic basis of enzyme inhibition and the practical techniques used to measure and analyze inhibitor effects. This knowledge forms a critical

foundation for exploring enzyme regulation in metabolic pathways and designing effective pharmaceutical inhibitors.

7.2.4.1: Irreversible Covalent Inhibition

Given what you already know about protein structure, it should be easy to determine how to inhibit an enzyme. Since structure mediates function, anything that would significantly alter an enzyme's structure would inhibit the enzyme's activity. Hence, extremes of pH and high temperature, all of which can denature the enzyme, would irreversibly inhibit the enzyme unless it could refold properly. Alternatively, we could add a small molecule, which interacts noncovalently with the enzyme to either change its conformation or directly prevent substrate binding. Finally, we could covalently modify certain side chains, which, if they are essential to enzymatic activity, would irreversibly inhibit the enzyme.

We discussed previously the types of reagents that would chemically modify specific side chains that might be critical for enzymatic activity. For example, iodoacetamide might abolish enzyme activity if a cysteine side chain is required for activity. However, these reagents will usually modify several side chains. Determining which is critical for the binding or catalytic conversion of the substrate can be difficult. One way would be to protect the active site with saturating concentrations of a ligand that binds reversibly at the active site. Then, the chemical modification can be performed at varying reaction times. The critical side chain would be protected from chemical modification, with the extent of protection depending on the K_D, the concentration of the protecting ligand, and the length of the reaction.

The rest of the chapter will deal with **reversible, noncovalent inhibition**

7.2.4.2: Competitive Inhibition

Reversible Competitive inhibition occurs when substrate (S) and inhibitor (I) both bind to the same site on the enzyme. In effect, they compete for the active site and bind in a mutually exclusive fashion. This is illustrated in the chemical equations and molecular cartoons shown in Figure 7.2.4.1.

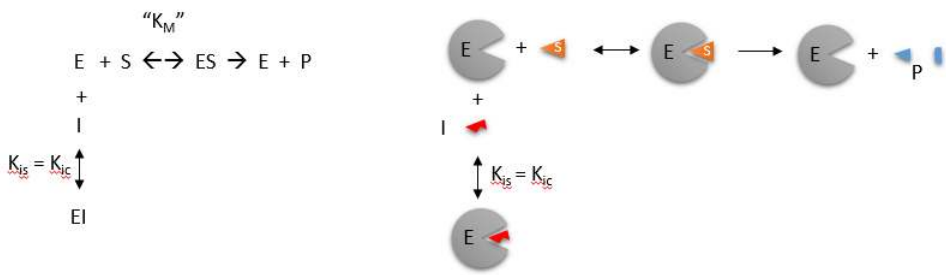


Figure 7.2.4.1: Competitive Inhibition

$$v_0 = \frac{V_M S}{K_M\left(1 + \frac{I}{Kis}\right) + S} \tag{7.2.4.1}$$

There is another type of inhibition that would give the same kinetic data. If S and I were bound to different sites, and S was bound to E and produced a conformational change in E such that I could not bind (and vice versa), then the binding of S and I would be mutually exclusive. This is called allosteric competitive inhibition. Inhibition studies are usually done at several fixed and non-saturating concentrations of I and varying S concentrations.

The key kinetic parameters to understand are V_M and K_M. Let us assume for ease of equation derivation that I binds reversibly and with rapid equilibrium to E, with a dissociation constant K_{IS}. The "s" in the subscript "is" indicates that the slope of the 1/v vs. 1/S Lineweaver-Burk plot changes while the y-intercept stays constant. K_{IS} is also named K_{IC}, where the subscript "c" stands for competitive inhibition constant.

A look at the top mechanism shows that even in the presence of I, as S increases to infinity, all E is converted to ES. That is, there is no free E to which I could bind. Now, remember that $V_M = k_{cat}E_0$. Under these conditions, ES = E_0; hence v = V_M. V_M is not changed. However, the apparent K_M, K_{Mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to E alone and not ES, it will shift the equilibrium of E + S → ES to the left. This would increase the K_{Mapp} (i.e., it would appear that

the affinity of E and S has decreased). The double reciprocal plot (Lineweaver-Burk plot) offers a great way to visualize the inhibition as shown in Figure 7.2.4.2.

$$v_0 = \frac{V_M S}{K_M\left(1 + \frac{I}{Kis}\right) + S}$$

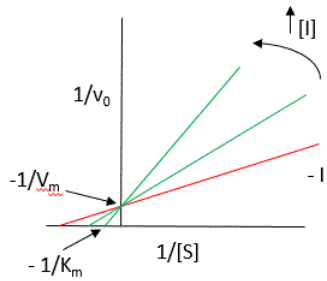
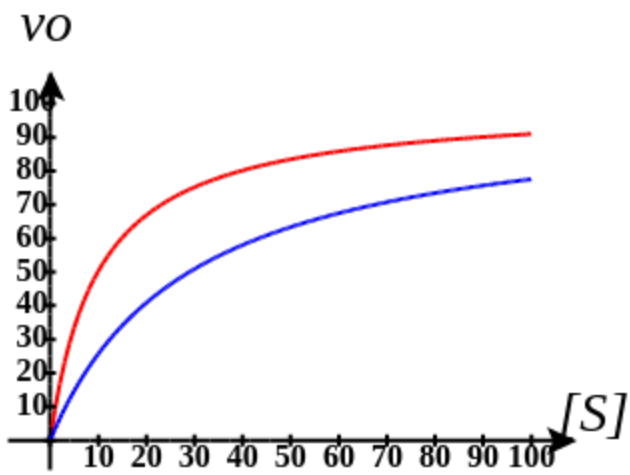


Figure 7.2.4.2: Competitive Inhibition: Lineweaver-Burk plots

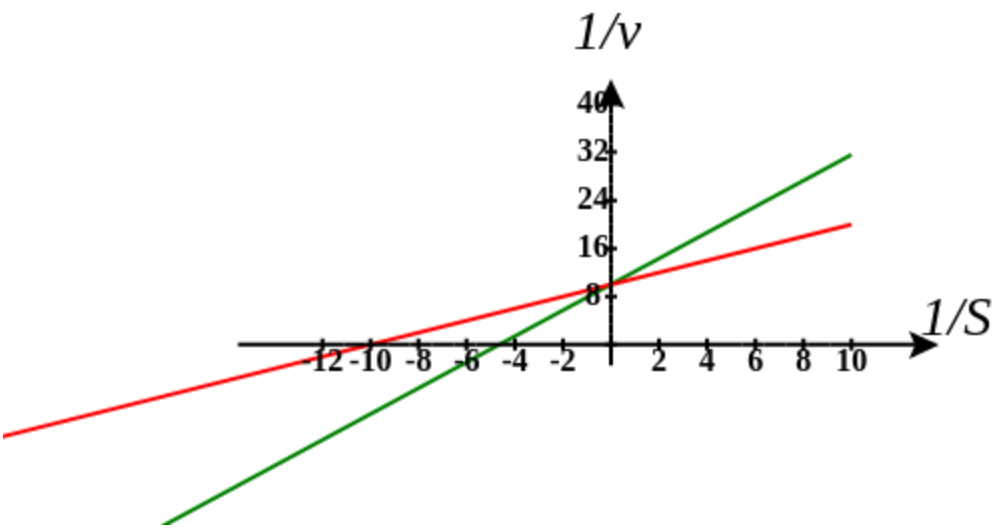
In the presence of I, V_M does not change, but K_M appears to increase. Therefore, $1/K_M$, the x-intercept on the plot, will get smaller and closer to 0. Therefore, the plots will consist of a series of lines, with the same y-intercept ($1/V_M$), and the x-intercepts ($-1/K_M$) closer and closer to 0 as I increases. These intersecting plots are the hallmarks of competitive inhibition.

Here is an interactive graph showing v_0 vs [S] for competitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is the interactive graphs showing $1/v_0$ vs 1/[S] for competitive inhibition, with Vm and Km both set to 10.

Competitive Inhibition



Note that in the first three inhibition models discussed in this section, the Lineweaver-Burk plots are linear in the presence and absence of an inhibitor. This suggests that plots of v vs. S in each case would be hyperbolic and conform to the usual form of the Michaelis-Menten equation, each with potentially different apparent V_M and K_M values.

An equation for v_0 in the presence of a competitive inhibitor is shown in the above figure. The only change compared to the equation for the initial velocity in the absence of the inhibitor is that the K_M term is multiplied by the factor $1+I/K_{is}$. Hence $K_{Mapp} = K_M(1+I/Kis)$. This shows that the apparent K_M does increase as we predicted. K_{IS} is the inhibitor dissociation constant. The apparent K_M changes with increasing I, which affects the slope of the double reciprocal plot.

v_0 vs log [Inhibitor] plots

If the data were plotted as v_0 vs. log S, the plots would be sigmoidal, as we saw for plots of ML vs. log L in Chapter 5B. In the case of a competitive inhibitor, the plot of v_0 vs log S in the presence of different fixed concentrations of inhibitor would consist of a series of sigmoidal curves, each with the same V_M, but with different apparent K_M values (where $K_{Mapp} = K_M(1+I/K_{is})$, progressively shifted to the right. Enzyme kinetic data is rarely plotted this way. These plots are mostly used for simple binding data for the M + L ↔ ML equilibrium in the presence of different inhibitor concentrations.

Reconsider our discussion of the simple binding equilibrium, M + L ↔ ML. For fractional saturation Y vs a log L graphs, we considered three examples:

1. $L = 0.01\ K_D$ (i.e., $L << K_D$), which implies that $K_D = 100L$. Then $Y = L/[K_D+L] = L/[100L + L] \approx 1/100$. This implies that irrespective of the actual [L], if $L = 0.01\ K_D$, then $Y \approx 0.01$.
2. $L = 100\ K_D$ (i.e,. $L >> K_D$), which implies that $K_D = L/100$. Then $Y = L/[K_D+L] = L/[(L/100) + L] = 100L/101L \approx 1$. This implies that irrespective of the actual [L], if L = 100 Kd, then $Y \approx 1$.
3. $L = K_D$, then $Y = 0.5$

These scenarios show that if L varies over four orders of magnitude ($0.01K_D < K_D < 100K_D$), or, in log terms, from $-2 + \log K_D < \log K_D < 2 + \log K_D$), irrespective of the magnitude of the K_D, that Y varies from approximately 0 - 1.

In other words, Y varies from 0-1 when L varies from $\log K_D$ by +2. Hence, plots of Y vs log L for a series of binding reactions of increasingly higher K_D (lower affinity) would reveal a series of identical sigmoidal curves shifted progressively to the right, as shown below in Figure 7.2.4.3.

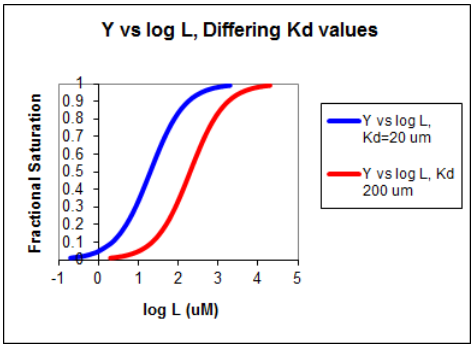


Figure 7.2.4.3: Plot of fractional saturation Y vs log L for two different K_D values.

The same would be true of v_0 vs. S in the presence of different concentrations of a competitive inhibitor, for initial flux, Jo vs ligand outside, in the presence of a competitive inhibitor, or ML vs. L (or Y vs L) in the presence of a competitive inhibitor.

In many ways, plots of v_0 vs. lnS are easier to visually interpret than plots of v_0 vs. S . As noted for simple binding plots, textbook illustrations of hyperbolas are often misdrawn, showing curves that level off too quickly as a function of [S] as compared to plots of v_0 vs lnS, in which it is easy to see if saturation has been achieved. In addition, as the curves above show, multiple complete plots of v_0 vs lnS at varying fixed inhibitor concentrations or for variant enzyme forms (different isoforms, site-specific mutants) over a broad range of lnS can be made, which facilitates comparisons of the experimental kinetics under these different conditions. This is especially true if Km values differ widely.

Now that you are more familiar with binding and enzyme kinetics curves, in the presence and absence of inhibitors, you should be able to apply the above analysis to inhibition curves where the binding or the initial velocity is plotted at varying competitive inhibitor concentrations at different fixed nonsaturating concentrations of ligand or substrate. Consider the activity of an enzyme. Let's say that at some reasonable substrate concentration (not infinite), the enzyme is approximately 100% active. If a competitive inhibitor is added, the activity of the enzyme decreases until at saturating (infinite) I, no activity would remain. Graphs showing this are shown below in Figure 7.2.4.4.

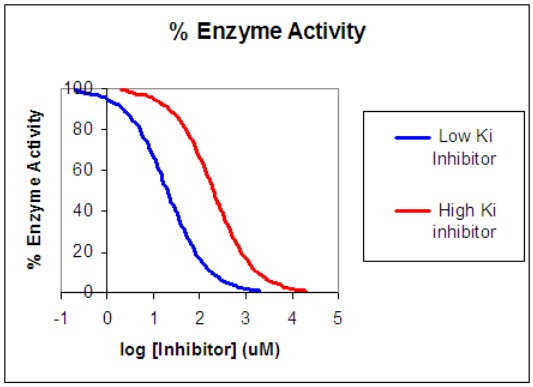


Figure 7.2.4.4: Inhibition of Enzyme Activity - % Activity vs log [Inhibitor]

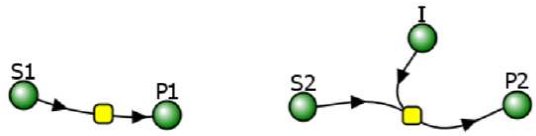
7.2.4.2.1: Progress Curves for Competitive Inhibition

In the previous section, we explored how important progress curve (Product vs time) analyses are in understanding both uncatalyzed and enzyme-catalyzed reactions. We are aware of no textbooks that cover progress curves for enzyme inhibition. Yet progress curves are what most investigators record and analyze to determine initial rates v_0 and to calculate V_M, K_M, and inhibition constants, as described above. We will use Vcell to produce progress curves for reversibly inhibited enzyme-catalyzed reactions.

VCell MODEL

Competitive Inhibition with constant [I]:

No inhibition (left) and competitive inhibition (right)



Initial conditions for no inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial conditions for competitive inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot I}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_CompInhib_MM_I

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed. For this model, select Vm, Km, Ki and I | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The graphs from your initial run show the concentrations of S, P, and I as a function of time for just the initial conditions shown above. In typical initial rate analyses of competitive inhibition, at least three sets of reactions are run with the same varying substrate concentrations and different fixed inhibitor concentrations. In the analyses above, [I] is fixed at 5 uM.

Conduct a series of runs at different values of I. Vary the K_I, the dissociation constant for the EI complex, as follows:

- I << K_I, the dissociation constant for the EI complex
- I >> K_I, the dissociation constant for the EI complex. Then, download the data and determine the initial rate for each of the initial conditions.

Figure 7.2.4.5 shows an interactive iCn3D model of human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT)

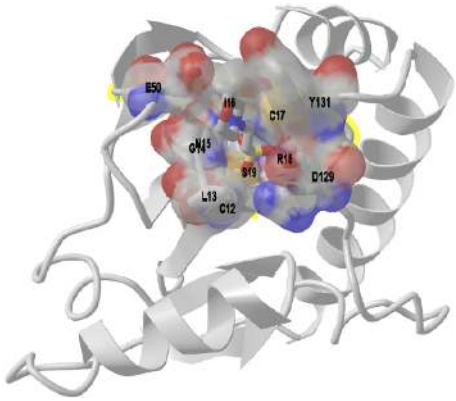


Figure 7.2.4.5: Human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XsEacG2tixDDi9

The competitive inhibitor, the deprotonated form of 2-(N-morpholino)-ethanesulfonic acid (MES), is actually the conjugate base of the weak acid (pKa = 6.15) of a commonly used component of a buffered solution. It is shown in color sticks with the negatively charged sulfonate at the bottom of the active site pocket. The amino acids comprising the active site binding pocket are shown as color sticks underneath the transparent colored surface of the binding pocket. The normal substrates for the enzyme are proteins phosphorylated on tyrosine side chains, so the sulfonate mimics the negatively charged phosphate group of the phosphoprotein target.

7.2.4.2.2: Two special cases of competition inhibition

Product Inhibition

Let's look at an enzyme that converts reactant S to product P. Since P arises from S, they may have structural similarities. For example, what if GTP is the reactant and GDP the product? If so, then P might also bind in the active site and inhibit the conversion of S to P. This is called **product inhibition**. It probably occurs in most enzymes, and when it does occur, it will start bending downward at the beginning of the progress curve for P formation. If the product binds very tightly, it might cause a significant underestimation of the initial velocity (v_0) or flux (J_0) of the enzyme. Let's use Vcell to explore product inhibition. The model will explore two reactions:

- E + R ↔ ER → E + Q (no product inhibition)
- E + S ↔ ES → E + P (with product inhibition)

The chemical equation above does not explicitly show the product P binding the enzyme to form an EP complex. An actual reaction diagram showing the inhibition of an enzyme by an inhibitor I and by the product P is shown in Figure 7.2.4.6 below.

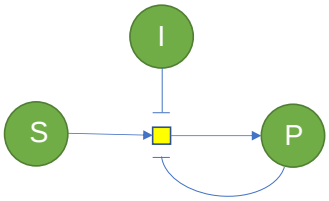


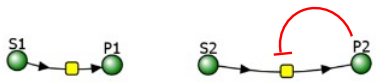
Figure 7.2.4.6: reaction diagram showing inhibition of an enzyme by an inhibitor I and by the product P

Vcell uses much simpler diagrams since it is most often used for modeling whole pathways or even entire cells. In the simpler Vcell reaction diagrams, the inhibitor is typically not shown since the inhibition is built into the equation for the enzyme, represented by the node or yellow square in the figure above.

Let's now explore product inhibition in Vcell. In the reaction without product inhibition, R and Q are the reactant and product, respectively. S and P are used for the reaction *with* product P inhibition.

VCell MODEL

Irreversible MM Kinetics - Without (left rx 1) and With (right, rx 2) Product Inhibition



Initial Conditions: No product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial Conditions: With product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot P2}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
P2	Species Concentration	☑	Variable

Select Load [model name] below

Load ProdInhib_NoProdInhib

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Inhibition by a competing substrate - the specificity constant

In the previous chapter, the specificity constant was defined as k_{cat}/K_M, which we also described as the second-order rate constant associated with the bimolecular reaction of E and S when S << K_M. It also describes how good an enzyme is at differentiating between different substrates. If an enzyme encounters two different substrates, one can be considered a competitive inhibitor of the

other. The following equation gives the ratio of initial velocities for two competing substrates at the same concentration equal to the ratio of their k_{cat}/K_M values.

$$\frac{v_A}{v_B} = \frac{\frac{k_{catA}}{K_A}A}{\frac{k_{catB}}{K_B}B} \tag{7.2.4.2}$$

? A derivation of the specificity constant for an enzyme with competing substrates

Here it is!

Derivation

$$v_A = \frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A} \quad v_B = \frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B} \tag{7.2.4.3}$$

$$\frac{v_A}{v_B} = \frac{\frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A}}{\frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B}} = \frac{\frac{V_A A}{K_A+\frac{K_A B}{K_B}+A}}{\frac{V_B B}{K_B+\frac{K_B A}{K_A}+B}} \tag{7.2.4.4}$$

Now, in the above equation:

multiple the top half of the right-hand expression by

$$\frac{\frac{1}{K_A}}{\frac{1}{K_A}} \tag{7.2.4.5}$$

multiple the bottom half of the right-hand expression by

$$\frac{\frac{1}{K_B}}{\frac{1}{K_B}} \tag{7.2.4.6}$$

replace V_A with $k_{catA}E_0$ and V_B with $k_{catB}E_0$

This gives the following expression for v_A/v_B:

$$\frac{v_A}{v_B} = \frac{\frac{k_{catA}}{K_A}A}{\frac{k_{catB}}{K_B}B} \tag{7.2.4.7}$$

7.2.4.3: Uncompetitive Inhibition

Reversible uncompetitive inhibition occurs when I binds only to ES and not free E. One can hypothesize that on binding S, a conformational change in E occurs, which presents a binding site for I. Inhibition occurs since ESI can not form the product. It is a dead-end complex that has only one fate: to return to ES. This is illustrated in the chemical equations and molecular cartoon shown in Figure 7.2.4.7.

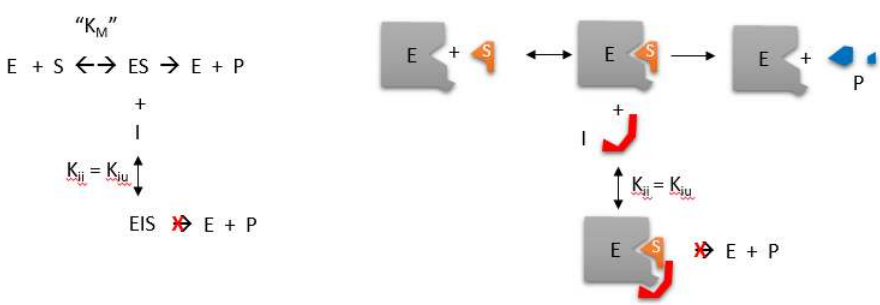


Figure 7.2.4.7: Uncompetitive Inhibition

$$v_0 = \frac{V_M S}{K_M + S\left(1 + \frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1+\frac{I}{Kii}}\right) S}{\left(\frac{K_M}{1+\frac{I}{Kii}}\right) + S} \tag{7.2.4.8}$$

Let us assume that I binds reversibly to ES with a dissociation constant K_{ii}. The second "i" in the subscript "ii" indicates that the intercept of the 1/v vs 1/S Lineweaver-Burk plot changes while the slope stays constant. K_{ii} is also named $K_{iu,}$ where the subscript "u" stands for the uncompetitive inhibition constant.

A look at the top mechanism shows that in the presence of I, as S increases to infinity, not all of E is converted to ES. There is a finite amount of ESI, even at an infinite S. Now remember that $V_m = k_{cat}E_0$ if and only if all E is in the form ES. Under these conditions, the apparent V_m, V_{mapp} is less than the real Vm without the inhibitor. In addition, the apparent K_m, K_{mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to ES alone, and not E, it will shift the equilibrium of E + S ↔ ES to the right, which would have the effect of decreasing the K_{mapp} (i.e., it would appear that the affinity of E and S has increased). The double reciprocal plot (Lineweaver-Burk plot) offers a great way to visualize the inhibition. In the presence of I, both Vm and Km decrease. Therefore, $-1/K_m$, the x-intercept on the plot, will get more negative, and $1/V_m$ will get more positive. It turns out that they change to the same extent. Therefore, the plots will consist of a series of parallel lines, which is the hallmark of uncompetitive inhibition, as shown in Figure 7.2.4.8.

$$v_0 = \frac{V_M S}{K_M + S\left(1 + \frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1+\frac{I}{Kii}}\right) S}{\left(\frac{K_M}{1+\frac{I}{Kii}}\right) + S}$$

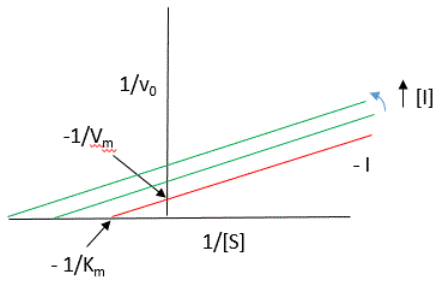
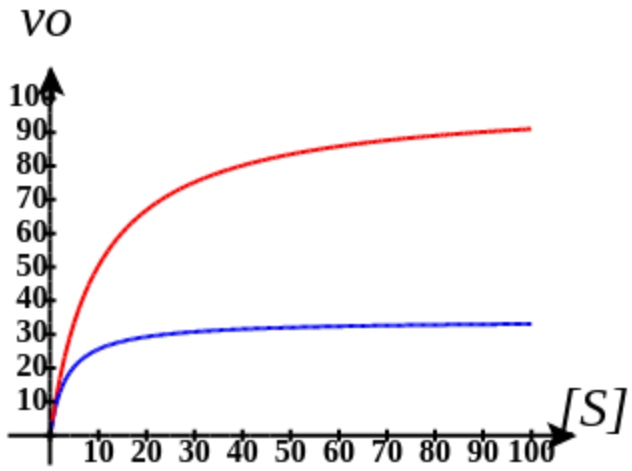


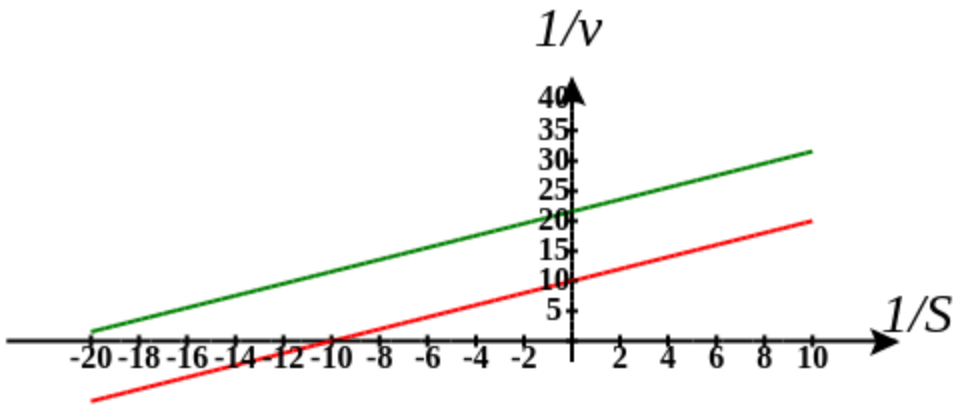
Figure 7.2.4.8: Uncompetitive Inhibition: Lineweaver-Burk plots

Here is an interactive graph showing v_0 vs [S] for uncompetitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is an interactive graph showing uncompetitive inhibition with Vm and Km both set to 10. Change the sliders for [I] and Kii and see the effect on the graph

Uncompetitive Inhibition



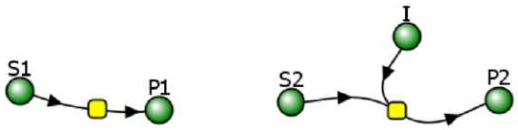
An equation, shown in the diagram above, can be derived to show the effect of the uncompetitive inhibitor on the velocity of the reaction. The only change is that the S term in the denominator is multiplied by the factor 1+I/Kii. We want to rearrange this equation to show how Km and Vm are affected by the inhibitor, not S. Rearranging the equation above shows that Kmapp = Km/(1+I/Kii) and Vmapp = Vm/(1+I/Kii). This shows that the apparent Km and Vm do decrease as we predicted. Kii is the inhibitor dissociation constant in which the inhibitor affects the intercept of the double reciprocal plot. Note that if I is zero, Km and Vm are unchanged.

7.2.4.3.1: Progress Curves: Uncompetitive Inhibition

Now, let's compare the progress curves for an enzyme-catalyzed reaction in the absence and presence of an uncompetitive inhibitor.

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



(Note the the Vcell reaction diagram is the same as for competitive inhibition. However, the mathematical equations differ as shown below.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_UncompInhibMM

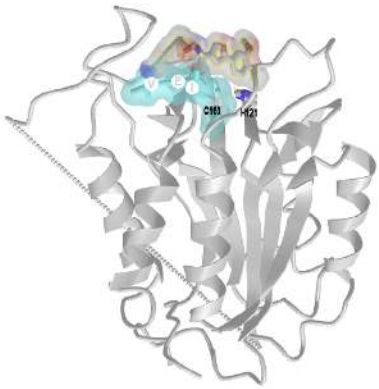
Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Figure **7.2.4.9** shows an interactive iCn3D model of an uncompetitive inhibitor of the cysteine protease caspase-6 (4HVA)



NCBI iCn3D Figure **7.2.4.9**: Uncompetitive inhibitor of the cysteine protease caspase-6 (4HVA) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TVfKio7NxcpX28

The "substrate" in this model (cyan, transparent surface, with labels on V, E, and I) is a substrate analog, VEI-CHO, in which the tripeptide substrate VEI ends not in a free carboxyl or amide group but an aldehyde, which causes this "substrate" to become

covalently attached to the enzyme and act as an inhibitor. The uncompetitive inhibitor (gray transparent surface) binds externally to the blue surface. Hence, it binds to the ES complex. Two active site residues, Cys 163, the active site nucleophile, and His 121, a catalytic acid/base, are shown in colored sticks and labeled.

Figure **7.2.4.10** below shows a Lineweaver-Burk plot showing $1/v_o$ vs 1/[S], where the substrate is the divalent compound $(VEID)_2R110$. Its N-terminus is capped with a benzyloxy (Z) group. R110 is a rhodamine-type fluorophore, which, on cleavage, gives a strong fluorescent signal that the benzyloxy group initially quenches in the uncleaved substrate.

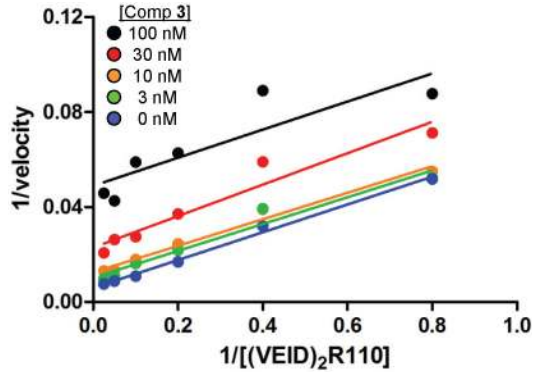


Figure **7.2.4.10**: Double-reciprocal Lineweaver-Burk plot of compound 3 with $(VEID)_2R110$ substrate showing uncompetitive inhibition. Heise CE et al. (2012) Mechanistic and Structural Understanding of Uncompetitive Inhibitors of Caspase-6. PLoS ONE 7(12): e50864. https://doi.org/10.1371/journal.pone.0050864. Creative Commons Attribution License.

7.2.4.4: Noncompetitive and Mixed Inhibition

Reversible noncompetitive inhibition occurs when I binds to both E and ES. We will look at only the special case in which the dissociation constants of I for E and ES are the same. This is called noncompetitive inhibition. It is quite rare, as it would be difficult to imagine a large inhibitor that inhibits the turnover of a bound substrate and does not affect the binding of S to E. However, the covalent interaction of protons with both E and ES can lead to noncompetitive inhibition. In the more general case, the Kd's are different, and the inhibition is called mixed. Since inhibition occurs, we will hypothesize that ESI can not form the product. It is a dead-end complex with only one fate: to return to ES or EI. This is illustrated in the chemical equations and the molecular cartoon in Figure **7.2.4.11**.

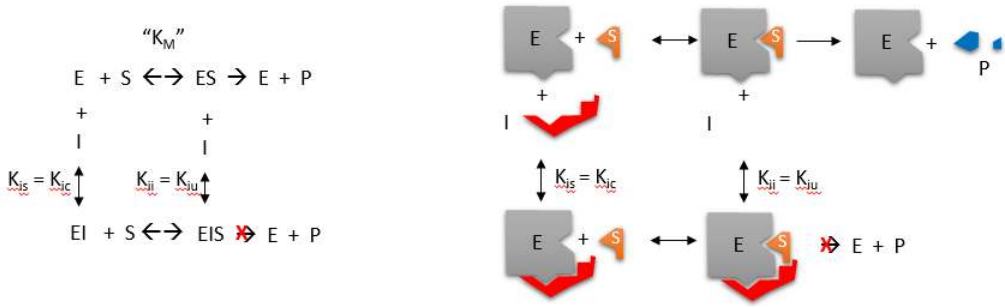


Figure **7.2.4.11**: Noncompetitive and Mixed Inhibition

Let us assume for ease of equation derivation that I binds reversibly to E with a dissociation constant of K_{is} (as we denoted for competitive inhibition) and to ES with a dissociation constant K_{ii} (as we noted for uncompetitive inhibition). Assume that for noncompetitive inhibition, $K_{is} = K_{ii}$. A look at the top mechanism shows that in the presence of I, as S increases to infinity, not all of E is converted to ES. There is a finite amount of ESI, even at infinite S. Now remember that $V_m = k_{cat}E_0$ if and only if all E is in the form ES. Under these conditions, the apparent V_m, V_{mapp}, is less than the real Vm without an inhibitor. In contrast, the apparent K_m, K_{mapp}, will not change since I binds to both E and ES with the same affinity and hence will not perturb that equilibrium, as deduced from LaChatelier's principle. The double reciprocal plot (Lineweaver-Burk plot) offers a great way to visualize the inhibition. In the presence of I, just V_m will decrease. Therefore, $-1/K_m$, the x-intercept will stay the same, and $1/V_m$ will get more positive. Therefore, the plots will consist of a series of lines intersecting on the x-axis, the hallmark of noncompetitive inhibition. You should be able to figure out how the plots would appear if K_{is} is different from K_{ii} (mixed inhibition).

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{Kis}\right)+S\left(1+\frac{I}{Kii}\right)} \tag{7.2.4.9}$$

An equation, shown in the diagram above, can be derived to show the effect of the noncompetitive inhibitor on the velocity of the reaction. Km is multiplied by $1+I/K_{is}$, and S by $1+I/K_{ii}$ in the denominator. We want to rearrange this equation to show how the inhibitor affects Km and Vm, not S. Rearranging the equation as shown above shows that $K_{mapp} = K_m(1+I/K_{is})/(1+I/K_{ii}) = K_m$ when $K_{is}=K_{ii}$, and $V_{mapp} = V_m/(1+I/K_{ii})$. This shows that the Km is unchanged, and Vm decreases as we predicted. The plot shows a series of lines intersecting on the x-axis as shown in Figure 7.2.4.12. Both the slope and the y-intercept are changed, which are reflected in the names of the two dissociation constants, K_{is} and K_{ii}. Note that if I is zero, $K_{mapp} = K_m$ and $V_{mapp} = V_m$. Sometimes, the K_{is} and K_{ii} inhibition dissociation constants are called $\mathbf{K_c}$ and $\mathbf{K_u}$ (competitive and uncompetitive inhibition dissociation constants.

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{KiS}\right)+S\left(1+\frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1+\frac{I}{Kii}}\right)S}{K_M+S}$$

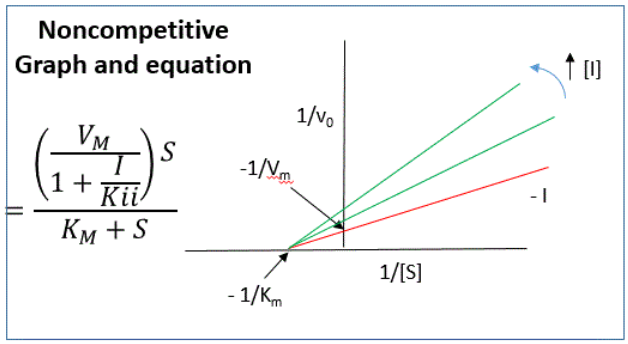
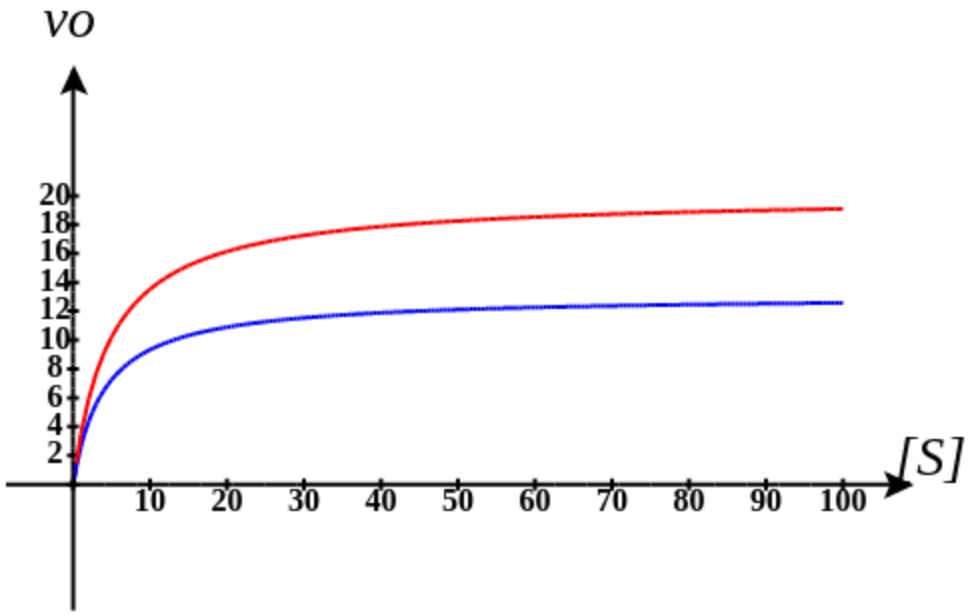


Figure 7.2.4.12: Noncompetitive Inhibition: Lineweaver-Burk plots

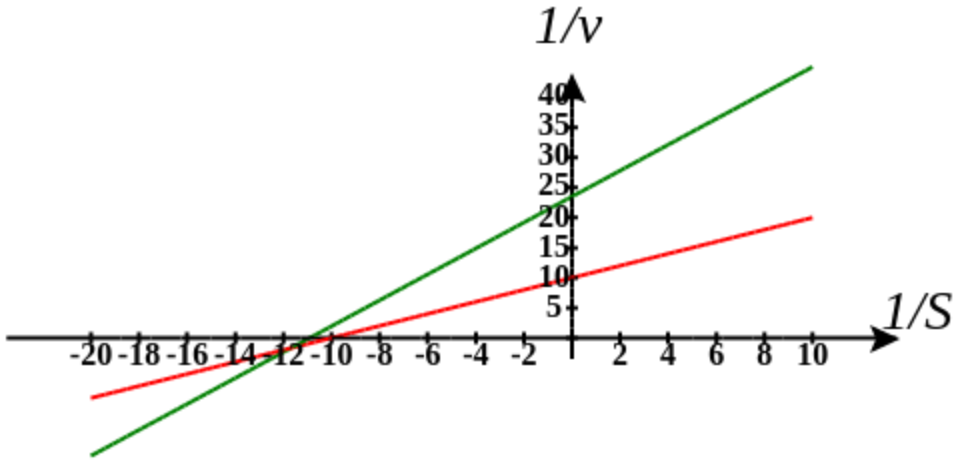
Move the sliders on this interactive graph below to show changes in K_{is} and K_{ii} affect the position on the graph where the lines intersect. Try to change their values to move the intersections of the graphs from the left top quadrant to the x-axis to the left bottom quadrant.

Here is an interactive graph showing v_0 vs [S] for mixed inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph



Here is an interactive graph showing mixed inhibition with Vm and Km both set to 10. Change the sliders for [I] and Kii and see the effect on the graph

Mixed Inhibition

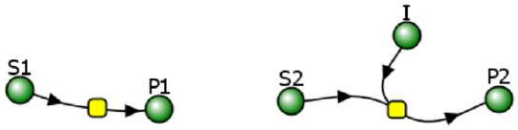


7.2.4.4.1: Progress Curves: Mixed/Noncompetitive Inhibition

Now let's compare the progress curves for an enzyme-catalyzed reaction in the absence and presence of a mixed inhibitor.

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



Note that the Vcell reaction diagram is the same as for competitive and uncompetive inhibition. It doesn't explicitly show that the mixed inhibitor binds to both free and substrate-bound enzymes. Those interactions are addressed in the mathematical equations for mixed inhibition.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate		$\frac{Vmax \cdot S2}{(Km \cdot (1.0 + \frac{I}{Kc}) + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Kc	user defined	☐	1.0
Ki	user defined	☐	1.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_MIXEDInhib_MM

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Mixed and noncompetitive inhibition (as shown by the mechanism above) differ from competitive and uncompetitive inhibition. The inhibitor binding is not simply a dead-end reaction in which the inhibitor can only dissociate in a single reverse step. In the above equilibrium, S can dissociate from ESI to form EI, so the system may not be at equilibrium. With dead-end steps, no flux of reactants occurs through the dead-end complex, so the equilibrium for the dead-end step is not perturbed.

Other mechanisms can commonly give mixed inhibition. For example, the product released in a ping pong mechanism (discussed in the next chapter) can give mixed inhibition. A ping pong reaction mechanism is shown and superficially explained in Figure 7.2.4.13.

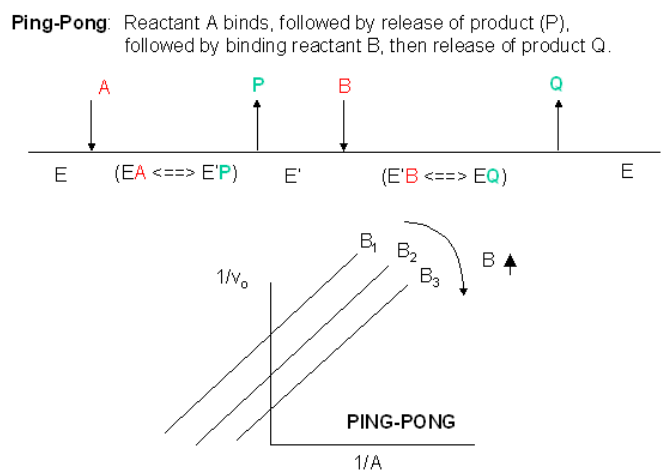
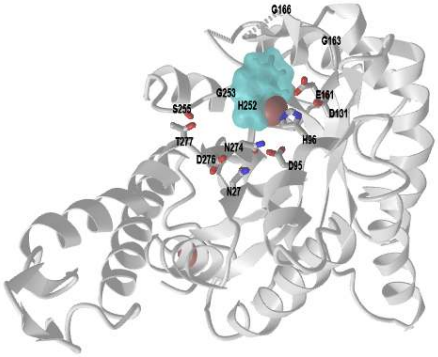


Figure 7.2.4.13: Ping pong reaction mechanism

If P, acting as a product inhibitor, can bind to two different forms of the enzyme (E' and also E), it will act as a mixed inhibitor.

Figure 7.2.4.14 shows an interactive iCn3D model of a noncompetitive inhibitor of M. tuberculosis's class IIa fructose 1,6-bisphosphate aldolase (4LV4).



NCBI iCn3D Figure 7.2.4.14: Noncompetitive inhibitor of M. tuberculosis's class IIa fructose 1,6-bisphosphate aldolase (4LV4). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wZDtkvtwqnrVy7

The non-competitive inhibitor, 8-hydroxyquinoline carboxylic acid (HCA), is shown with a cyan surface. Active site side chains of the Zn^{2+}- containing enzyme are shown as colored sticks and labeled.

The noncompetitive inhibitor does not enter into the pocket where the substrate (not shown in the model above) binds. The binding of the noncompetitive inhibitor causes a conformational change that moves the Zn-binding loop (Z-loop), which contains His 212, which coordinates the Zn^{2+} ion. Hence, the Zn-loop is part of the active site, and when it moves on binding of the inhibitor, the access to the empty pocket where the substrate binds is hindered. When the Zn^{2+} moves away from the active site, it can no longer engage in catalytic activity.

7.2.4.5: Enzyme Inhibition *in Vivo*

The pharmaceutical industry is devoted to finding drug molecules that affect biological processes. Typically, this means the development of small-molecule inhibitors for target proteins. Recent work has expanded to developing inhibitory RNA molecules that affect DNA transcription and mRNA translation. Using combinatorial synthetic techniques and computational modeling, developing small molecule inhibitors (especially competitive ones) that inhibit proteins in vitro using purified enzymes, substrates, and inhibitors in lab testing has gotten easier. Assuming that the inhibitor could pass through the membrane and accumulate to a sufficient concentration, would it have the same inhibitory properties in the cell as in the test tube? The answer turns out to be maybe. Remember that a cell is tightly packed with a multitude of other small molecules and macromolecules. In addition, the enzyme targeted for inhibition is most likely part of a pathway of enzymes that feed reactants into the enzyme and remove the product. Hence, the flux of substrate and product is controlled by the entire pathway and not just the single target enzyme. The target enzyme's product concentration is determined by kinetic parameters for the enzyme and available substrate concentration.

The conditions under which the enzymes are studied (in vitro) and operate (in vivo) are very different.

- In vitro (in the lab), the enzyme is held at a constant concentration while the substrate is varied (i.e., the substrate concentration is the independent variable). The velocity is determined by the substrate concentration. When inhibition is studied, the substrate is varied while the inhibitor is held constant at several different fixed concentrations.
- In vivo (in the cell), the velocity might be held at a relatively fixed level, with the substrate concentration determined by the velocity. The substrate can't build up at the enzyme to avoid a bottleneck in flux, so the enzyme processes it in a steady state fashion to produce the product as determined by the Michael-Menten equation.

What happens when an inhibitor is added *in vivo*? Let's assume that the enzyme is running at $v = V_m/2$. How might in vivo inhibition plots look at constant velocity (for example, $v=V_m/2$) when both I and S can vary, and in which S for an enzyme in the middle of a pathway is determined by v?

The equations and graph below show the ratio of S/K_m vs I/K_{ix} for inhibition at constant v, a condition encountered when an enzyme in a metabolic pathway is subject to flux controls imposed by the entire pathway. The x-axis reflects the ratio of inhibitor

concentration to its inhibition constant. Likewise, the y-axis reflects the relative amount of substrate compared to its K_m. The graph for *in vivo* competitive inhibition is linear, but it "blows up" for uncompetitive inhibition, as shown in Figure 7.2.4.15.

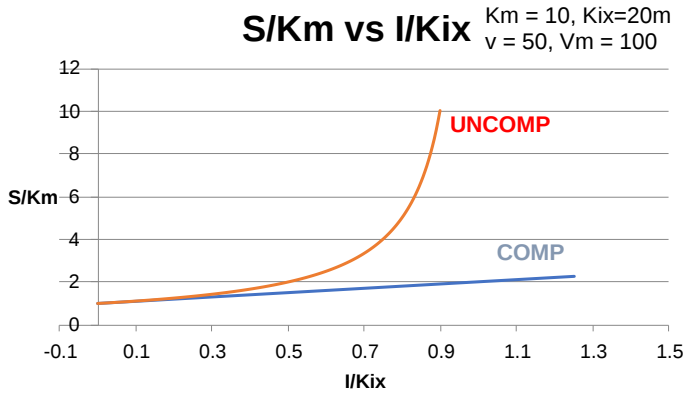


Figure 7.2.4.15: S/K_m vs $1/K_{ix}$ plots for competitive and uncompetitive inhibition

Here are the derivations used to produce the graphs in Figure 7.2.4.13.

? A derivation of the competitive and uncompetitive inhibition in vivo when v = constant

Here it is!

Derivation

Competitive Inhibition at constant velocity v:

Let's start with the equation of competitive inhibition.

$$\mathrm{v} = \frac{\mathrm{V_M S}}{\mathrm{K_M}\left(1 + \frac{\mathrm{I}}{\mathrm{K_{is}}}\right) + \mathrm{S}} \tag{7.2.4.10}$$

Let

$$\left(1 + \frac{\mathrm{I}}{\mathrm{K_{is}}}\right) = \mathrm{y} \tag{7.2.4.11}$$

Then,

$$\begin{gathered} v = \frac{V_M S}{K_M y + S} \\ v\left(K_M y + S\right) = V_M S \\ v K_M y + v S = V_M S \\ v K_M y = V_M S - v S = S\left(V_M - v\right) \\ K_M = \frac{S\left(V_M - v\right)}{v y} \end{gathered} \tag{7.2.4.12}$$

which gives:

$$\begin{gathered} \frac{S}{K_M} = \frac{v y}{\left(V_M - v\right)} = \frac{y}{\frac{V_M}{v} - 1} = \frac{1 + \frac{I}{K_{is}}}{\frac{V_M}{v} - 1} = \frac{1}{\frac{V_M}{v} - 1} + \frac{\frac{I}{K_{is}}}{\frac{V_M}{v} - 1} \\ \frac{S}{K_M} = \frac{1}{\frac{V_M}{v} - 1} + \left(\frac{1}{\frac{V_M}{v} - 1}\right) \frac{I}{K_{is}} \end{gathered} \tag{7.2.4.13}$$

Note from the last equation that the graph of S/K_M vs I/K_{is} is linear (at a fixed v), as shown in the above figure.

Uncompetitive Inhibition at constant velocity v:

Let's start with the equation of uncompetitive inhibition.

$$\mathrm{v} = \frac{\mathrm{V_M S}}{\mathrm{K_M} + \mathrm{S}\left(1 + \frac{\mathrm{I}}{\mathrm{K_{ii}}}\right)} \tag{7.2.4.14}$$

Let

$$\left(1 + \frac{1}{\mathrm{K_{ii}}}\right) = \mathrm{y} \tag{7.2.4.15}$$

then:

$$\begin{gathered} v = \frac{V_M S}{K_M + Sy} \\ v\left(K_M + Sy\right) = V_M S \\ vK_M = V_M S - vSy = S\left(V_M - vy\right) \end{gathered} \tag{7.2.4.16}$$

which gives:

$$\frac{S}{K_M} = \frac{v}{V_M - vy}\left(\frac{\frac{1}{v}}{\frac{1}{v}}\right) = \frac{1}{\frac{V_M}{v} - y} = \frac{1}{v - 1 - \frac{I}{K_{ii}}} \tag{7.2.4.17}$$

This graph is not a linear function of I/K_{ii} as it was for in vivo competitive inhibition.

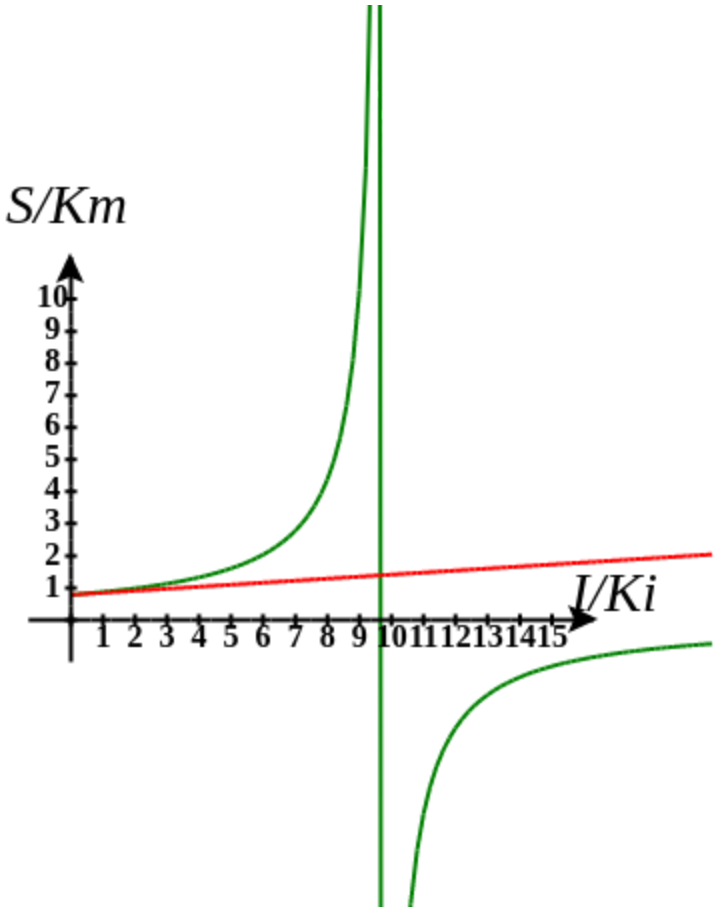
- for competitive inhibition, the graph of S/K_M is a linear function of I/K_{is}
- for uncompetitive inhibition, the graph of S/K_M is **NOT** a linear function of I/K_{ii} but rather "blows up" to infinity.

These graphs and associated equations are **dramatically** different from the very similar forms of inhibition equations and curves for in vitro inhibition at varying S and different fixed values of inhibitor. Consider the uncompetitive graph and equation. In the absence of an inhibitor, if S=Km, then Vm/v = 2, so the calculated value from the equation above of S/Km = 1. The y-intercept of the graph above is 1 for uncompetitive (and competitive) inhibition. If I is allowed to increase to a value of Kii (so I/Kii = 1), again at constant v=Vm/2, then the right-hand side goes to infinity.

In a linked series of reactions, if the middle reaction is inhibited, the substrate for that enzyme builds, whether the inhibition is competitive or uncompetitive. With competitive inhibition, the substrate concentration can be raised to meet the requirements of the enzyme. However, as the above figure shows, this can't happen for uncompetitive inhibition since, as more substrate accumulates, the reaction reaches a point where the steady state is lost.

Obviously, this limiting case can't be realistically reached, but it does suggest that uncompetitive inhibitors would be more effective *in vivo* in controlling a metabolic pathway than competitive inhibitors. Cornish-Bowden argues that purely uncompetitive inhibitors are rare because of the degree of inhibition they can hypothetically produce (1986). Likewise, he suggests that medicinal chemists should synthesize uncompetitive inhibitors if their goal is to maximally inhibit a metabolic pathway under the kind of flux control described above. Although it is more difficult to synthesize a purely uncompetitive inhibitor (as it can't be easily modeled after the structure of a natural ligand that binds to the active site and are competitive inhibitors), he notes that synthesizing mixed (and noncompetitive) inhibitors whose Kii values are of reasonable size compared to their Kis values, would be one approach.

Move the sliders on the interactive graph below to show how the graphs change. Disregard the graph in the right lower quadrant, as that location would require negative values of either S or K.



7.2.4.6: Inhibition by Temperature and pH Changes

From 0 to about 40-50 ^{o}C, enzyme activity usually increases, as do the rates of most reactions in the absence of catalysts. (Remember the general rule of thumb that reaction velocities double for each increase of 10 ^{o}C). At higher temperatures, the activity decreases dramatically as the enzyme denatures. These features are illustrated in Figure 7.2.4.16.

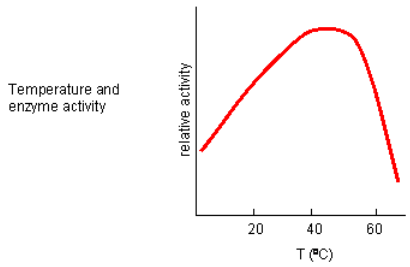


Figure 7.2.4.16: Temperature effects on enzyme activity

Likewise, pH has a marked effect on the velocity of enzyme-catalyzed reactions as illustrated in Figure 7.2.4.17.

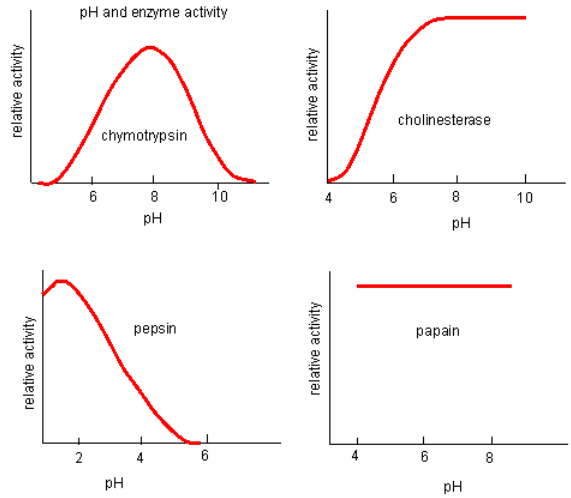


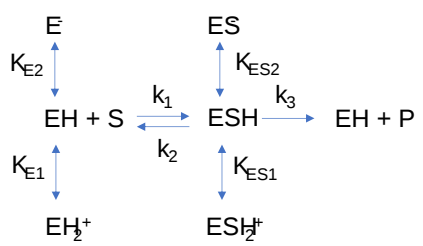
Figure 7.2.4.17: pH effects on enzyme activity

Think of all the things that pH changes might affect. It might

- affect E in ways to alter the binding of S to E, which would affect K_m
- affect E in ways to alter the actual catalysis of bound S, which would affect k_{cat}
- affect E by globally changing the conformation of the protein
- affect S by altering the protonation state of the substrate

The easiest assumption is that key side chains necessary for catalysis must be in the correct protonation state. Thus, a sidechain, with an apparent pKa of around 6, must be deprotonated for optimal activity of trypsin, which shows increases in activity with pH centered at pH 6. Which amino acid side chain would be a likely candidate?

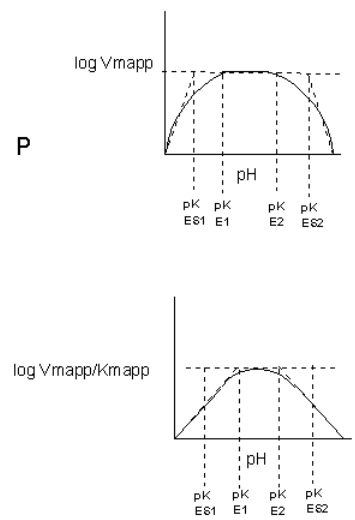
Table 7.2.4.1 below shows how pH effects on enzyme kinetics can be modeled at the chemical and mathematical level.



$$v_0 = \frac{V_{mapp}S}{K_m\ app} + S$$

$$V_{m\ app} = \frac{V_m}{1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}} \qquad (7.2.4.18)$$

$$K_{m\ app} = \frac{K_m\left(1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}\right)}{1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}}$$



7.2.4.7: Summary

Chapter Summary: Enzyme Inhibition – From Irreversible Covalent Modifications to Reversible Inhibition Mechanisms

This chapter explores the diverse strategies by which enzymes can be inhibited, emphasizing the molecular basis of both irreversible and reversible inhibition. Understanding these mechanisms is crucial for appreciating how enzyme activity is regulated

in biological systems and for the design of therapeutic inhibitors.

1. **Irreversible Covalent Inhibition**
   The chapter begins by discussing how enzyme activity can be permanently abolished through irreversible modifications. These modifications, whether induced by extremes of pH, high temperature, or covalent binding of small molecules (e.g., iodoacetamide modifying critical cysteine residues), result in structural changes that render the enzyme inactive. Techniques to identify essential catalytic side chains using protecting groups are also introduced.

2. **Reversible Inhibition Mechanisms**
   The discussion then shifts to reversible inhibition, which is subdivided into several mechanistic types:

   - **Competitive Inhibition:**
     Substrate and inhibitor compete for the same active site on the enzyme. Kinetic analysis shows that while VMV_M remains unchanged, the apparent K_M increases (i.e., $K_{M\ app}=K_M(1+I/K_{is})$. Graphical representations (e.g., Lineweaver-Burk plots) illustrate that the lines intersect on the y-axis.
   - **Uncompetitive Inhibition:**
     The inhibitor binds exclusively to the enzyme–substrate complex, reducing both V_M and K_M proportionally, which is reflected as parallel lines in Lineweaver-Burk plots.
   - **Noncompetitive and Mixed Inhibition:**
     When an inhibitor binds both to free enzyme and the ES complex (with equal or different affinities), the kinetic effects differ. In pure noncompetitive inhibition, V_M decreases without changing K_M, while mixed inhibition shows changes in both parameters. These distinctions are crucial for understanding enzyme regulation and designing inhibitors.

3. **Graphical and Mathematical Analysis of Inhibition**
   The chapter emphasizes the use of various kinetic plots (e.g., Michaelis-Menten, Lineweaver-Burk, and progress curves) to analyze how inhibitors affect enzyme kinetics. Interactive graphs and simulations illustrate how changing inhibitor concentrations and their dissociation constants (e.g., K_{is} and K_{ii}) influence parameters like K_M and V_M, as well as the overall reaction velocity.

4. **In Vivo versus In Vitro Considerations**
   The text contrasts inhibition studies conducted in vitro—where enzyme and substrate concentrations are controlled—with the complex, flux-driven environment in vivo. It discusses how metabolic pathways influence enzyme kinetics and how different inhibition mechanisms may have distinct effects under cellular conditions. For example, while competitive inhibitors can be overcome by increasing substrate concentration in vitro, uncompetitive inhibitors may exert more pronounced effects in vivo by preventing the accumulation of substrate.

5. **Environmental Effects on Enzyme Activity**
   Finally, the chapter reviews how external factors such as temperature and pH influence enzyme activity. Moderate increases in temperature typically enhance reaction rates, whereas excessive heat leads to irreversible denaturation. Similarly, pH changes can alter the protonation state of key catalytic residues, thereby affecting both substrate binding and catalytic turnover.

In summary, this chapter provides a comprehensive understanding of enzyme inhibition by detailing both irreversible covalent modifications and reversible inhibition mechanisms. It integrates mathematical derivations with graphical analyses and computational models to illustrate how inhibitors modulate enzyme kinetics. These insights are essential for the rational design of inhibitors in drug discovery and for a deeper understanding of metabolic regulation in living systems.

This page titled 7.2.4: Enzyme Inhibition is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.4: Enzyme Inhibition** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.2.5: Enzymatic Reaction Mechanisms

6.05: Enzymatic Reaction Mechanisms

This page titled 7.2.5: Enzymatic Reaction Mechanisms is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

7.2.6: Enzymes and Protein Regulation

Learning Goals (ChatGPT o1, 1/30/25)

- **Proteome Complexity Beyond the Genome**
  - Explain how a relatively modest number of genes can give rise to a proteome that is two to three orders of magnitude more complex through alternative splicing and post-translational modifications (PTMs).
- **Diversity and Function of Post-Translational Modifications (PTMs)**
  - Identify and describe common enzymatic and nonenzymatic PTMs—including phosphorylation, acetylation, glycosylation, ubiquitination, sumoylation, oxidation, and methylation—and their effects on protein structure, stability, activity, and localization.
  - Explain how PTMs create proteoforms and modulate protein–protein interactions and signal transduction pathways.
- **Detection and Analysis of PTMs**
  - Evaluate the various methods for detecting PTMs (e.g., antibody-based techniques, Western blotting, mass spectrometry) and discuss the advantages and limitations of each approach for both targeted and global PTM analysis.
- **Allosteric Regulation and Conformational Control**
  - Describe the principles of allosteric regulation, including how binding of small molecules, ions, or proteins at sites distinct from the active site can induce conformational changes that modulate enzymatic activity.
  - Differentiate between allosteric changes induced by small molecules, phosphorylation, protein–protein interactions, and alterations in disulfide bonds.
- **Zymogen Activation and Enzyme Maturation**
  - Explain the concept of zymogens (proenzymes) and the mechanisms by which they are activated, with examples such as the activation of pancreatic proteases and caspases during apoptosis.
- **Isoenzymes (Isozymes) and Differential Regulation**
  - Define isozymes and distinguish them from allozymes, highlighting their origins via gene duplication or alternative splicing and their roles in fine-tuning metabolic pathways.
  - Illustrate how isozymes, such as cyclooxygenases (COX-1 and COX-2), differ in kinetic properties, regulatory mechanisms, and tissue-specific expression, and discuss their relevance to therapeutic interventions.
- **Cyclooxygenases and Prostaglandin Biosynthesis**
  - Describe the structural and functional characteristics of COX-1 and COX-2, including their bifunctional enzyme activities (cyclooxygenase and peroxidase), and explain how their regulation influences the synthesis of prostaglandins and thromboxanes.
  - Discuss how selective and non-selective NSAIDs target COX isozymes and the physiological implications of disrupting the balance between prostacyclin and thromboxane in cardiovascular and gastrointestinal contexts.

By mastering these goals, students will be able to integrate molecular, structural, and regulatory aspects of protein control and appreciate the intricate layers that modulate protein function in response to cellular and environmental cues.

Many different mechanisms control protein activity within a cell. The primary sequence of a protein is a main determinant of protein folding and final conformation, as well as biochemical activity, stability, and half-life. However, at any given moment, the proteome, the full complement of proteins within a cell, is up to two or three orders of magnitude more complex than the encoding genomes would predict. This chapter will give an overview of the major mechanisms biological systems utilize to regulate protein functions after the protein has been synthesized. Note that these mechanisms seldom work in isolation. Multiple levels of protein control function at any given time and in response to many different environmental cues and signals. We will discuss many ways to

regulate protein activity. At the end, we will discuss regulation through the use of different **isozymes** of an enzyme, which are variants of an enzyme arising from differential splicing of a gene or from slightly different genes that arose from a common precursor gene. We will focus on cyclooxygenases, the target of so many medicinal drugs.

7.2.6.1: Post-translational modifications (PTMs)

The human genome contains approximately 20,000 genes. When analyzing the transcriptome, it becomes apparent that the genome becomes amplified by the wide array of splice variants that can occur when primary RNA transcripts are spliced to form mature messenger RNAs. There are about 200,000 expressed transcripts, of which about 146,00 encode proteins, with about 7.4 transcripts per gene. A summary of human gene transcripts is shown in Table 7.2.6.1 below.

Type	GTEx dataset	Expressed all genes	Expressed protein-coding genes
Transcripts per Gene	3.42	3.63	7.43
Gene counts	58,219	53,539	19,591
Transcript counts	199,324	194,146	145,571

Table 7.2.6.1: **Numbers of transcript and gene in the GTEx dataset.** Tung, KF., Pan, CY., Chen, CH. *et al.* Top-ranked expressed gene transcripts of human protein-coding genes investigated with GTEx dataset. *Sci Rep* **10**, 16245 (2020). https://doi.org/10.1038/s41598-020-73081-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The number of distinct primary structures of proteins (**proteoforms**) from 20,000 human genes is amplified again through posttranslational modifications (PTMs) of proteins, which produce up to a million different variants. PTMs are present in both eukaryotes and prokaryotes. Still, they are more common in eukaryotic cells, in which about 5% of the genome is dedicated to enzymes that carry out posttranslational modifications of proteins. We discussed the chemical modification of specific amino acids in Chapter 3.1: Amino Acids and Peptides.

Protein PTM results from the enzymatic or nonenzymatic attachment of specific chemical groups to amino acid side chains. Such modifications occur either following protein translation or concomitant with translation. PTM influences both protein structure and physiological and cellular functions. Examples of enzymatic PTMs include phosphorylation, glycosylation, acetylation, methylation, sumoylation, palmitoylation, biotinylation, ubiquitinylation, nitration, chlorination, and oxidation/reduction. Nonenzymatic PTMs include glycation, nitrosylation, oxidation/reduction, acetylation, and succinylation. Some rare and unconventional PTMs, such as glypiation, neddylation, siderophorylation, AMPylation, and cholesteroylation, are also known to influence protein structure and function. Note that many of these modifications are not made in isolation. It is common for proteins to have several different types of modifications, and these modifications can differ depending on the tissue type and environmental circumstances present. The major PTMs in eukaryotes, their target amino acid residue(s), and the types of enzyme(s) or protein(s) involved are shown in Table 7.2.6.2 below.

Posttranslational modification	Target amino acid residue(s)	Enzyme(s) or proteins involved
Phosphorylation	Tyrosine, serine, threonine	Kinases, phosphatases
Glycosylation N-linked	Asparagine	Glycosyltransferases, deglycosylases
Glycosylation O-linked	Serine/threonine	Glycosyltransferases, deglycosylases
Acetylation	Lysine	Acetyltransferases, deacetylases
Methylation	Lysine, arginine	Methyltransferases, demethylases
Ubiquitination	Lysine	Ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, ubiquitin ligases, deubiquitinases
Sumoylation	Lysine	Ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, ubiquitin ligases, deubiquitinases
Myristoylation	Glycine	N-Myristoyltransferases
Prenylation	Cysteine	Farnesyltransferases, geranyl geranyltransferases
Palmitoylation	Cysteine	DHHC protein acyltransferases, acyl-protein thioesterases
Sulfation	Tyrosine	Sulfatases, desulfatases
S-Nitrosylation	Cysteine, methionine	
Glycation	Lysine	
Nitration	Tyrosine	Denitrases
Chlorination	Tyrosine	Myeloperoxidases
Oxidation/reduction	Cysteine	Peroxidases, oxidases, glutathione, thioredoxin
Carbonylation	Lysine, proline, arginine, threonine	

Table 7.2.6.2: **Common Protein Post-Translational Modifications, Their Target Amino Acid Residues, and the Enzyme(s) or Proteins Involved.** *Santos, A.L, and Lindner, A.B. (2017) Oxidatve Medicine and Longevity, Article ID: 5716409*

PTMs can exert their effect through many different structural changes, including opening and closing the active site, changing the conformation and electrostatic properties of binding sites, changing the chain's flexibility, increasing or decreasing intrinsically disordered regions, altering protein-protein interactions, etc. These effects can also arise from the binding of small molecules.

7.2.6.1.1: Methods to Detect Protein Posttranslational Modifications

Specific amino acid residues are subjected to PTMs depending on the chemistry of the reaction and the sequence specificity of the enzyme involved. Initially, PTMs were detected by various analytical methods, such as radiolabeling of the proteins, thin-layer chromatography, column chromatography, and/or polyacrylamide gel electrophoresis. Other methods have since been developed, such as protein sequencing by Edman degradation and Western blotting using protein-specific antibodies. Antibody-based detection methods and mass spectrometry-based proteomic analysis are the predominant methods used to detect and analyze PTMs. However, mass spectrometric methods are the only available tools for global or large-scale PTM analysis.

Antibody-based methods mostly rely on the availability of antibodies that can specifically recognize a modified amino acid residue within a protein or peptide. Such antibodies can be polyclonal or monoclonal and are developed against either the modified peptide/protein or against the modified amino acid. Moreover, antibody-based detection and quantification of PTMs on protein/peptide samples can be performed by chemiluminescence-based Western blotting and absorbance/fluorescence-based ELISA. However, the detection of PTMs depends entirely on the recognition site of the antibody used. If the antibody detects only the modified amino acid, additional analysis—for instance, protein/peptide isolation and sequencing—should be performed to detect the sequence context of the modification. However, suppose the antibody detects the PTM within a specific sequence

context. In that case, the presence of PTM at other sites will remain undetected (ie, the antibody will be specific for only that single modification).

Mass spectrometric detection of specific PTMs is based on mass changes. Depending on the type of modification, a specific change in the mass of the modified amino acid or peptide occurs. Subsequently, the change in mass is detected by the mass spectrometer to identify the presence of a PTM in a peptide sample. Using tandem mass spectrometric methods, identification of the specific site of PTM can be achieved by subsequent fragmentation and sequencing of the relevant peptide. Yet, technical challenges hamper MS-based investigation of biologically important PTMs, such as ADP-ribosylation, one of the key signaling molecules that regulate DNA repair, a critical process in maintaining genome stability that is compromised in cancer and aging.

Data increasingly implicate PTMs during aging and/or under pathological conditions, and also in the cell's normal functioning. PTMs are increasingly studied for their role in health and disease. For example, the precise and accurate measurement of distinct PTM-containing moieties offers potential biomarker utility to aid early diagnosis, prognosis, monitoring response to therapy, and decisions regarding inclusion in clinical trials as new medicines are developed. However, technical difficulties limit these studies, leaving many unanswered questions. Identifying unknown/unexpected PTMs by proteomic data reanalysis is an emerging subfield of proteomics recently boosted by the increased availability of raw data shared in public repositories. Notably, a sampling of the proteome in a given organism or cell provides only a snapshot of a highly dynamic process, confounding the analytical problem and ultimately arguing for time-resolved inventories. Thus, while many tools are currently available for studying PTMs, new methods are needed to advance the study of these modifications further.

7.2.6.1.2: Examples of PTMs

Protein PTMs involve the covalent addition of some chemical groups by enzymatic catalysis. Typically, an electrophilic fragment of a co-substrate is added to an electron-rich protein side chain, which acts as a nucleophile in the transfer. Common covalent protein PTMs include phosphorylation, acylation, methylation, sumoylation, ubiquitination, glycosylation, lipidation, oxidation, and disulfide bond formation (internal within a single protein or linking two protein/peptide chains together). Examples of common PTMs are provided in Figure \(\PageIndex{1}\ below.

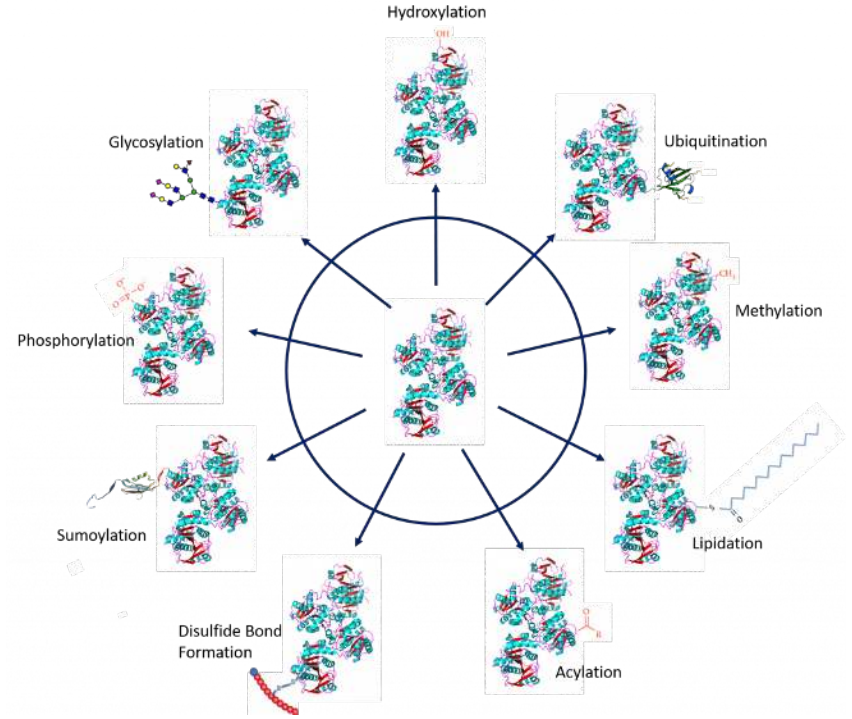


Figure **7.2.6.1**: **Major types of covalent additions to protein side chains.** Common modifications include oxidation, ubiquitination, methylation, lipidation, acylation, disulfide bond formation, sumoylation, phosphorylation, and glycosylation. *Figure modified from: Santos, A.L, and Lindner, A.B. (2017) Oxidative Medicine and Longevity, Article ID: 5716409, Goettig, P.*

(2016) Int. J. Mol. Sci. 17(12), 1969, Rogerdodd, Jakob Suckale, Akinlade, A., et al (2014) Int. Archives of Med 7(50):28, WilsonNR, and Sivart13

Protein Phosphorylation

One of the most common posttranslational modifications, protein phosphorylation, is the reversible addition of a phosphoryl group from adenosine triphosphate (ATP) to amino acid side chains such as serine, threonine, and tyrosine residues, as shown in Figure 7.2.6.2. This modification causes conformational changes that either (1) affect the catalytic activity to activate or inactivate the protein and/or cause the tendency of a protein to misfold and aggregate or (2) recruit other proteins to bind; both responses result in altered protein function and cell signaling. Phosphorylated proteins have critical and well-known functions in diverse cellular processes across eukaryotes, but phosphorylation also occurs in prokaryotic cells. In humans, about one-third of proteins are estimated to be substrates for phosphorylation. Indeed, phosphorylated proteins are now identified and characterized by high-throughput phosphoproteomics studies. The reversibility of protein phosphorylation is attributed to the actions of ***kinases*** and ***phosphatases***, which phosphorylate and dephosphorylate substrates, respectively. The temporal and spatial balance of kinase and phosphatase concentrations within a cell mediates the size of its phosphoproteome.

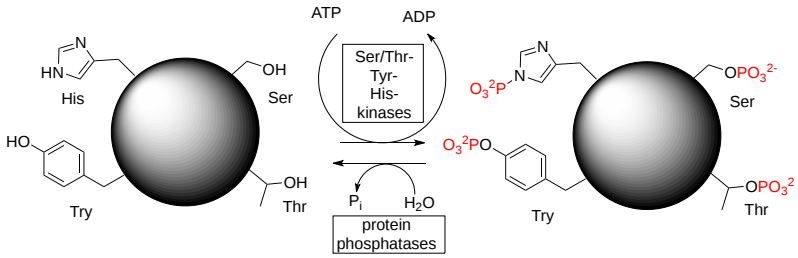


Figure 7.2.6.2: Amino acid phosphorylation in proteins

Protein Acylation

The simplest form of acylation is protein N-acetylation. This occurs at the amino terminus amine and the ε-amino group of the lysine side chains through the action of acetylases. The acetylation of lysine side chains can be reversed through deacetylases (similar to the combined actions of protein kinases and phosphatases). Figure 7.2.6.3 shows the general reactions of protein N-acetylation. Interestingly, 80–90% of eukaryotic proteins are acetylated, yet the underlying biological significance remains unclear.

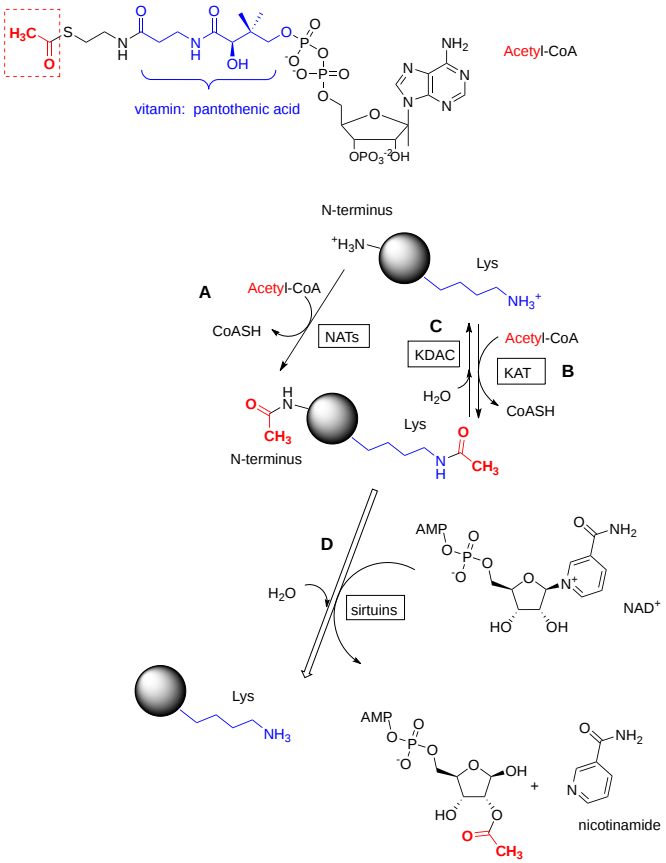


Figure 7.2.6.3: Protein N-acetylation and deacetylation reactions

The acetyl donor is acetyl-CoA. The N-terminus (reaction A) is acetylated by **N-terminal acetyltransferases (NATs**). The ε-amino group of lysine (reaction B) is acetylated by **lysine acetyltransferases (KATs).** The latter modification is reversed through two different enzymes. One is a simple hydrolysis (reaction C) catalyzed by lysine deacetylases (KDACs). The other (reaction D) requires NAD^+ and is catalyzed by enzymes called **sirtuins** (silent information regulators). Since sirtuins use NAD^+, one of the main oxidizing agents in the biological world, they link the protein **acetylome** with metabolic status and promote metabolic health.

Histones, positively charged proteins found in eukaryotic cell nuclei, pack and order the DNA into structural units called nucleosomes, as shown in Figure 7.2.6.4.

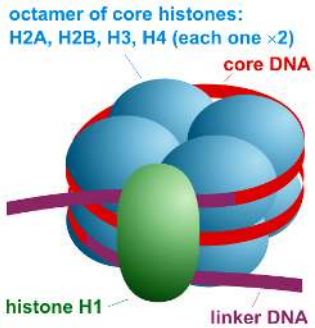


Figure 7.2.6.4: Nucleosome organization. https://commons.wikimedia.org/wiki/F...ganization.png. Creative Commons Attribution-Share Alike 3.0 Unported license.

The histones have multiple lysine and arginine residues that interact with the negatively charged phosphate groups of the DNA backbone. The nucleosome core comprises two H2A-H2B dimers and an H3-H4 tetramer, forming two nearly symmetrical halves by tertiary structure (C2 symmetry; one macromolecule is the mirror image of the other). The 4 'core' histones (H2A, H2B, H3, and H4) are relatively similar in structure and highly conserved through evolution, all featuring a 'helix-turn-helix' motif (a DNA-

binding protein motif that recognizes specific DNA sequence). They also share the feature of long 'tails' on one end of the amino acid structure, which is the location of post-translational modification, specifically N-acetylation.

Histone acetylation typically results in transcriptional activation; deacetylation typically results in transcriptional suppression. Acetylation occurs via ***histone acetyltransferases (HATs)*** and is reversible via the action of ***histone deacetylases (HDACs)***. One group of histone deacetylases is the **sirtuins** (silent information regulator), which maintain gene silencing via hypoacetylation. Sirtuins have been reported to aid in maintaining genomic stability. Figure 7.2.6.5 shows the effect of acetylation on histone:DNA packing.

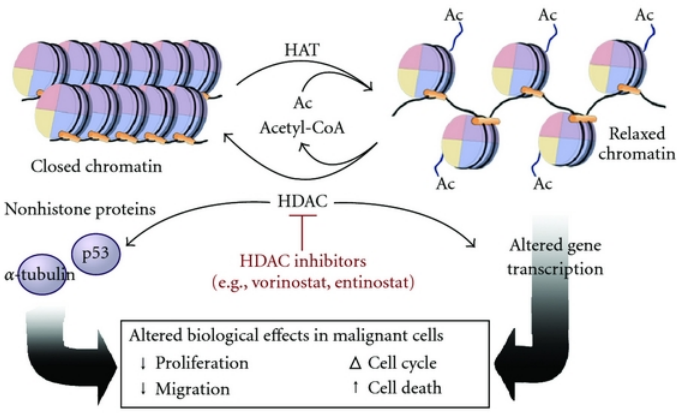


Figure 7.2.6.5: **Histone octamers form the backbone for chromosomal structure.** Acetylation and methylation affect histone-DNA interactions. DNA methylation increases histone-DNA affinity and blocks transcriptional activation. Acetylation of the histone tails disrupts histone-DNA interactions and facilitates gene expression. https://commons.wikimedia.org/wiki/F...cetylation.jpg. Creative Commons Attribution 3.0 Unported

Although first described in histones, acetylation is also observed in cytoplasmic proteins. Acetylated proteins can also be modulated by cross-talk with other posttranslational modifications, including phosphorylation, ubiquitination, and methylation. Therefore, acetylation may contribute to cell biology beyond transcriptional regulation.

Protein Glycosylation

Protein glycosylation involves adding diverse sugar moieties to the protein core. Glycosylation significantly affects protein folding, conformation, distribution, stability, and activity. Glycosylated proteins can have additions of simple monosaccharides. For example, many nuclear transcription factors are modified in this way. Alternatively, some proteins are modified with highly complex branched polysaccharides, such as those on cell surface protein receptors.

More than half of all mammalian proteins are believed to be glycosylated. While proteins exhibit improved stability and trafficking after glycosylation *in vivo*, glycan structures can alter protein functions or activities. Glycan structures are often modified by glycan-processing enzymes working within a cell at any given time. However, the structures are sometimes protein-specific, depending on protein trafficking properties and interactions with other cellular factors.

There are three types of protein glycosylation in higher eukaryotes: N-linked, O-linked, and C-linked. These types reflect their glycosidic linkages to amino acid side chains. In N-linked glycosylation, β-N-acetylglucosamine (GlcNAc) is attached through an amide linkage to the side chain of Asn within a consensus sequence of AsnXaaSer/Thr (Figure 8.8). N-linked glycans have multiple functions. While they act as ligands for glycan-binding proteins in cell-cell communication, they also can regulate glycoprotein aggregation in the plasma membrane and affect the half-life of antibodies, cytokines, and hormones in serum.

O-linked glycosylation in higher eukaryotes occurs through several different mechanisms. The most abundant type of O-linked glycosylation is the ***mucin-type***, which involves the attachment of an α-N-acetylgalactosamine (GalNAc) to the hydroxyl group of Ser/Thr side chains. ***Mucins*** are a family of high molecular weight, heavily glycosylated proteins (glycoconjugates) produced by epithelial tissues in most animals. A key characteristic of mucin proteins is their ability to form gels; therefore, they are a key component in most gel-like secretions, serving functions from lubrication to cell signaling to forming chemical barriers. Aberrant expression of mucin-type O-linked glycans occurs in cancer cells and may provide targets for anticancer vaccines.

O-linked glycosylation, occurring with the addition of α-O-mannose, is the only form of O-linked glycosylation in yeast. Still, it also occurs in the brains of higher eukaryotes. Higher eukaryotes also have an α-O-fucose modification of Ser/Thr residues. This type of glycosylation modulates signaling pathways during eukaryotic development. Another modification, β-O-galactosylation, may contribute to rheumatoid arthritis.

Finally, C-linked glycosylation involves the addition of α-mannose (Man) to the 2-position of the indole side chain of tryptophan residues. While first identified on ribonuclease 2, it also occurs on other proteins, including some mucin proteins, thrombospondin, and the Ebola virus soluble glycoprotein. Figure 7.2.6.6 shows examples of glycan adducts.

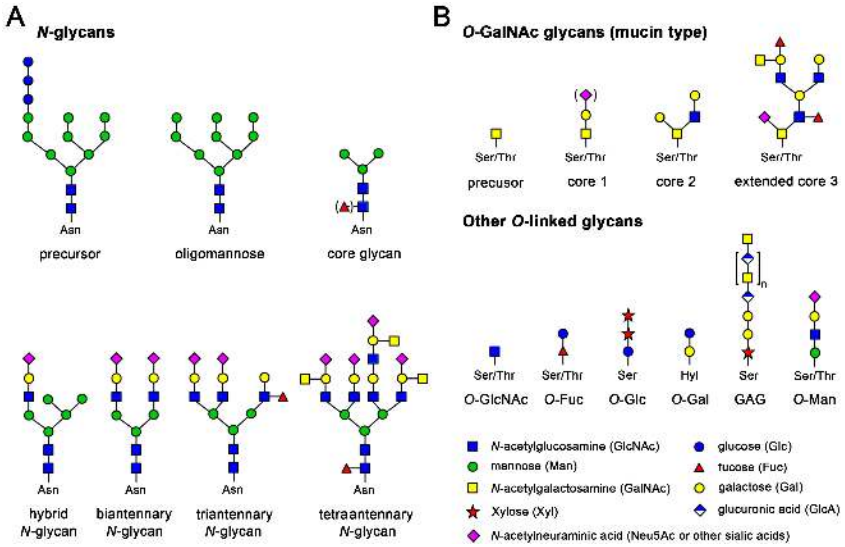


Figure 7.2.6.6: Glycan adducts to proteins Goettig, P. (2016) Int. J. Mol. Sci. 17(12), 1969

We will discuss glycoproteins in great detail in the next chapter.

Protein Ubiquitination and Sumoylation

Ubiquitination (or ubiquitylation) is the addition of an eight kDa polypeptide, called ubiquitin, to lysine residues of target proteins via the C-terminal glycine residues of ubiquitin. Adding one ubiquitin can lead to adding more to form ubiquitin chains on the target protein. The addition of one ubiquitin can modify the activity of the target protein. This can regulate the activity of a protein, its location within the cell, and its interaction with other proteins. Adding a single ubiquitin to lysine 120 (K120) in humans and to K123 in yeast alters gene transcription through the remodeling of chromatin. However, if multiple ubiquitins are added to the target protein, forming a polyubiquitinated protein, the protein is tagged for degradation by the 26S proteasome. Figure 7.2.6.7 shows different patterns of ubiquitylation for proteins.

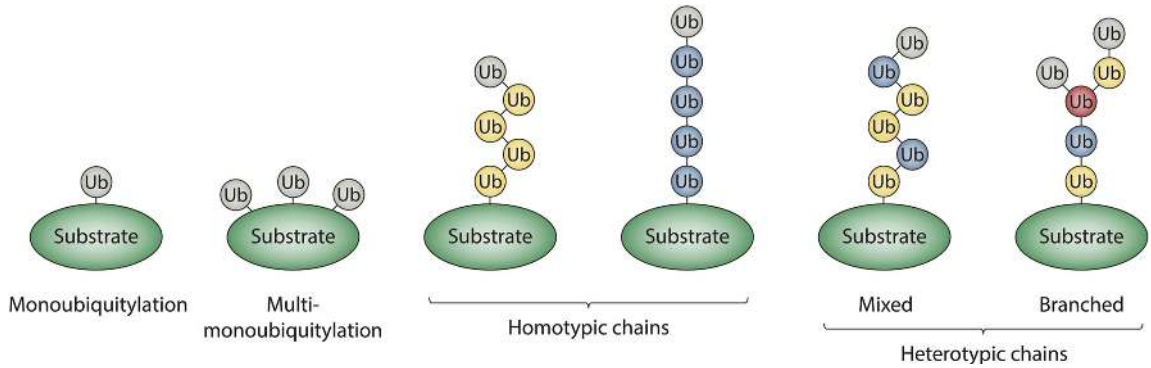


Figure 7.2.6.7: Classification of ubiquitin modifications. Ubiquitins modified on one acceptor site are colored in blue or yellow; the branch point ubiquitin is colored in red; unmodified or "terminal" ubiquitins are colored gray. French et al. Cell Discovery (2021) 7:6 Cell Discovery. https://doi.org/10.1038/s41421-020-00237-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

An elaborate system of enzymes attaches a ubiquitin through its C-terminal Gly 76 to an internal lysine (or cysteine) in a target protein. Then it adds more ubiquitins to form the structures shown above. Figure 7.2.6.8 shows the first step in the process, which

requires three enzymes: E1, E2, and E3.

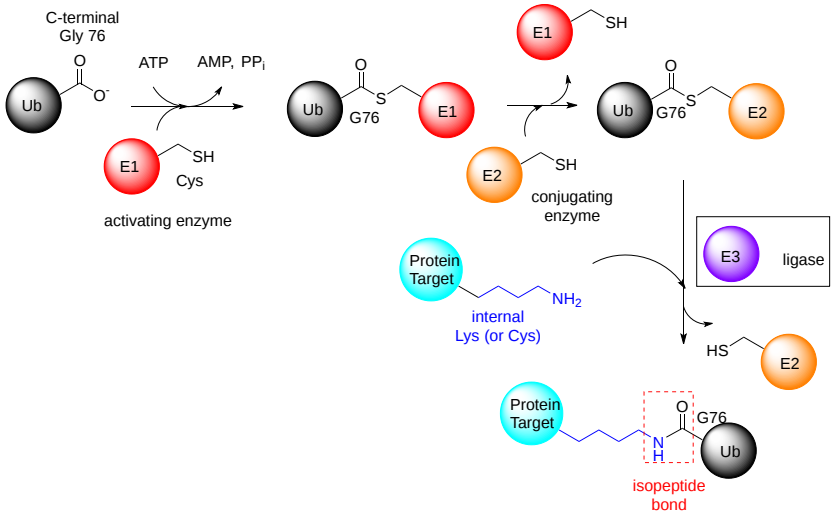


Figure 7.2.6.8: The first step in the process, which requires three enzymes, E1 (Ub-activating), E2 (Ub-conjugating), and E3 (Ub-ligating) enzymes

Once the first (or proximal) ubiquitin is added, more can be added to form chains. The steps involved in adding a second ubiquitin to the proximal one are shown in Figure 7.2.6.9.

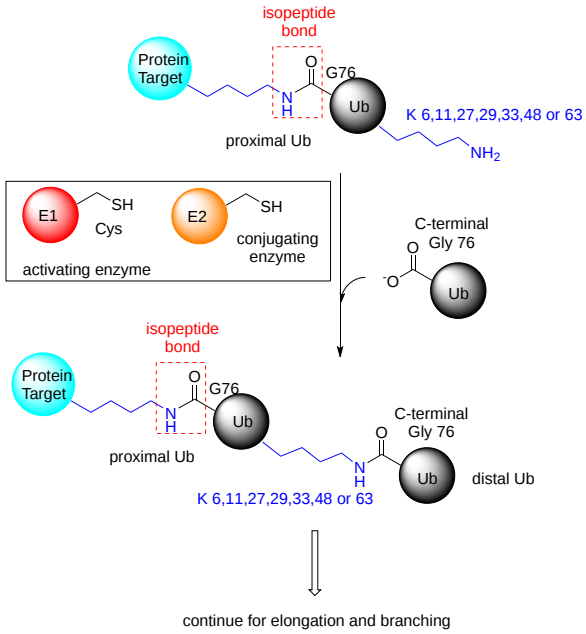


Figure 7.2.6.19: Addition of additional ubiquitin to form chains

Similarly, protein sumoylation is a reversible posttranslational modification whereby a **s**mall **u**biquitin-like **mo**difier (**SUMO**) protein is covalently attached to a target. Like ubiquitin, SUMO is linked to a lysine side chain of target proteins and is removed by SUMO-specific isopeptidases. Sumoylation controls many aspects of nuclear function. Protein sumoylation is involved in many extranuclear neuronal processes and potentially in many neuropathological conditions.

Figure 7.2.6.10 shows ubiquitin (1UBQ, magenta) and SUMO-1 (1A5R, cyan) superimposed. Note the similarity in their structures even though they share only 18% sequence identity. Key amino acids are shown in spacefill.

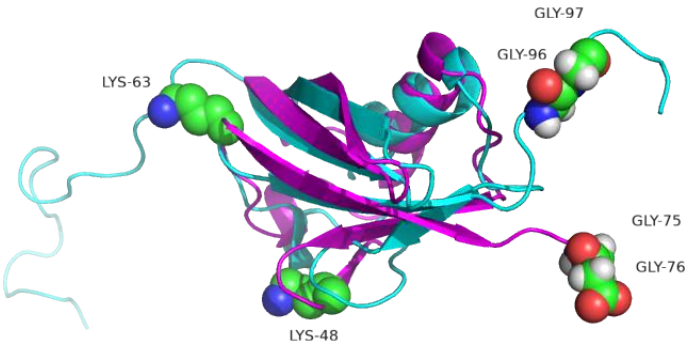


Figure 7.2.6.10: Ubiquitin (1UBQ) and SUMO-1 (1A5R) superimposed

SUMO-1 doesn't target proteins for degradation. One of its roles is in the transport of proteins between the nucleus and cytoplasm, and it is involved in the modulation of protein-protein interactions. Both ubiquitin and SUMO-1 have a C-terminal glycine involved in isopeptide bond formation between ubiquitin or SUMO and the target protein. A significant difference between ubiquitin and SUMO-1 is along the disordered N-terminal end in SUMO. Lysines 48 and 63, found in ubiquitin, are also replaced by other amino acids. This may account for the observation that SUMO-1 does not form polymers.

The interplay between ubiquitylation and SUMOylation is complex. If the ubiquitination of proteins is inhibited, SUMOylation of newly synthesized proteins increases. These can end up in phase-separated condensates called Promyelocytic Leukemia Nuclear Bodies (PML nuclear bodies), 0.1–1 μm condensates without a membrane found in most mammalian cell nuclei. These are somehow involved in gene expression in the nucleus.

Protein Oxidation

The reaction of proteins with a variety of free radicals and reactive oxygen species (ROS) leads to oxidative protein modifications such as the formation of protein hydroperoxides, hydroxylation of aromatic groups and aliphatic amino acid side chains, oxidation of sulfhydryl groups, oxidation of methionine residues, conversion of some amino acid residues into carbonyl groups, cleavage of the polypeptide chain, and formation of cross-linking bonds. Aromatic and sulfur-containing amino acid residues are particularly susceptible to oxidative modification.

Unless repaired or removed from cells, oxidized proteins are often toxic and can impair cellular viability since oxidatively modified proteins can form large aggregates. Oxidatively damaged proteins undergo selective proteolysis, primarily by the 26S proteasome, in a ubiquitin- and ATP-independent way. Upon extensive protein oxidation, these aggregates can become progressively resistant to proteolytic digestion, bind the 20S proteasome, and irreversibly inhibit its activity.

Protein carbonylation is defined as an irreversible posttranslational modification (PTM) whereby a reactive carbonyl moiety, such as an aldehyde, ketone, or lactam, is introduced into a protein. The first identified source of protein-bound carbonyls was metal-catalyzed oxidation (MCO). MCO results from the Fenton reaction when transition metal ions are reduced in the presence of hydrogen peroxide, generating highly reactive hydroxyl radicals. These hydroxyl radicals can oxidize amino acid side chains or cleave the protein backbone, leading to numerous modifications, including reactive carbonyls. For example, the oxidation of proline and arginine results in the production of glutamic semialdehyde, while lysine is oxidized to aminoadipic semialdehyde and threonine to 2-amino-3-ketobutyric acid. Direct oxidation of other amino acid residues can also lead to protein-bound carbonyls. Tryptophan oxidation by ROS produces at least seven oxidation products. Among them are kynurenine and N-formyl kynurenine and their hydroxylated analogs, which contain aldehyde or keto groups formed by oxidative cleavage of the indole ring.

Another important source of protein-bound carbonyls is reactive lipid peroxidation products, which are produced during the oxidation of polyunsaturated fatty acids. Protein carbonylation can also occur via glycoxidation. Reactive α-carbonyls formed during glycoxidation, such as glyoxal, methylglyoxal, and 3-deoxyglucosone, can then modify the basic residues Lys and Arg to generate, for example, pyrralines and imidazolones. Glycation (i.e., the reaction of reducing sugars such as glucose or fructose with the side chains of lysine and arginine residues) forms Amadori and/or Hynes products. ROS can further change these glycated residues into advanced glycation end products (AGE) carrying carbonylated moieties.

Some oxidative modifications are made enzymatically and have key regulatory or structural functions within the modified proteins. For example, proline can be converted to hydroxyproline and lysine to hydroxylysine, as shown in Figure 7.2.6.11. 4-Hydroxyproline makes up about 13.5% of the residues within the mammalian collagen family of proteins. Recall that collagen is the main protein of connective tissue and represents about one-fourth of the total protein content in many animals. Hydroxyproline contributes to the stability of the triple helix and also aids in cross-linking collagen fibers to form larger macromolecular complexes.

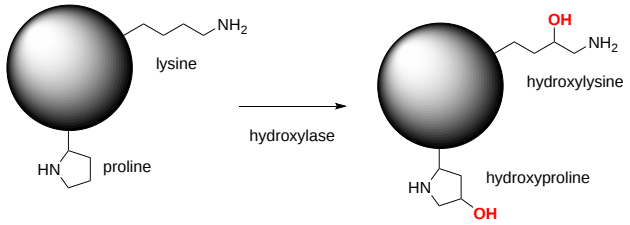


Figure 7.2.6.11: Hydroxlation of lysine and proline. Xu, Y., et. al. (2014) Int J Mol Sci 15(5):7594-7610.

Protein Methylation

Alkyl substituents are also a common posttranslational modification. Introducing such alkyl groups results in the alteration of the hydrophobicity of the modified protein. The most common type of protein alkylation is protein methylation, which is mediated by methyltransferase enzymes. One-carbon methyl groups are added to nitrogen or oxygen (N- and O-methylation, resp.) on amino acid side chains, increasing protein hydrophobicity or neutralizing a negative charge when bound to carboxylic acids. While N-methylation is typically irreversible, O-methylation is potentially reversible. Methylation occurs so often that its primary methyl donor, S-adenosyl methionine (SAM), is one of the most-used enzymatic substrates after ATP. The methylation of a histone protein is shown in Figure 7.2.6.12.

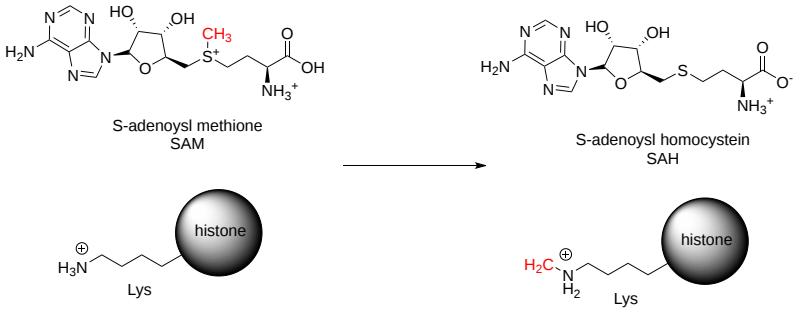


Figure 7.2.6.12: Methylation of a histone protein by SAM

This modification (not unlike phosphorylation reactions) plays a role in regulating protein-protein interactions. For instance, the arginine methylation of proteins can either inhibit or promote protein-protein interactions depending on the type of methylation. Protein methylation is also a common modification found in ***histone proteins.*** The transfer of methyl groups from S-adenosyl methionine to histones is catalyzed by enzymes known as ***histone methyltransferases***. The N-terminal tails of histones H3 and H4 receive methyl groups on specific lysines. Methylation then determines if gene transcription is activated or repressed, thus leading to different biological outcomes.

Histone methylation was traditionally thought to be irreversible. However, ***histone demethylases*** demonstrate the reversibility of this PTM. The simultaneous removal of one histone methylation and the addition of another can enable transcriptional fine-tuning.

In addition to methylation, histone acetylation, deacetylation, and mono-ubiquitination are essential parts of gene regulation, as shown below in Figure \(\PageIndex{13\). Acetylation removes the positive charge on the histones, thereby decreasing the interaction of the N termini of histones with the negatively charged phosphate groups of DNA. Consequently, condensed chromatin is transformed into a more relaxed structure associated with greater gene transcription levels.

Nonhistone proteins also exhibit methylation as a common PTM, which is used to regulate signal transduction pathways. Furthermore, methylation works in concert with other types of PTMs to exert influence not only on chromatin remodeling but also on gene transcription, protein synthesis, and DNA repair.

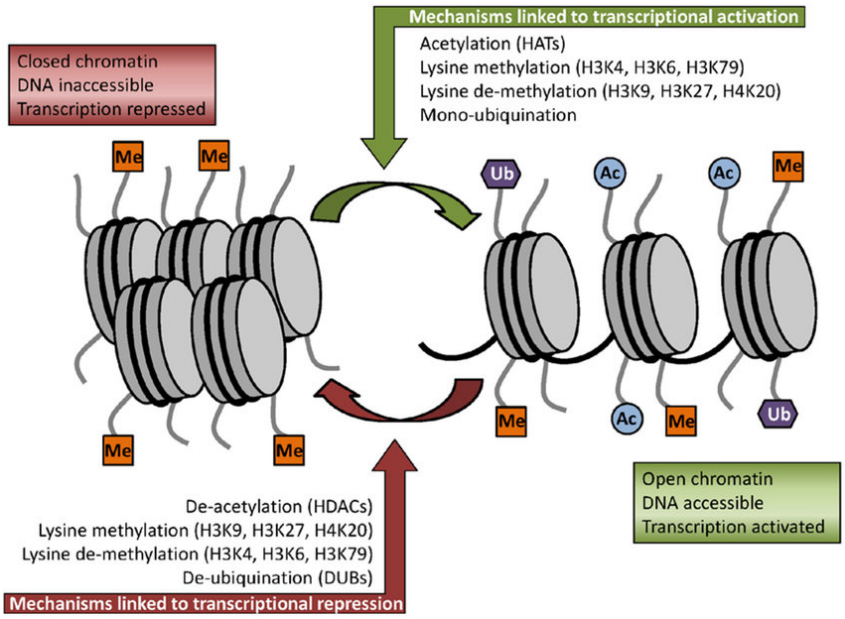


Figure \(\PageIndex{13\): Some of the key histone modifications influencing gene expression (Me: methylation, Ub: ubiquination, Ac: acetylation). Figure from: Ullah, M.F. (2015) Medicines 2:141-156.

7.2.6.2: Allosteric Regulation

Allosteric regulation fine-tunes most biological processes, including signal transduction, enzyme activity, metabolism, and transport. ***Allostery***, an intrinsic property of a protein, is referred to as the regulation of activity at one site (also known as an ***orthosteric site***) in a protein by a topographically and spatially distant site; the latter is designated as an ***allosteric site. Allosteric regulation*** occurs through binding a modulator (e.g., small molecule or protein) at an allosteric site to engender a conformational change that affects function at the orthosteric site. This effect may cause the re-distribution of the conformational ensemble by either stabilizing an active conformation (***allosteric activation***) or destabilizing an inactive conformation (***allosteric inhibition***) in response to allosteric perturbations (Figure 8.13). Traditionally, the repertoire of allostery was primarily confined to determining the allosteric effects or mechanisms in individual multi-subunit or monomer proteins by conformational transitions. Recently, increasing evidence has indicated that allosteric signals can propagate across multiple proteins to sculpt allosteric networks.

Allosteric regulation is particularly important in the cell's ability to adjust enzyme activity based on the surrounding environmental conditions. Feedback control loops, such as feedback inhibition from downstream products or feedforward from upstream substrates, are common allosteric regulatory mechanisms found in nature. Another example of allostery includes oxygen binding to one of the subunits of hemoglobin that prompts cooperative binding to other subunits.

We mentioned before that PTMs can lead to conformational changes in proteins. Other methods to induce such allosteric changes include the binding of small molecules, the binding of proteins, changes in the redox environment of a protein, which affects disulfide bonds, and general changes in protein flexibility and dynamics. Let's look at some examples of allostery brought about by these changes. Many of these examples are presented by Laskowski et al. (https://doi.org/10.1016/j.febslet.2009.03.019).

Allosterism by small molecules: opening and closing active sites

Small molecule binding at an allosteric site can lead to small and large conformation changes in a distal active site. An often observed motion is a partial or full hinge-clamping conformational change that brings critical catalytic groups into a more organized and effective active site, which excludes water as a competing hydrolytic substrate (for example). One example of a small hinge-clamping change is seen in phosphoglycerate dehydrogenase (PGDH, which catalyzes the first step (oxidation by NAD^+) in synthesizing serine. It converts 3-phosphoglycerate into 3-phosphohydropyruvate. The end product of the pathway, serine, binds to an allosteric site on PGDH and inhibits it, a classic example of feedback inhibition of the first reaction of a pathway by the end product. PGDH has a regulatory binding domain (RBD) that binds serine, a substrate binding domain (SBD), and an NAD^+/NADH binding domain (NBD), which is where the allosteric inhibitor binds, plus binding domains for substrate (SBD) and the NAD nucleotide (NBD). Figure 7.2.6.14 shows the small hinge-bending conformational change on bind conversion of

monomeric *apo*-PGDH (1psd) to *holo*-PGDH with serine (not shown) bound in the RBD allosteric domain. The RBD-SBD domains, relative to the NBD domain, undergo a 15° rotation.

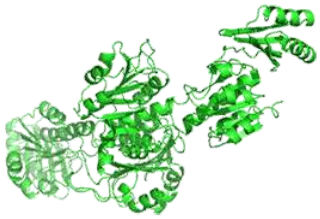


Figure 7.2.6.14: Small hinge-bending conformational changes on bind conversion of monomeric *apo*-PGDH (1psd) to *holo*-PGDH on binding the allosteric end-product inhibitor serine (not shown)

The functional enzyme is a tetramer (only the monomer is shown in the above figure). The rotation in the full tetramer aligns catalytic residues and closes off the active site.

Conformational changes are ubiquitous on the binding of ligands to proteins. These changes also occur on substrate binding at the active site. It is important to recognize that substrate-induced changes in the active site are **not** allosteric since the conformational changes, no matter how big, are caused by binding at the active site, not at an allosteric site. A classic example of a large conformational change on binding occurs when glucose or glucose analogs bind to the active site of hexokinase, the first enzyme in the glycolytic pathway. Figure 7.2.6.15 shows the large conformational change on binding a glucose analog to hexokinase (1hkg, no glucose; 2yhx, with glucose analog). The background is shown in black to remind readers that this is **NOT** an example of allosterism. The changes facilitate catalysis by excluding water, a competing substrate, from the active site after glucose and ATP (not shown) bind. On binding the glucose analog in the active site, the site becomes sequestered ("jaw" clamping down on binding) from the solvent.

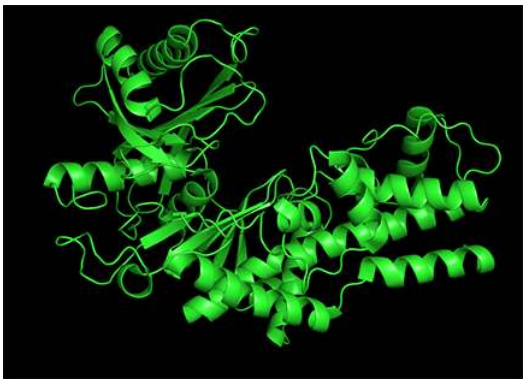


Figure 7.2.6.15: Large, substrate-induced (non-allosteric) conformational change on binding a glucose analog to (1hkg, no glucose to 2yhx, with a glucose analog.

Allosterism by small molecules: Subtitle electrostatic effects

Often, no large conformation change is evident in the protein. In those cases, subtle rearrangements of key residues in the active site or near it (which might promote access) may result from allosteric effector binding. A simple change in the electrostatic environment might account for the effect. Such appears to be the case with chorismate mutase, an enzyme in the bacterial, fungal, and plant pathways for synthesizing aromatic amino acids tyrosine and phenylalanine. The enzyme catalyzes the conversion of chorismate to prephenate, an intermediate in the synthesis of aromatic amino acids. An offshoot pathway takes chorismate to the last aromatic amino acid, tryptophan. The enzyme chorismate mutase is key in the metabolic decision pathway to make tyrosine/phenylalanine or tryptophan. The enzyme is activated by tryptophan and inhibited by tyrosine. The aligned sequences of two crystal structures of the enzyme with the bound allosteric activator tryptophan (Trp 502 in pdb structure 1csm, showing a magenta glutamic acid 23) and with bound inhibitor (Tyr 300 in pdb structure 2csmA showing cyan Glu 23) are shown in Figure 7.2.6.16:

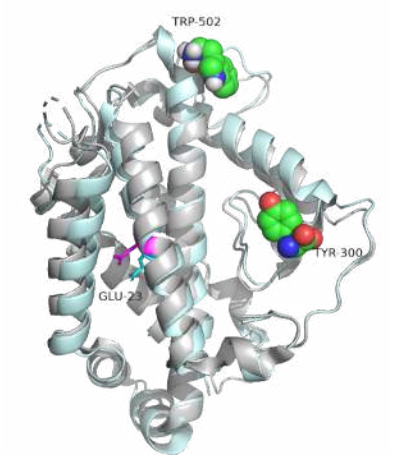


Figure 7.2.6.16: Aligned structures of chorismate mutase with bound allosteric activator tryptophan (Trp 502 , pdbID 1csm , magenta glutamic acid 23) and with bound inhibitor (Tyr 300 pdbID 2csmA, cyan Glu 23)

No significant changes in the protein's overall structure or active site occur. However, the alignment of glutamic acid 23 is different, which may be the basis for the observed allosteric effects. When active, glutamic acid 23 is buried in the active site pocket, but in the inhibited site, it swings into the binding site. Since the chorismate is a charged dicarboxylic acid, Glu 23 probably repels its binding, inhibiting the enzyme's activity.

Allosterism by phosphorylation

We'll focus on one example of a huge conformational change on phosphorylation of a serine side chain in glycogen phosphorylase (GP). This enzyme catalyzes the phosphorolysis of an acetal link between a terminal glucose on a glucose polymer, glycogen, and the next glucose in the chain. This reaction is not a hydrolysis, in which water acts as a nucleophile to cleave acetal bonds. This is a key reaction in metabolism since it cleaves a major energy storage molecule, glycogen.

The enzyme is a multimer that exists in two major states, a T-state and an R-state glycogen phosphorylase (GP). We explored T and R states and allosterism in their interconversion when we discussed oxygen binding to hemoglobin in Chapter 5.3. Crystal structures of the T and R-state bound to allosteric activators IMP and AMP are known. High AMP concentrations imply low ATP concentrations and the need to mobilize glycogen reserves. Conversion to the R state is also promoted by the phosphorylation of Serine 14, which activates the enzyme. The allosteric activators AMP and IMP bind to a disordered loop (313-326) and change it to a form promoted by nucleotide binding. which is disordered in the free structure and adopts a conformation dictated mainly by the type of nucleotide that binds at this site. Figure 7.2.6.17 shows a monomer of the European rabbit glycogen phosphorylase in the R (active) state bound to AMP (3e3n).

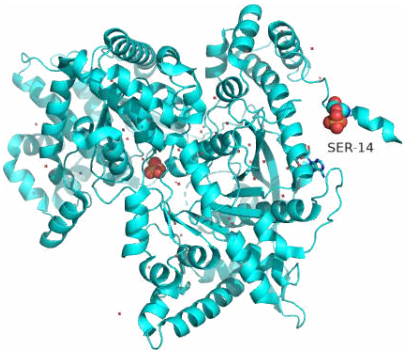


Figure 7.2.6.17: Monomer of the European rabbit glycogen phosphorylase in the R (active) state bound to AMP (3e3n).

Only one chain of the active homodimer is shown. The helix containing the phospho-Ser 14 swings away from the rest of the protein.

Figure 7.2.6.18 shows the difference between the T (inactive) state of rabbit muscle glycogen phosphorylase (7P7D) and the R (active) state (3e3n, neither phospho-Ser 14 nor AMP shown). Note the huge conformational state elicited on phosphorylation of Ser 13.

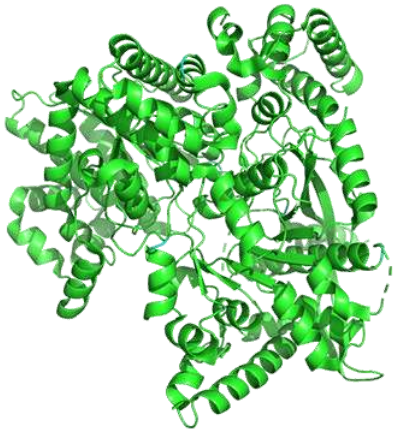


Figure 7.2.6.18: Conformational changes going from the T (inactive) state of rabbit muscle glycogen phosphorylase (7P7D) to the R (active) state (3e3n) on phosphorylation of Ser 13.

Allosterism by small molecules: control of quaternary structure

The binding of small molecules at allosteric sites may promote or inhibit the formation of a protein's correct functional multimeric structure. One example is ATP phosphoribosyltransferase (APRT), which catalyzes the first step in synthesizing histidine. As we saw with phosphoglycerate dehydrogenase (PGDH), APRT has three domains. I and II comprise the active site that is located between them, and a regulatory domain, III (the functional protein is a dimer). As with PGDH, APRT is allosterically inhibited by the end product of the pathway, histidine. On histidine binding to the inactive form (1nh7), a large conformational change results (1nh8), as shown in Figure 7.2.6.19.

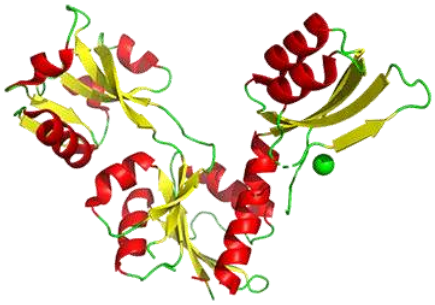


Figure 7.2.6.19: Conformational changes on pathway end product histidine binding to the inactive form (1nh7) form of ATP phosphoribosyltransferase (APRT) inactivating the enzyme (1nh8)

This change drives the formation of a hexamer - $(dimer)_3$, which closes off the active site and inactivates the enzyme. The enzyme ribonuclease reductase behaves similarly.

Allosteric by protein binding

Another protein can bind to an enzyme and activate it by promoting an allosteric change. An example is the binding of a regulatory protein, CyclinA, to cyclin-dependent kinase 2 (CDK2). A large conformational change occurs on binding cyclin to one side of CDK2's catalytic cleft, activating the enzyme by altering active site geometry and making the active site more accessible. CDK enzymes are involved in cell-cycle progression. For more control of the cell cycle, the binding of cyclin and the phosphorylation of CDK2 are required for activity. On binding cyclin, the active site is more available for substrate (ATP) binding. Next, the amino acid that needs to be phosphorylated for activation, Thr 16, is made accessible. Figure 7.2.6.20 shows the conformational change in apo-CDK2 (1hc1) on cyclin binding, which allows ATP binding (1fin).

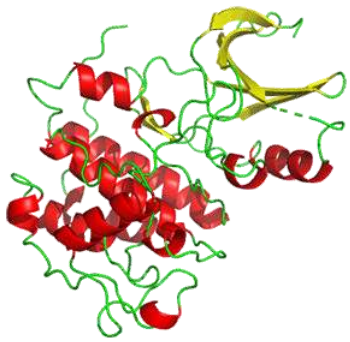


Figure 7.2.6.20 show the conformational change in apo-CDK2 (1hc1) on cyclin binding which allows ATP binding (1fin).

Cyclin (the regulatory protein) is shown in gray. ATP is shown in spacefill. CDK2 is shown as a colored cartoon.

Allosterism by disulfide bonds

It should make sense that cleaving an intrachain or interchain disulfide, whose presence constrains a specific protein conformation, should significantly alter protein structure and function. One example is the protein botulinum neurotoxin type A, a neurotoxin and Zn protease. It cleaves proteins like synaptobrevins, syntaxin, and SNAP-25 in the neurosynapse required for neurotransmitter release.

It has a catalytic (Zn peptidase/protease) and a translocation domain in a single protein chain that contains two disulfide bonds when the protein is synthesized (PDB code 3bta). On cleavage at a select peptide bond, it is split into A and B chains, which remain connected by a single disulfide bond, with the A chain containing the catalytic domain. The B chain contains the translocation domain, which effectively still blocks the active site of the peptidase/protease domain. The disulfides are cleaved when the protein enters an endosome, which contains a more reducing environment. The separate A chain now separates from the inhibiting B domain and expresses protease activity. The active protein then enters the cytoplasm.

Figure 7.2.6.21 shows the domain structure and disulfide bonds of the unprocessed form. The N-terminal Zn proteinase domain (Peptidase_M27, amino acids 1-409) is followed by an intrachain disulfide between C429 and C 453.

Figure 7.2.6.21: The domain structure and disulfide bonds of the unprocessed form of botulinum neurotoxin type A

The mature form is proteolyzed between amino acids 447 and 448 (within the sequences connected by the disulfide (429-453). The protein remains inactive until the disulfide bond is cleaved.

Figure 7.2.6.22 shows an interactive iCn3D model of botulinum neurotoxin type A (3bta).

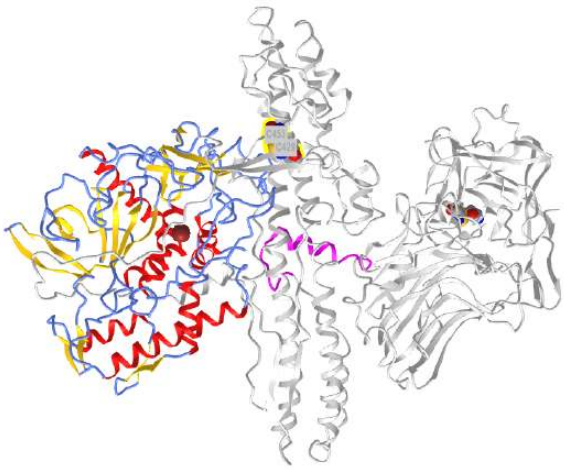


Figure 7.2.6.22: Botulinum neurotoxin type A (3bta) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CfSsBH94jcPyb8

The N-terminal Zn peptidase/protease domain is shown in color, and the other regions of the protein are shown in gray. The two disulfide bonds are shown in spacefill, with CPK color. On cleavage of the disulfide bond between C429 and C453, the Zn protease domain, which is inhibited in the presence of the constraining disulfides, is released and activated.

Allosterism by changes in protein flexibility and population shifts

Sometimes, the crystal structure shows little difference between the apo- and holo-forms of the protein containing a separate allosteric binding site. One example is the enzyme dihydrodipicolinate synthase (DHDPS), which catalyzes the first step in the bacterial and plant pathways for lysine synthesis. DHDPS catalyzes the synthesis of 4-hydroxy-2,3,4,5-tetrahydro-2-dipicolinic acid from aspartate-β-semialdehyde and pyruvate. One explanation for allosterism in the protein is that it occurs on time scales too short to be seen in crystal or NMR structures. Rather, some transient conformation might cause allosterism. Such changes are detectable in the molecular dynamics simulation of the protein in the presence and absence of the allosteric effector molecule.

This possibility suggests that the allosteric effector might change the distribution of possible transient conformational states within a population of such states. If an allosteric effector bound one of the transient states in the population, that state would be "removed" from the ensemble of states, shifting the dynamic equilibrium among the populated states to produce more of the specific conformation that binds the allosteric effector. This sounds like the T to R state conversion in the MWC model oxygen binding to hemoglobin and the allosterism discussed above with the enzyme glycogen phosphorylase, only on a much shorter time scale. This model is consistent with a conformational selection model in which the allosteric molecule binds preferentially to a different preexisting conformational state that exists transiently.

7.2.6.3: Zymogen Activation

A ***zymogen,*** also called a ***proenzyme***, is an inactive precursor of an enzyme. A zymogen requires a biochemical change (such as a hydrolysis reaction revealing the active site or changing the configuration to reveal the active site) to become an active enzyme.

Protease enzymes secreted by the pancreas are initially synthesized as zymogens. The pancreas secretes zymogens to help prevent the enzymes from inappropriately digesting proteins in the pancreatic cells in which they are synthesized. Enzymes like Trypsin are synthesized as proenzymes. For trypsin, trypsinogen is an inactive precursor translated by the rough endoplasmic reticulum and transported to the Golgi apparatus for sorting. Trypsinogen is always co-synthesized and packed with a pancreatic secretory trypsin inhibitor (PSTI) that inhibits premature activation. Thus, there are two mechanisms in place to maintain the inactivity of the protease within the pancreas: (1) synthesis of the zymogen or proenzyme form, and (2) co-expression of a trypsin inhibitor protein that will bind and inhibit any prematurely cleaved trypsin until it has reached the small intestine. An animation showing structural differences between bovine trypsinogen (magenta) with just amino acids 10-15 of the pre-sequence (spacefill orange) and mature trypsin (cyan) is shown in Figure 7.2.6.23.

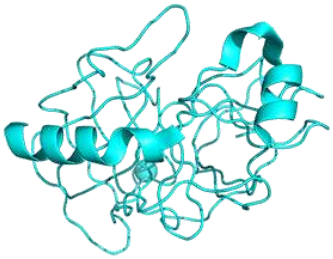


Figure 7.2.6.23: Animation showing structural differences between bovine trypsinogen (magenta) with just amino acids 10-15 of the pre-sequence (spacefill orange) and mature trypsin (cyan)

During packaging within the Golgi system, the trypsinogen and other digestive enzymes condense into core particles and are packed in zymogen granules. The condensed enzymes are stable, and minimal activation occurs within the zymogen granules. Once the pancreatic cells receive secretory stimuli, these zymogen granules are released into the lumen of the pancreatic duct, which carries the digestive enzymes into the duodenum. Once in the duodenum, enteropeptidase activates trypsinogen by removing the 7-10 amino acids trypsinogen activation peptide (TAP) from the N-terminal region (see above). Removal of TAP induces a conformational change, resulting in active trypsin. TAP is immunologically distinct from the same sequence within trypsinogen, thereby allowing the detection of trypsinogen activation *in situ*.

Once activated, trypsin will cleave and activate other zymogen proteases and lipases in the duodenum. These include the activation of elastase, chymotrypsin, carboxypeptidase, and lipase, as shown in Figure 7.2.6.24. Zymogens are also found in other cellular processes. For example, caspases, intracellular proteases, are activated similarly during cellular apoptosis or programmed cell death. The process of blood clotting also involves the activation of zymogens.

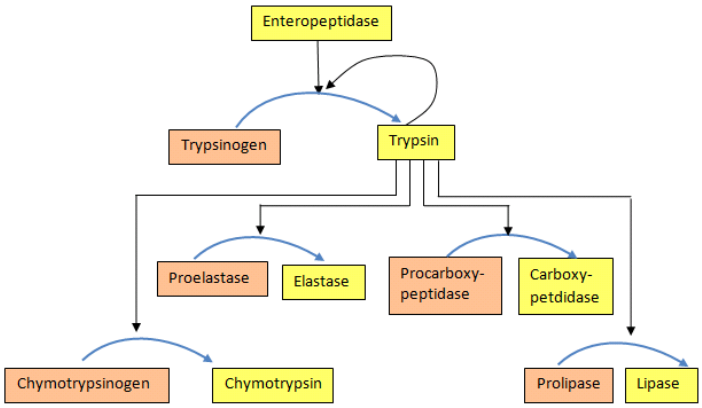


Figure 7.2.6.24: Activation of proenzymes

7.2.6.4: Isozymes

Isozymes (also known as **isoenzymes**) are enzymes that differ in amino acid sequence but catalyze the same chemical reaction. These enzymes usually display different kinetic parameters (e.g. different K_M or K_{cat} values), or different regulatory properties. Isozymes permit the fine-tuning of metabolism to meet the particular needs of a given tissue or developmental stage. Often, isozymes are coded for by homologous genes duplicated within the genome and then diverged over time. *Isozymes* should not be confused with **allozymes, *allelic variants of the same gene locus*** found within a population. Allozymes represent enzymes from different alleles of the same gene. Isozymes represent enzymes from different genes that process or catalyze the same reaction or from the same gene that produces, through differential splicing of primary RNA transcripts, sequences of similar but different sequences. We will focus on isozymes within this section. Part of the regulation derived from the production of different isozymes can arise from the differential expression of isozymes.

Isozymes are usually the result of gene duplication. Over evolutionary time, if the function of the new variant remains *identical* to the original, then it is likely that one or the other will be lost as mutations accumulate, resulting in a pseudogene. However, if the mutations do not immediately prevent the enzyme from functioning, but instead modify either its function or its pattern of

expression. The two variants may both be favored by natural selection and become specialized for different functions. For example, they may be expressed at different stages of development or in different tissues. Some isozymes may also arise from convergent evolution and may not share a high degree of sequence homology or common ancestry.

7.2.6.4.1: The Cyclooxygenase I and II (Cox-1 and Cox-2) Isozymes

As an example of isozymes, we will discuss **cyclooxygenases COX-1** and **COX-2**, which are also called **Prostaglandin Synthases.** They regulate a key step in prostaglandin and thromboxane synthesis and are the targets of **nonsteroidal anti-inflammatory drugs (NSAIDs)**, such as aspirin, ibuprofen, naproxen, and Celebrex (Figure 7.2.6.25). Prostaglandins (PG) are a group of physiologically active lipid compounds called eicosanoids with diverse hormone-like effects in animals. Prostaglandins have been found in almost every tissue in humans and other animals. They are derived enzymatically from the fatty acid arachidonic acid. Every prostaglandin contains 20 carbon atoms, including a 5-carbon ring. They are a subclass of eicosanoids and the prostanoid class of fatty acid derivatives.

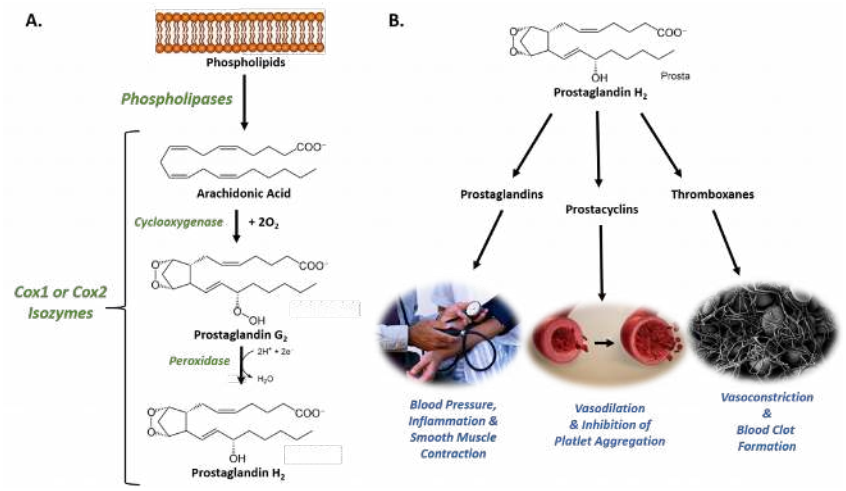


Figure 7.2.6.25: The Role of COX-1 and COX-2 in Prostaglandin Biosynthesis. (A) COX-1 and COX-2 are bifunctional enzymes that mediate the cyclooxygenase and peroxidase reactions that convert Arachidonic Acid to Prostaglandin H_2. (B) Some of the physiological effects that prostaglandins, prostacyclins, and thromboxanes have on biological processes.Images from: Администрация Волгоградской области , scientific animations, and Fuzis

Cycloxygenases are a class of enzymes called dioxygenases, which incorporate both atoms of O_2 into a substrate. We will explore the mechanism of cyclooxygenase in more detail in Chapter 13.5: Biological Oxidation and Reduction Reactions.

The structural differences between prostaglandins account for their different biological activities. A given prostaglandin may have different and even opposite effects in different tissues. The ability of the same prostaglandin to stimulate a reaction in one tissue and inhibit the same reaction in another tissue is determined by the type of receptor to which the prostaglandin binds. They act as autocrine or paracrine factors, with their target cells present near the site of their secretion. Prostaglandins differ from endocrine hormones in that they are not produced at a specific site but in many places throughout the human body and tend to act locally once secreted. Prostaglandins are implicated in various physiological processes such as gastrointestinal cytoprotection, hemostasis, thrombosis, and renal hemodynamics.

Through their role in vasodilation, prostaglandins are also involved in inflammation and can trigger the onset of a fever or the sensation of pain. They are synthesized in the walls of blood vessels. They prevent needless clot formation and regulate the contraction of smooth muscle tissue. The prostacyclins, a special class of prostaglandins, are powerful, locally-acting vasodilators and inhibit the aggregation of blood platelets. Conversely, thromboxanes (produced by platelet cells) are vasoconstrictors and facilitate platelet aggregation. Their name comes from their role in clot formation or thrombosis.

The cyclooxygenases COX-1 and COX-2 regulate the first two steps in prostaglandin and are bifunctional enzymes containing two active sites. The first active site performs the *bis*-oxygenation and cyclization of arachidonic acid, whereas the second active site mediates a peroxidase (reduction) reaction to form PGH_2. The enzyme is an example of a dioxygenase, which uses O_2 as a substrate. The enzyme contains a heme cofactor. The functional enzymes of both COX-1 and COX-2 are homodimers of 70 kDa

subunits, each having an N-terminal epidermal growth factor domain, a membrane binding domain, and a C-terminal catalytic domain. The cyclooxygenase active site is on the opposite side of the peroxidase active site in the catalytic domain.

Figure 7.2.6.26 shows an interactive iCn3D model of arachidonic acid bound to V349I murine COX-2 (6OFY). Just one chain of the dimer is shown.

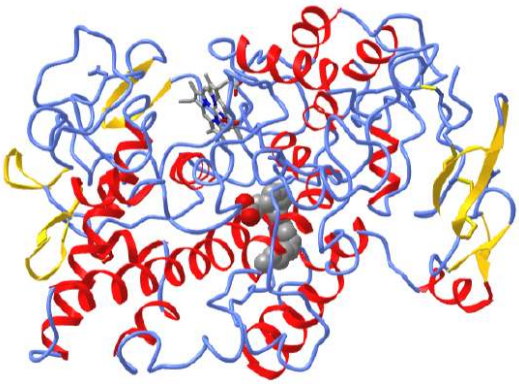


Figure 7.2.6.26: Arachidonic acid and heme bound to V349I murine COX-2 (6OFY) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?gUtX3uAtwoozBvJUA

The heme is shown in stick while the arachidonic acid is shown in spacefill. Note that arachidonic acid is very kinked and not extended due to its four cis double bonds. Three additional amino acids that play key roles in NSAID inhibition of Cox are highlighted. These include Ser 530 (in model 531), which is covalently acetylated by aspirin, Arg 120 (in model 121), and Tyr 355 (in model 356). These play a key role in the binding of NSAIDs and access to the arachidonic acid binding site.

The COX-1 enzyme is widely distributed in many tissues where it is constitutively expressed. Expression of the COX-2 isoform (shown in Figure 7.2.6.2), on the other hand, is normally undetectable in most tissues (except for the central nervous system, kidneys, and seminal vesicles). Various inflammatory and mitogenic stimuli induce COX-2. A third isoform named COX-3 was recently identified as a COX-1 splicing variant. This new variant may play a role in processes such as fever and pain. Additionally, a high level of COX-2 expression is found in several forms of cancer. For example, COX-2 overexpression is related to poor prognosis in certain breast cancers and endometrial adenocarcinomas.

COX-2, unlike COX-1, is induced in inflammatory cells and activated by various inflammatory and mitogenic stimuli. Under these conditions, COX-2 activity leads to the production of prostanoid mediators that trigger important inflammatory processes. Although inflammation is initially a necessary process to fight infection, if it is maintained or remains uncontrolled, it can provoke chronic pathologies and tissue damage. This is why COX proteins are key targets for anti-inflammatory and pain management. Their actions are illustrated in Figure 7.2.6.27.

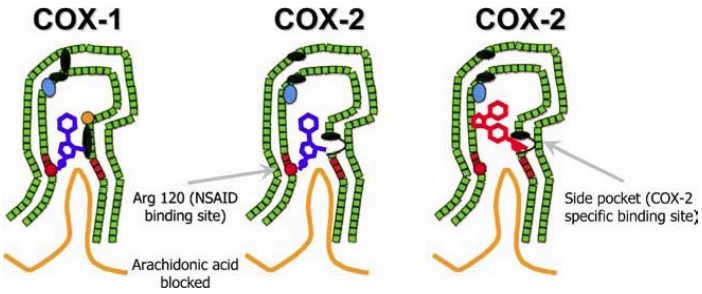


Figure 7.2.6.27: **NSAID Inhibition of COX-1 and COX-2 Enzymes.** The schematic representation of the COX-1 (large green figure) active site inhibited by a nonselective NSAID (central blue figure). The NSAID blocks the entrance channel to COX-1. The binding and transformation of arachidonic acid (bottom yellow figure) within COX-1 is prevented. *Middle Lower Panel* shows the inhibition of COX-2 by a nonselective NSAID (central blue figure). The nonselective binding uses an amino acid residue, Arg120, that is conserved in both enzymes. The *right panel* shows the inhibition of COX-2 by COX-2 selective NSAIDs (central red figure). The COX-2 side pocket allows specific binding of the COX-2 selective NSAIDs. The entrance channel to COX-2 is

blocked. The bulkier COX-2-selective NSAID will not fit into the narrower COX-1 entrance channel, allowing COX-1 to remain active. Upper Figure by Saiz, M., Gonzalez, R., & Garcia, E., (2019) Protopedia and Lower Figure by Meek, I.L., et al. (2010) Pharmaceuticals 3(7):2146-2162.

7.2.6.4.2: Nonsteroidal Antiinflammatory Drugs (NSAIDs)

Clinically available NSAIDs can be separated into three different classes based on their mechanism of action:

- **ASPIRIN:** - Acts to irreversibly inhibit COX 1 & COX-2 by covalent acetylation of Ser 530 in the active site. Most notably, low doses of aspirin can suppress platelet COX-1 activity by 95% or more, an effect that is permanent for the lifetime of the platelet since platelets lack DNA and cannot synthesize new enzymes. Due to aspirin's antithrombotic properties at low doses, this treatment has been found to have cardioprotective effects and is often prescribed for patients at high risk of myocardial infarction. All other NSAIDs interact with COX isoforms reversibly and produce variable COX inhibition (ranging from 50% to 95%) in a time-dependent fashion based on how quickly they are metabolized in the body.
- **NON-SELECTIVE COX INHIBITORS:** Different non-selective NSAIDs have varying inhibitory effects against COX-1 & COX-2 (Figure 8.3). The two most commonly used over-the-counter drugs in this group (ibuprofen & naproxen) produce reversible platelet inhibition ranging from 50 to 95% in a reversible, time-dependent manner (see below). These NSAIDs may be insufficient to provide cardio-protection throughout a commonly used dosing interval and are not commonly used for this purpose. Ketorolac (Toradol ®), an NSAID most commonly used in a hospital setting to treat moderately severe pain, is classified as a non-selective NSAID. However, it is, arguably, a very selective COX-1 inhibitor (Figure 8.3). Inhibition of COX-1 can result in unwanted side effects, such as gastrointestinal discomfort and, in severe cases, ulceration.
- **COXIBS:** Selective COX-2 inhibitors were designed and marketed to treat pain and inflammation while avoiding the GI side effects. However, soon after they were introduced into the market, their use led to the first reported incidence of increased cardiovascular events (myocardial infarction and stroke) in 2004. Rofecoxib (Vioxx ®), one of the most selective COX-2 inhibitors, was removed from the market because of mounting evidence for significant cardiovascular toxicity (Drazen, 2005). Celecoxib (Celebrex ®) is currently the only FDA-approved coxib available in the US, and it has been given a black box warning indicating the potential risk of cardiovascular toxicity. It has a 10-20-fold selectivity for COX-2 over COX-1. Etoricoxib (Arcoxia ®) is a second coxib with ~10^6-fold selectivity for COX-2 over COX-1 that is available outside the United States.

Figure **7.2.6.28** shows the Selectivity and Treatment Efficacy of Nonsteroidal Anti-inflammatory Drugs (NSAIDs).

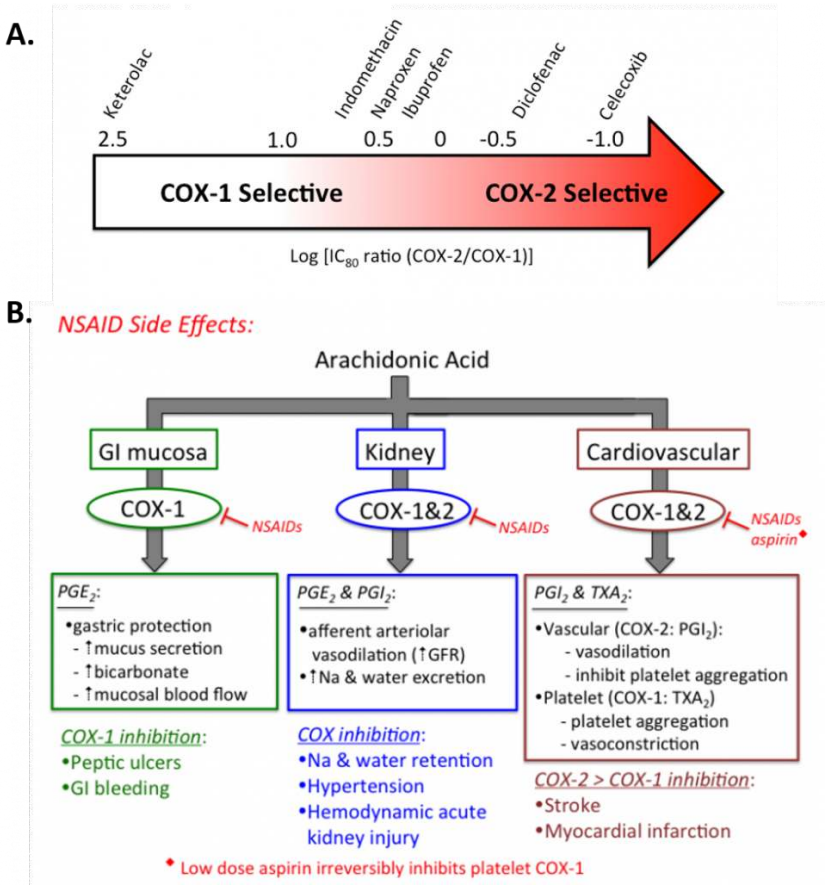


Figure 7.2.6.28: Selectivity and Treatment Efficacy of Nonsteroidal Anti-inflammatory Drugs (NSAIDs). (A) Relative COX-1 & COX-2 selectivity for commonly used non-aspirin NSAIDs. Celecoxib (Celebrex ®) is the only COX-2 selective NSAID on the market in the US. Adapted from Danelich et al (2015). (B) Major physiological roles for COX-1 & COX-2, and mechanisms underlying drug-induced side effects. PGI_2: prostacyclin, TXA_2: thromboxane.*Figures from: Clarkson, C.W. (2018) TUSOM Pharmwiki*

The structures of ibuprofen (Advil), naproxen, aspirin, Celebrex, and acetaminophen (Tylenol, which reduces fever and pain but not inflammation, so it's not an NSAID) are shown in Figure 7.2.6.29.

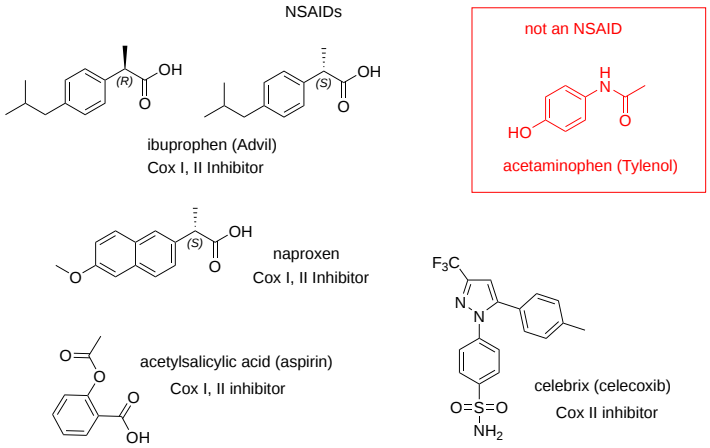


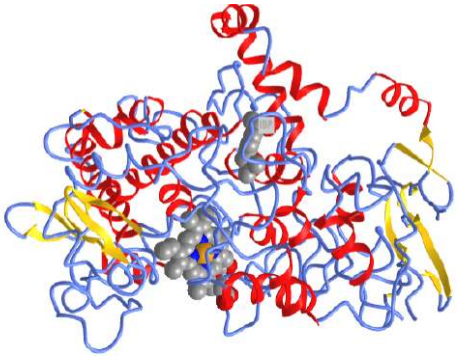
Figure 7.2.6.29: Structures of ibuprofen, naproxen, celebrix (NSAIDs) and acetaminophen (not an NSAID)

Acetominophen - Tylenol

Although this extremely popular drug relieves pain and reduces fever, it does **not** reduce peripheral inflammation and is hence **not classified as an NSAID**. It does have effects on the central nervous system. It does **not** appear to bind to the active site of

COX-1 or COX-2 (no PDB structures available), but it can reduce their activities, probably either upstream or downstream of the pathways these enzymes are part of. It does decrease prostaglandin production in the CNS, which reduces CNS pain and fever. It might act as a reducing substrate in the peroxidase site. Many proteins bind this drug. which is understandable given its relatively simple structure shown in Figure 7.2.6.5. A metabolite of it, p-aminophenol, can be esterified to arachidonic acid to produce the fatty acid amide, which can work through cannabinoid receptors to reduce pain.

Figure 7.2.6.30 shows an interactive iCn3D model of the NSAID ibuprofen bound to cyclooxygenase-2 (4PH9). Just one chain of the functional dimer is shown.



NCBI iCn3D Figure 7.2.6.30: Ibuprofen bound to cyclooxygenase-2 (4PH9). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...syYYjSoeZRGNU8

Ibuprofen is sold as the racemic mixture of R and S enantiomers. Still, it was thought that only the S isomer's inhibition of Cox II led to its antiinflammatory effects, while the other isomer had no effects and no side effects. The crystal structures of IBP bound to both Cox I and II are known. At least in the mouse Cox2, the S-isomer binds more tightly than the R-isomer. Arg-120 and Tyr-355 at the entrance of the cyclooxygenase channel are important in the action of IBP in Cox2 (which are labeled in Figure 7.2.6.6.

As mentioned above, aspirin inhibits COXs through the covalent acetylation of a key serine side chain. The other NSAIDs inhibit prostaglandin H(2) synthase by blocking the channel for arachidonic acid binding. This should make you think they act as competitive inhibitors, which is true for ibuprofen and a methyl flurbiprofen derivative. However, the effect of two other NSAIDs, aclofenac and flurbiprofen, seems to have an added time component not found in the simple competitive inhibition model. They are called slow tight-binding inhibitors, and a simple model describing it is shown in Figure 7.2.6.31.

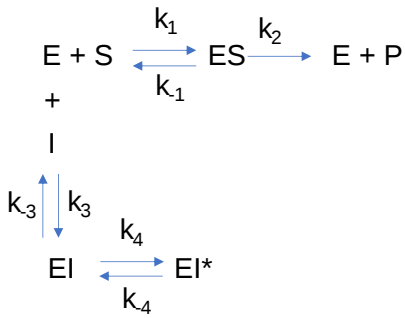


Figure 7.2.6.31: Model for slow binding inhibition

A time step suggests a conformational change after initial binding. Yet crystal studies of the four inhibitors mentioned above show that the NSAID-occupied active sites are essentially the same, suggesting no major global conformational change. It may be that the rate involves slow hydrogen bonding around two key residues, Arg 120 and Tyr 355.

Slow binding and slow tight-binding inhibition

All the enzyme inhibition models we discussed in Chapter 6.4 are based on reversible rapid equilibrium binding of an inhibitor to E (competitive), ES (uncompetitive), and E and ES (mixed/noncompetitive). Equations can be derived for **slow binding**

and **slow tight-binding** inhibition. In these models, the degree of inhibition at a fixed inhibitor concentration changes with time to establish an "equilibrium" among all species. The model above works for both. Both types of inhibitors (slow and slow-tight) inhibit enzymes in a time-dependent fashion. In slow binding, it takes longer to establish equilibrium among the three species, E, EI, and EI*. In slow binding, the steps determining the concentration of EI (k_3I and k_{-3}) are assumed to be fast compared to those involved in determining the concentration of EI* (k_4 and k_{-4}). When [I] >> [E], the following dissociation constants for pure competitive ($K_{IS} = K_C$) and for the slow binding inhibition (K_I), which accounts for mass balance, hold.

$$K_{IS} = K_C = \frac{[E][I]}{[EI]} \tag{7.2.6.1}$$

$$K_I^* = \frac{[E][I]}{[EI] + [EI^*]} \tag{7.2.6.2}$$

In initial rate Michaelis-Menten kinetics, v_0 is measured as a function of [S]. The initial rate v_0 is determined by measuring [P] vs. t at the beginning of the reaction when the substrate is not depleted. Valid v_0 values require that [P] vs t curves are linear. The slope of that line is the initial velocity, v_0. This is true for uninhibited and reversibly inhibited (competitive, uncompetitive, and mixed) reactions. However, with slow binding inhibition, the [P] vs t curves bend with time as the slower formation of EI* occurs. It's a bit like the pre-steady state assumption in enzyme kinetics. At very short times (msec), the P vs. t curves bend as a steady state emerges. It's the same with slow-binding inhibition. This is illustrated in Figure 7.2.6.32.

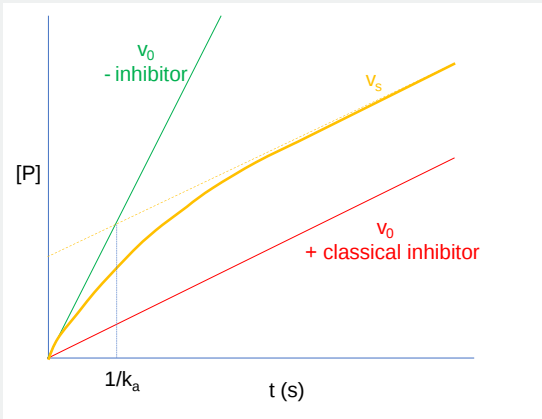


Figure 7.2.6.32: P vs. t curves for uninhibited (green), classical inhibition (red), and slow binding inhibition (gold)

A progress curve equation showing [P] vs. t for slow binding in inhibition is shown below without inhibition.

$$[P] = v_s t + \frac{(v_o - v_s)\left(1 - e^{-k_a t}\right)}{k_a} + C \tag{7.2.6.3}$$

In this equation, k_a is the apparent first-order rate constant to the development of the steady state at a given substrate and inhibitor concentration, v_0 is the very first initial rate, and v_s is the steady-state rate.

Slow tight-binding inhibition occurs when the rate constants for net production of EI* are fast compared to the step for forming EI. In this case, even small amounts of I will produce EI* as the reaction is pulled to EI*, and more complicated equations must be used to determine product vs time equations.

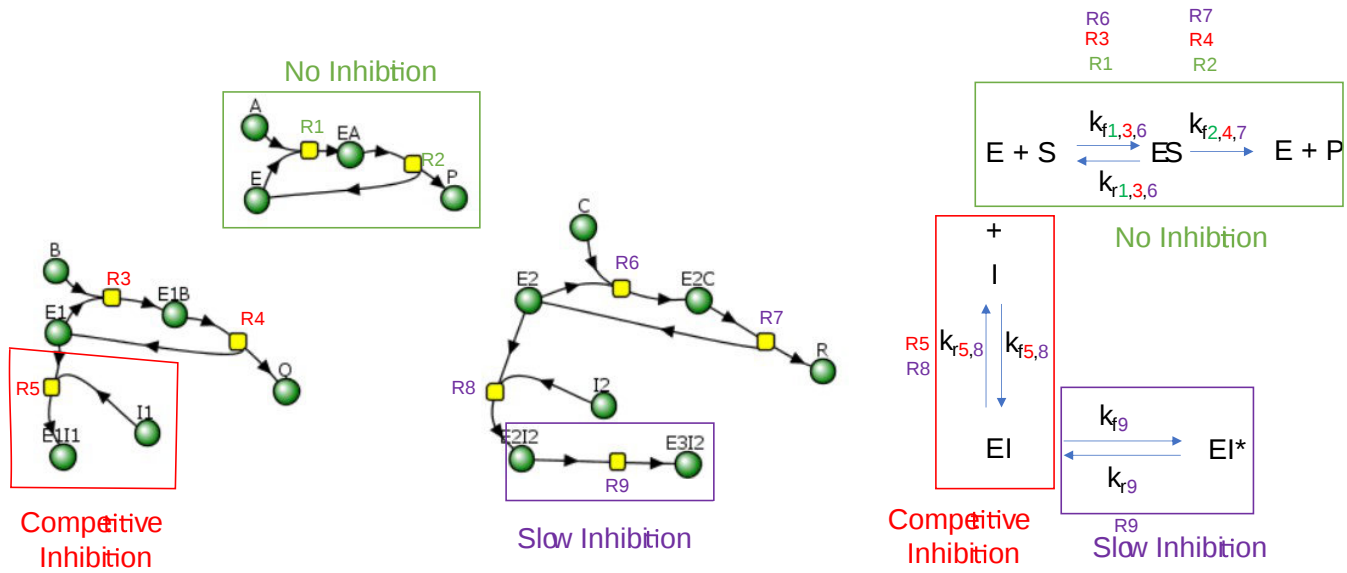
Other chemical models could also account for slow-binding inhibition. These include a slow binding rate constant k_4, a slow isomerization of EI after fast binding of I, or a slow binding to a specific, low population conformation of the enzyme in a process called conformational selection. Slow-binding inhibitors of enzymes (such as neural acetylcholine esterase) are known.

Here is a Vcell simulation showing progress curves for an uninhibited reaction, one in the presence of a competitive inhibitor, and one in the presence of a slow (competitive) inhibitor.

VCell MODEL

Comparison No Inhibition, Competitive Inhibition, and Slow (Competitive) Inhibition

Vcell reaction diagram with all equations based on mass action (not Michael-Menten kinetics)



Concentrations

- A, B and C (substrates), t_0 = 1 uM
- I_1 and I_2 (inhibitors) = 1 uM
- E, E_1 and E_2 (enzymes) at t_0 = 0.1 uM
- P, Q and R (products) at t_0 = 0 uM

Rate Constants

All forward (k_f) and reverse (k_r) rate constants are set initially to 1 except for

- k_{r2} for all reactions = 0 and k_{f4} = 1.

Select Load [model name] below

Load SlowInhibition

Select **Start** to begin the simulation.

Only plots of P, Q and R (products) are initially shown. Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

7.2.6.4.3: Coxibs and the Thromboxane/Prostacyclin Imbalance Hypothesis

Previous research indicates that in the cardiovascular system, more significant inhibition of COX-2 vs. COX-1 (as produced by COX-2 selective “coxibs”) can tip the normal balance between the effects produced by prostacyclin & thromboxane, resulting in an increased likelihood for platelet aggregation and vasoconstriction. These effects can help to explain the higher incidence of myocardial infarction and stroke observed when these drugs have been used clinically. The mechanisms involved are illustrated in Figure **7.2.6.33**.

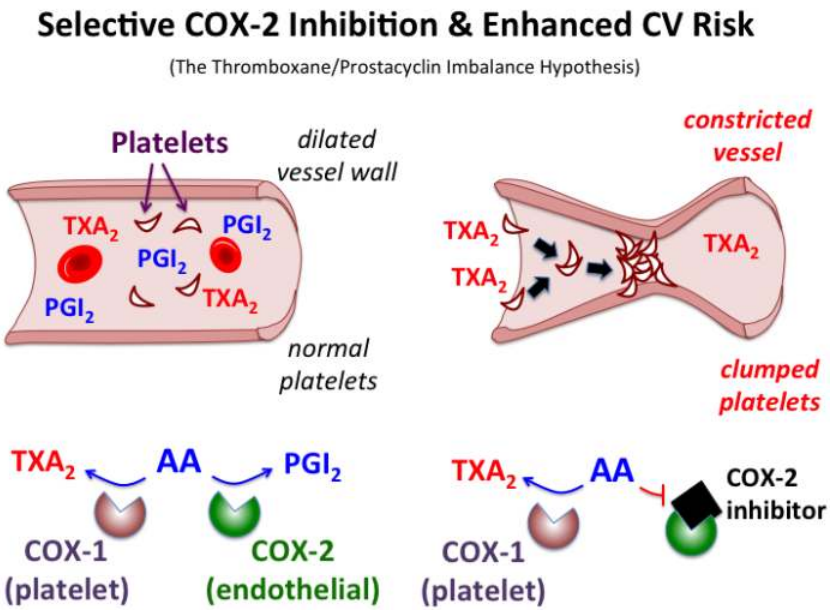


Figure 7.2.6.33: **COX-2 Inhibitors & Cardiovascular Risk.** *Figure from: Clarkson, C.W. (2018) TUSOM Pharmwiki*

The left graphic illustrates the normal balanced effect between prostacyclin (PGI_2) and Thromboxane (TXA_2). PGI_2 is produced primarily by COX-2 activity in the endothelial cell wall of blood vessels. PGI_2 produces vasodilation and inhibits platelet activation. In contrast, TXA_2 is produced primarily by COX-1 activity inside platelets, producing vasoconstriction and enhanced platelet aggregation. Normal vascular homeostasis is maintained when there is a balanced effect of both PGI2 & TXA2. However, when the balance is tipped in favor of TXA_2 formation after selective inhibition of COX-2 (right graphic), vasoconstriction and platelet clumping are more likely to occur, potentially causing an increased risk for cardiovascular events such as myocardial infarction and stroke. Overall, the COX-1/COX-2 isozyme example sheds light on the complexity of biological systems and the ability for slight adjustments in gene expression to create varied and tissue-specific responses.

7.2.6.5: Summary

This chapter explores the multiple layers of regulation that contribute to the dynamic and complex proteome, far exceeding the number of protein-coding genes. It highlights the critical roles of post-translational modifications (PTMs), allosteric regulation, zymogen activation, and the expression of isozymes in fine-tuning protein function in response to cellular and environmental cues.

Proteome Complexity and Post-Translational Modifications (PTMs):
The chapter begins by illustrating how approximately 20,000 human genes generate a vast array of proteoforms through alternative splicing and an extensive repertoire of PTMs. PTMs, which include phosphorylation, acetylation, glycosylation, ubiquitination, sumoylation, oxidation, methylation, and more, dramatically diversify protein structure, stability, activity, and interactions. Both enzymatic and nonenzymatic PTMs contribute to the formation of up to a million different protein variants, enabling precise control over cellular functions. Detailed tables and figures outline common PTMs, the amino acid residues they target, and the enzymes involved.

Detection and Analysis of PTMs:
Advances in analytical techniques, particularly mass spectrometry and antibody-based methods, now allow for large-scale and site-specific identification of PTMs. The chapter discusses how these tools, along with methods such as Western blotting and protein sequencing, have become essential for mapping the dynamic landscape of PTMs and understanding their biological roles.

Allosteric Regulation:
Beyond PTMs, the chapter examines allosteric regulation—how binding events at sites distant from an enzyme's active site can induce conformational changes that modulate activity. Examples include small molecule binding that triggers hinge-clamping motions, subtle electrostatic rearrangements, and changes in protein quaternary structure. The concept of conformational selection is also discussed, emphasizing how shifts in the ensemble of protein conformations can regulate activity and signal transduction.

Zymogen Activation:
To prevent unwanted enzymatic activity, many enzymes are synthesized as inactive precursors known as zymogens. The chapter describes how zymogen activation, through specific proteolytic cleavage or conformational changes, ensures that enzymes such as digestive proteases or caspases become active only under appropriate conditions. This mechanism is crucial for maintaining cellular integrity and preventing self-digestion.

Isozymes and Differential Regulation:
Isozymes, or isoenzymes, are different forms of an enzyme that catalyze the same reaction but possess distinct kinetic properties, regulatory mechanisms, and tissue-specific expressions. These variations, often arising from gene duplication or alternative splicing, allow organisms to finely tune metabolic processes. The chapter concludes with a detailed discussion of cyclooxygenases (COX-1 and COX-2), key enzymes in prostaglandin synthesis that serve as primary targets for nonsteroidal anti-inflammatory drugs (NSAIDs). The distinct roles and regulatory mechanisms of COX isozymes illustrate how subtle differences in enzyme structure and expression can have significant physiological consequences, including effects on cardiovascular health and inflammation.

Overall, this chapter provides a comprehensive overview of the sophisticated strategies employed by cells to regulate protein function post-synthesis. By integrating discussions on PTMs, allosteric modulation, zymogen activation, and isozymes, the chapter underscores the complexity and precision of cellular regulation, setting the stage for further exploration of metabolic control and signal transduction in biochemistry.

7.2.6.6: Some References

1. Shen, Q., Wang, G., Li, S., Liu, X., Lu, S., Chen, Z., Song, K., Yan, J., Geng, L, Huang, Z., Huang, W., Chen, G., and Zhang, J. (2016) ASDv3.0: Unraveling allosteric regulation with structural mechanisms and biological networks. Nucleic Acids Research 44(D1):D527-D535. Available at: https://academic.oup.com/nar/article/44/D1/D527/2503129
2. Wikipedia contributors. (2020, January 19). Isozyme. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, May 19, 2020, from https://en.Wikipedia.org/w/index.php?title=Isozyme&oldid=936548836
3. Wikipedia contributors. (2020, April 30). COX-2 inhibitor. In *Wikipedia, The Free Encyclopedia*. Retrieved 07:00, May 22, 2020, from en.Wikipedia.org/w/index.php?title=COX-2_inhibitor&oldid=954080651
4. Clarkson, C.W. (2018) Major Side Effects of NSAIDs and COX-2 Selective Inhibitors. *TUSOM Pharmwiki*. Available at: http://tmedweb.tulane.edu/pharmwiki/doku.php/nsaid_side_effects?do=
5. Santos, A.L. and Lindner, A.B. (2017) Protein Porsttranslational Modifications: Roles in Aging and Age-Related Disease. *Oxidative Medicine and Cellular Longevity*, Article ID: 5716409. Available at: https://www.hindawi.com/journals/omcl/2017/5716409/#copyright
6. Wikipedia contributors. (2020, May 12). Protein phosphorylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:38, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Protein_phosphorylation&oldid=956228409
7. Szylveszter, K.P., Németh, T., and Mócsai, A. (2019) Tyrosine Kinases in Autoimmune and Inflammatory Skin Diseases. Front. Immunol. 10.3389(2019.01862). Available at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01862/full
8. Wikipedia contributors. (2020, May 7). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:12, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Histone&oldid=955458038
9. Wikipedia contributors. (2020, May 13). Mucin. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:10, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Mucin&oldid=956387296
10. Wikipedia contributors. (2020, March 15). Allosteric regulation. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:07, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Allosteric_regulation&oldid=945637073
11. Dixit, Ajay. Dawra, Rajinder K. Dudeja, Vikas. Saluja, Ashok K. (2016). Role of trypsinogen activation in genesis of pancreatitis.
    Pancreapedia: Exocrine Pancreas Knowledge Base, DOI: 10.3998/panc.2016.25
12. Coll-Martinez, B., and Crosas, B. (2019) How the 26S Proteasome Degrades Ubiquitinated Proteins in teh Cell. *Biomolecules* 9(9):395. Retrieved from: https://www.mdpi.com/2218-273X/9/9/395
13. Liu, W., Tang, X., Qi, X., Ghimire, S., Ma, R., Li, S., Zhang, N., and Si H. (2020) The Ubiquitin Conjugating Enzyme: An Important Ubiquitin Transfer Platform in Ubiquitin-Proteosome System. *Int J Mol Sci* 21(8):2894. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7215765/

This page titled 7.2.6: Enzymes and Protein Regulation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.6: Enzymes and Protein Regulation** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.2.7: Ribozymes - RNA Enzymes

Learning Goals (ChatGPT o1, 1/30/25)

- **Foundations of Catalysis by RNA (Ribozymes):**
  - Explain the concept that catalysis is not exclusive to proteins and that RNA molecules can fold into defined secondary and tertiary structures to form active catalytic centers.
  - Describe the basic principles of electron pushing in chemical reactions, including the roles of electron donors (sources) and electron acceptors (sinks), and how these principles apply to ribozyme catalysis.
- **Classification and Structural Diversity of Ribozymes:**
  - List and classify the different types of ribozymes, including small self-cleaving ribozymes (hammerhead, viroid, hairpin, and riboswitches) as well as larger self-splicing introns (Group I and Group II) and RNP complexes (e.g., RNase P, spliceosome, ribosome).
  - Compare the sizes, structures, and functional roles of small versus large ribozymes, including their dependence on metal ions and other cofactors for catalytic activity.
- **Mechanisms of Ribozyme Catalysis:**
  - Illustrate the general mechanism of RNA-catalyzed phosphodiester bond cleavage, including the roles of general acid and base catalysis, SN2 transesterification, and transition state stabilization.
  - Analyze specific reaction mechanisms for well-characterized ribozymes (e.g., hammerhead, hairpin) by mapping out electron flow (arrow-pushing) and identifying critical nucleotides (e.g., G8, G12) that facilitate catalysis.
- **Functional Examples and Applications:**
  - Examine the catalytic strategies of ribozymes such as self-cleaving reactions in viroids and hairpin ribozymes, noting how they process their own RNA or cleave external substrates (cis versus trans catalysis).
  - Discuss the roles of larger ribozymes (Group I and II introns, spliceosomal introns, RNase P, and the ribosome) in RNA processing and gene expression, and compare their mechanisms to those of small self-cleaving ribozymes.
- **Evolutionary Implications:**
  - Evaluate the hypothesis that ribozymes may have preceded protein enzymes in early evolution, considering their dual role as both genetic material and catalysts.
  - Discuss the significance of ribozyme catalysis in the RNA world hypothesis and how modern ribozymes (and engineered artificial ribozymes) inform our understanding of early biocatalysis.
- **Advanced Functional Ribozymes:**
  - Explore specialized ribozymes such as the methyltransferase ribozyme (MTR1), detailing how small molecule cofactors (e.g., O6-methylguanine) assist in methyl transfer reactions and expand the catalytic repertoire of RNA.
  - Discuss how ribozymes integrate cofactor binding, structural dynamics, and catalytic mechanisms to perform complex chemical transformations.
- **Analytical and Structural Techniques:**
  - Identify and discuss the experimental methods (such as X-ray crystallography, cryo-EM, and interactive structural modeling via tools like iCn3D) used to elucidate ribozyme structures and catalytic mechanisms.
  - Interpret interactive models and figures to understand the active site organization and the role of key residues in ribozyme function.

By achieving these learning goals, students will be well-equipped to understand the chemical and biological principles underpinning ribozyme catalysis, appreciate the structural and mechanistic diversity of RNA enzymes, and recognize their evolutionary significance and practical applications in modern biochemistry.

7.2.7.1: Ribozymes

Any molecule that displays any of the catalytic motifs seen in the earlier chapters (general acid/base catalysis, electrostatic catalysis, nucleophilic catalysis, intramolecular catalysis, and transition state stabilization) can be a catalyst. So far, we have only examined protein catalysts. These can fold to form unique 3D structures, with active sites with appropriate functional groups or nonprotein "cofactors" (metal ions, vitamin derivatives) participating in catalysis. There is nothing special about the ability of proteins to do this. RNA can also form secondary and tertiary structures, as seen in Chapter 8. RNA molecules that act as enzymes are called **ribozymes**.

We are presenting the section before Chapter 8 for a few reasons. Most readers have encountered the structures of RNA and DNA before. You most likely know about three different types of RNA: ribosomal RNA (rRNA), transfer RNA (tRNA), and messenger RNA (mRNA). Likewise, you know from introductory biology classes that the essential dogma of biology, DNA, the holder of the genetic code, is transcribed into RNA, which is translated into a protein sequence. Finally, most have studied (even at the high school level) that the DNA of many species has exons and introns (intervening sequences), the latter of which are spliced out of RNA transcripts to form mature RNA. In this section, we will discuss the catalytic properties of ribozymes, so the introductory background we just mentioned, although important, takes a "second" seat to the chemistry of catalysis, which is the main topic of Chapter 6.

There are 12 classes of ribozymes

- small self-cleaving RNAs (9 classes)
- Group I introns
- Group II introns
- Ribonuclease P

The large ribonucleoprotein nanoparticles, the spliceosome and ribosome, are also functionally ribozymes.

The term ribozyme is used for RNA that can act as an enzyme. Ribozymes are mainly found in selected viruses, bacteria, plant organelles, and lower eukaryotes. Ribozymes were first discovered in 1982 when Tom Cech’s laboratory observed Group I introns acting as enzymes. This was shortly followed by the discovery of another ribozyme, Ribonuclease P, by Sid Altman’s laboratory. Both Cech and Altman received the Nobel Prize in chemistry in 1989 for their work on ribozymes.

Ribozymes can be categorized based on size. Small ones, which usually don't require metal ions for activity, vary from 30-150 nucleotides, while large ones can be a few thousand nucleotides in length. This translates into approximate molecule weights (using this formula for single-stranded RNA: (# nucleotides x 320.5) + 159.0) into 9800 for a 30-mer and 640,000 for a 2000-mer, typical of small and very large proteins/protein complexes, respectively. into two groups depending upon their size – small and large. Large ribozymes, which require metal ions for activity, can vary in size from a few hundred to several thousand nucleotides. Examples of small ones include hammerhead, viroid, hairpin, and riboswitch ribozymes. Large ones include type I and II self-splicing introns, bacterial ribonuclease P, and the RNA in spliceosomes and ribosomes. Many are not true enzymes since they catalyze their own cleavage, although some can cleave presented RNA substrates. Large ones act as true catalysts.

Since RNAs can carry genetic information and act as enzymes, they probably evolved before proteins, which require nucleic acids for their synthesis. In addition, DNA required a special enzyme (ribonucleotide reductase) encoded by DNA to reduce the 2'OH to a 2'H. Artificial ribozymes have been made to catalyze many reactions that require protein enzymes. We will explore some ribozymes in several classes.

7.2.7.2: Small self-cleaving RNAs

We'll consider four small self-cleaving RNAs- hammerhead, viroid, hairpin ribozymes, and the glucosamine-6-phosphate riboswitch (glmS). All catalyze the cleavage of an internal phosphodiester bond (cis catalysis) or in a presented substrate (trans catalysis) by a transesterification reaction. Internal cis-catalysis reactions cleave the ribozyme into two fragments, activating their catalytic activity. They don't act as true catalysts since they engage in only one cleavage cycle. Trans catalysis, in which a substrate RNA binds to and is cleaved by the ribosome, would be considered true catalysis.

The cleavage reaction in an internal cis cleavage is a S_N2 trans-esterification reaction, as shown in Figure **7.2.7.1**.

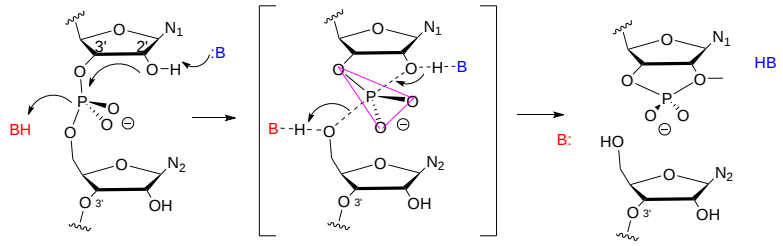


Figure 7.2.7.1: SN2 trans-esterification cleavage of an internal phosphodiester bond in small RNA ribozymes

In the reaction, an adjacent general base (:B) abstracts a proton of the C'2 OH of nucleotide N_1. The resulting 2' O^- acts as a nucleophile in a S_N2 reaction and attacks the δ^+ phosphorous in the phosphodiesterase bond, forming a pentavalent, trigonal pyramidal sp^3d hybridized intermediate/transition state, which collapses, breaking the phosphodiesterase bond between nucleotide N_1 and N_2. This is an inline mechanism: the incoming nucleophilic O in C2' and the exiting one on the O of 5' CH_2OH of nucleotide 2 are axial, separated by 180^0. A general acid, BH, facilitates the departure of the exiting nucleophile by its protonation. Bound metal ions may facilitate the reaction and can be considered cofactors, but might be more involved in maintaining a catalytically active structure. Self-cleaving small RNAs are also found in humans and may be part of long noncoding RNAs.

a. Hammerhead RNA

The hammerhead ribozyme is a small RNA ribozyme with a conserved core with three helical stems. It has a structure similar to the head of a hammerhead shark. One predicted secondary structure of a hammerhead ribozyme is shown below in Figure 7.2.7.2

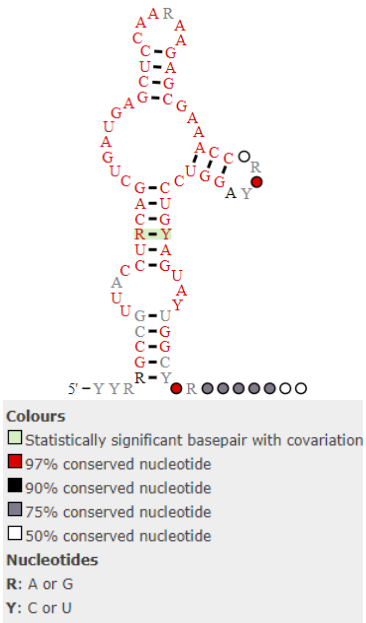


Figure 7.2.7.2: Predicted secondary structure and sequence conservation of the HH9 ribozyme found conserved from lizard to human genomes.

A possible trigonal pyramidal intermediate/transition state in the Hammerhead ribosome is shown in Figure 7.2.7.3.

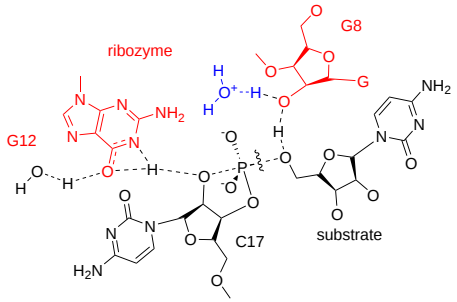
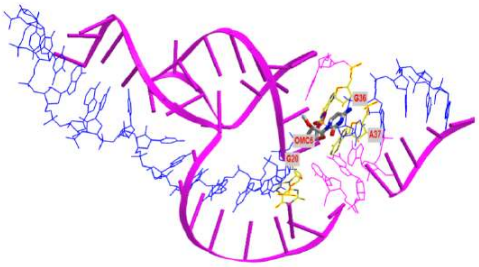


Figure 7.2.7.3: Role of G8 and G12 in active site of full-length Schistosoma mansoni hammerhead ribozyme. After Martick and Scott, Cell. 126, 2006 DOI:https://doi.org/10.1016/j.cell.2006.06.036

A deprotonated G-12 in the ribozyme probably acts as a general base that activates the 2'-OH to form the incoming nucleophile that attacks the trans-substrate RNA. The 2′-OH of G-8 in the ribozyme appears to hydrogen bond to the 5′-O of the departing nucleophile in the substrate, where bond scission occurs. The ribozyme increases the rate by 1000-fold.

Figure 7.2.7.4 shows an interactive iCn3D model of the full-length Schistosoma manson catalytically active hammerhead ribozyme (3dz5). This is an example of a ribozyme that acts in trans as the cleaved phosphodiester bond is a bound RNA single-stranded substrate.



NCBI iCn3D Figure 7.2.7.4: Full-length Schistosoma manson catalytically active hammerhead ribozyme (3dz5). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?qNBw2MY1RZ7xKogH7

The numbering system is a bit different in the iCn3D model above. PDB requires sequential numbering, whereas ribozyme sequences are numbered in discontinuous ways. Core residues that are conserved are given common numbers. However, the different projecting double-stranded RNA regions vary among ribozymes and are numbered differently. The G8 and G12 in Figure 7.2.7.3 are numbered G20 and G36, respectively.

b. ***Viroids***

Viroids are small, single-stranded circular RNAs that infect plant cells. They are not packaged with the viral capsid protein. Some enter cells along with viruses and are called **virusoids** or **viroidlike satellite RNAs**. Intrastrand pairing occurs, and they are synthesized as tandem repeats containing multiple adjacent copies of the viroid. These repeats are cut and ligated to form the individual mature viroid by internal ribozyme sequences. One example is the hepatitis delta virus (HDV), a satellite of the hepatitis B virus. A possible mechanism of catalysis of the hepatitis delta virus ribozyme involving general acid/base catalysis is shown in Figure 7.2.7.5.

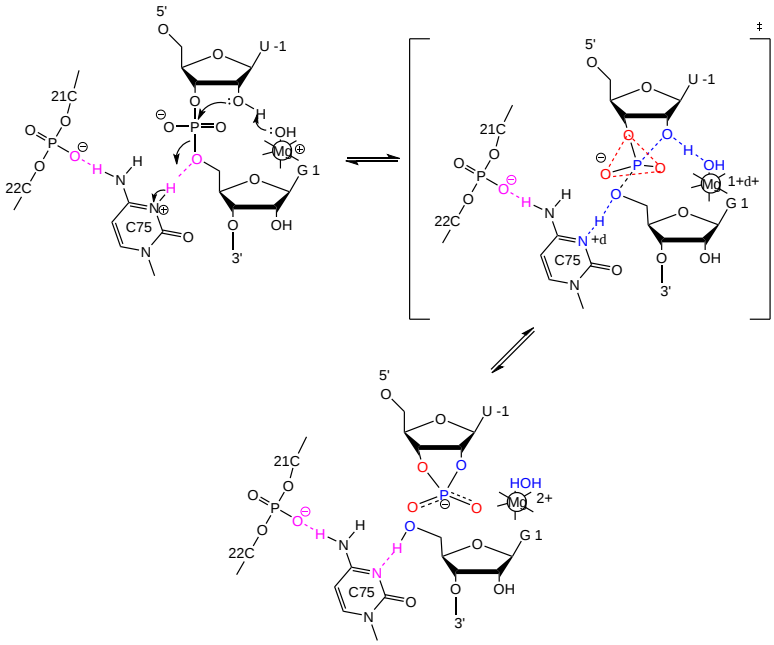


Figure 7.2.7.5: General acid/base catalysis by Hepatitis delta virus (HDV) ribozyme

c. ***Hairpin ribozyme***

The satellite RNA of plant viruses encodes hairpin ribozymes. They are about 50 nucleotides long and can cleave themselves internally or, in a truncated form, can cleave other RNA strands in a transesterification reaction. The structure consists of two domains, stem A, which is required for binding (self or other RNA molecules), and stem B, which is required for catalysis. Self-cleavage in the hairpin ribozyme occurs in stem A between an A and G bases (which are splayed apart) when the 2' OH on the A attacks the phosphorus in the phosphodiester bond connecting A and G to form a pentavalent intermediate.

Rupert et al. solved the crystal structure of a hairpin ribozyme with a non-cleavable substrate analog containing a 2'-O-CH_3 group on the ribose. This acts as a nucleophile in the transesterification cleavage of RNA.

A38 in Stem B appears to be able to interact with the products (the cleaved A now in the form of a cyclic phosphodiester with itself) and the departing G. With a transition state pentavalent analog of the sessile A-G bond in which the phosphodiester linking A and G in the substrate is replaced with a pentavalent vanadate bridge between A and G. This is Illustrated in Figure 7.2.7.6.

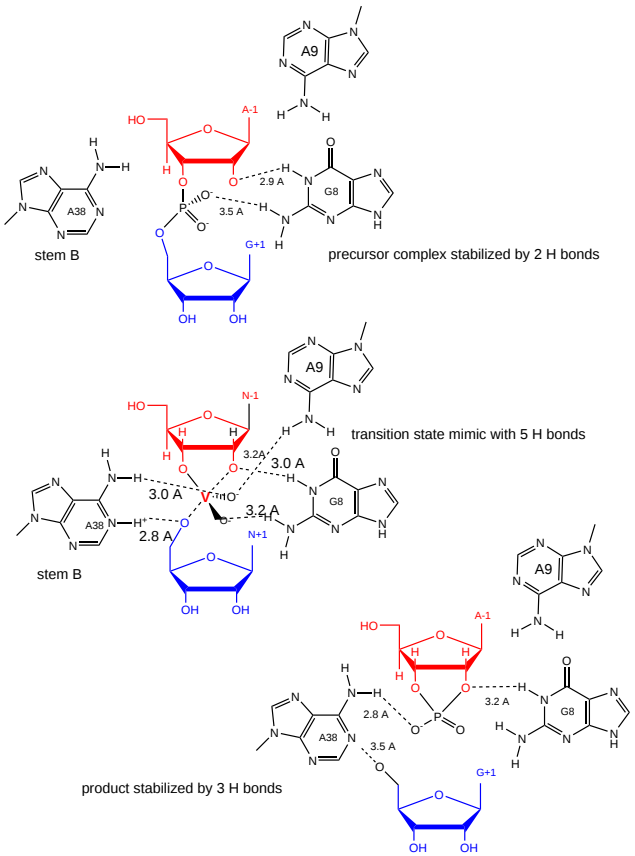


Figure 7.2.7.6: Transition state binding stabilization of hairpin ribozyme (adapted from Rupert et al, Science, 298 pg 1423 2002)

However, A 38 does not appear to react with the sessile A-G groups in the normal substrate, indicating that the main mechanism used by this ribozyme is transition-state binding. Since RNA molecules have fewer groups available for acid/base and electrostatic catalysis (compared to protein enzymes), ribozymes, presumably the earliest type of biological catalyst, probably use transition state binding as their predominant mode of catalytic activity.

Figure 7.2.7.7 shows an interactive iCn3D model of the hairpin ribozyme in the catalytically-active conformation (1M5K).

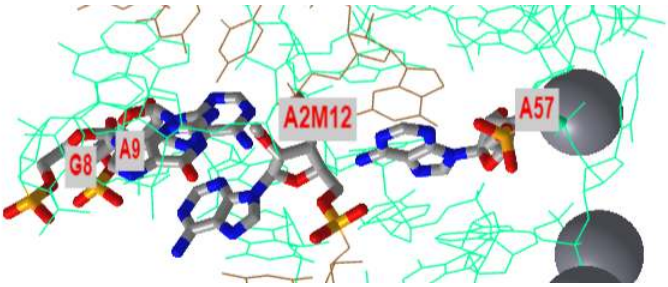


Figure 7.2.7.7: Hairpin ribozyme (human) in the catalytically-active conformation (1M5K). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...goBWNHUbLU7my8

The hairpin ribozyme is shown in cyan sticks, and the inhibitor substrate is shown in brown sticks. The inhibitor contains a 2'-O-methyl adenosine (A2M12), so it can not be cleaved and instead acts as an inhibitor. The A38 shown in the catalytic mechanism is labeled A57 in the iCn3D.

d. ***Glucosamine-6-phosphate riboswitch (glmS)***:

A novel use of ribozymes was recently reported by Winkler et al. They discovered that the 5' end of the mRNA of the gene glmS (from Gram-positive bacteria) is a ribozyme. The GlmS gene encodes glucosamine-6-phosphate synthetase (GlmS), which catalyzes the reaction of fructose-6-phosphate and glutamine to glucosamine-6-phosphate (GlcN6P) and glutamate. This is the first committed step in bacterial cell wall synthesis. Glucosamine-6-phosphate binds to the ribozyme (3' end of the mRNA) and acts as a

cofactor, leading to self-cleavage of the ribozyme. What an amazing mechanism for pathway inhibition. At high GlcN6P concentrations, it binds to the ribozyme, inhibiting its own synthesis. G40 in the active site appears to act as a general base. Figure 7.2.7.8 shows an interactive iCn3D model of GlmS Ribozyme Bound to Its Catalytic Cofactor, glucosamine6-phosphate (GlcN6P) **(2NZ4).**

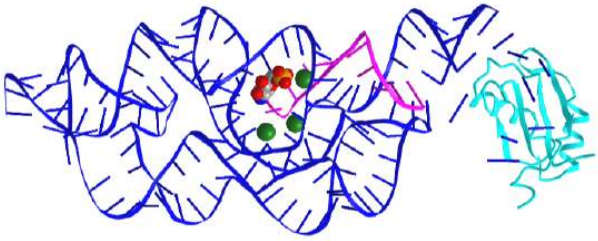


Figure 7.2.7.8: GlmS ribozyme bound to its catalytic cofactor glucosamine 6 phosphate (GlcN6P) (2NZ4) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...J19zdfwEujH4c8

Riboswitches are discussed in greater detail in Chapter 28.1: Regulation of Gene Expression in Bacteria.

7.2.7.3: Group I and Group II Introns

Introns present in RNA molecules must be removed to form mature RNA. About 80% of introns in humans are less than 200 nucleotides long, but some can be 10,000 or longer. Before we discuss introns, we'll provide a quick background on RNA splicing. There are two major types of self-splicing introns, Groups I and II. Other introns are removed by a ribonucleoprotein called the spliceosome. Some call these Group III introns. A simple two-step mechanism for the self-splicing Group I and II introns is shown in Figure 7.2.7.9.

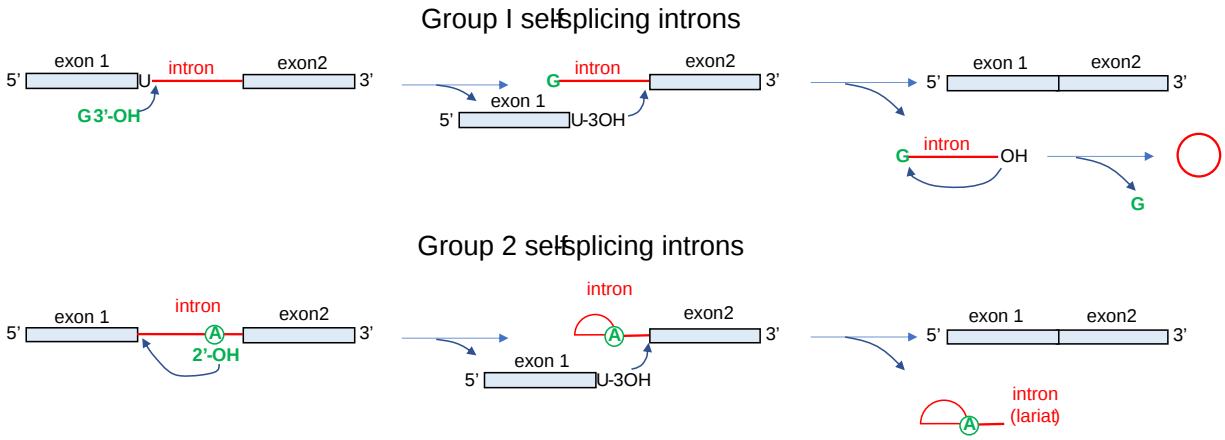


Figure 7.2.7.9: Self-splicing of Group I and II introns. (after N.K. Tanner /FEMS Microbiology Reviews 23 (1999)

Both required first a scission of the RNA strand, followed by ligation of the two exons to form the mature RNA. Note that Group I introns require an external guanosine nucleophile, and the removed intron forms a circular RNA. In contrast, in Group II introns, an internal A residue acts as the first nucleophile in the scission reaction, and the intron, on removal, forms a branched lariat structure.

A simple two-step mechanism for Group II introns, which are spliced out from pre-mRNA in the nucleus by a ribonucleoprotein complex called the **splicesome,** is shown in Figure 7.2.7.10.

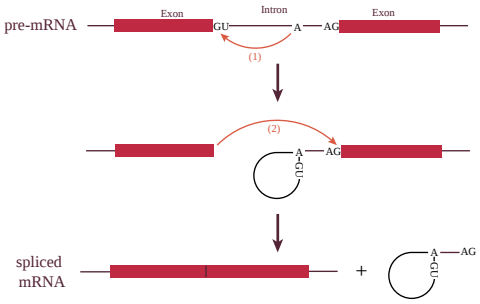


Figure 7.2.7.10: Two steps of canonical RNA processing, from pre-mRNA to spliced RNA and the branched lariat intron. https://commons.wikimedia.org/wiki/F...g_reaction.svg. Creative Commons Attribution-Share Alike 3.0 Unported

Note that the mechanism is extremely similar to the auto-removal of Group II introns, suggesting an evolutionary relationship between the two.

Figure 7.2.7.11 below shows a more detailed view of the catalytic cycle of the spliceosome. Five small ribonucleoproteins (U1, U2, U4/U6, and U5 snRNPs) assemble on the nuclear pre-mRNA and facilitate the intron removal, but the main mechanism involves ribozyme activity.

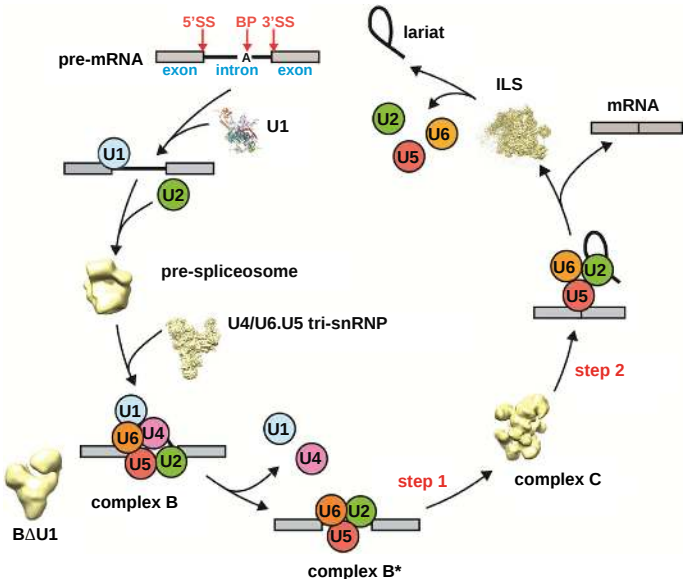


Figure 7.2.7.11: Step-wise spliceosome assembly from its U-snRNP components. Nguyen et al. Current Opinion in Structural Biology 2016, 36:48–57. http://dx.doi.org/10.1016/j.sbi.2015.12.005. CC BY license (http://creativecommons.org/licenses/by/4.0/).

7.2.7.3.1: Group I introns

These are found in bacteria, lower eukaryotes (including mitochondrial and chloroplast RNA), and higher plants and are in ribosomal RNA (rRNA), mRNA, and tRNA. They are also found in Gram-positive bacteriophages (viruses that attack bacteria). As shown in Figure 7.2.7.9, they require guanosine as a cofactor and have a single active site for both scission and ligation to produce the mature mRNA, tRNA, or rRNA. Mg^2 is not needed for catalysis per se, but to maintain the correct tertiary structure of the ribozyme with the correct secondary structure. In Group I introns, a guanosine cofactor initiates the **splicing reaction**. They have one active site that catalyzes the initial cleavage of the phosphodiester bond and the final religation after cleavage.

The Group I catalytic core from Tetrahymena thermophila has two domains. A cleft is formed between them when they pack, which can bind the short helix containing the 5' splice site. and the guanosine cofactor. This "active" site is pre-formed without substrates, similar to the active sites of protein enzymes. Figure 7.2.7.12 shows the secondary structure and the reaction of the group I intron ribozyme from *Tetrahymena*.

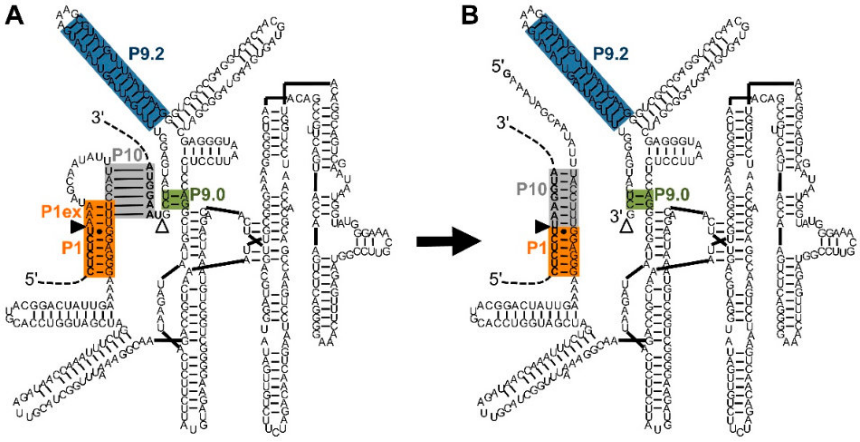
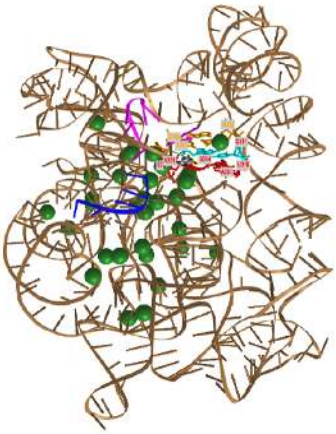


Figure 7.2.7.12: Secondary structure and reaction of the natural, *cis*-splicing group I intron ribozyme from *Tetrahymena*. Müller. *Molecules* **2017**, *22*(1), 75; https://doi.org/10.3390/molecules22010075. Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

The intron runs between the two triangles, which show the 5' (filled triangle) and 3' (open triangle) splice sites. The orange P1 helices contain the 5′-splice site (G:U). The P10 helix (grey), P9.0 helix (green), and the P9.2 helix (blue) are structured to present the appropriate 3′-splice site. During splicing (from (**A**) to (**B**)), the P1 helix extension (orange) is opened to expose the 3′-hydroxyl group of the terminal uridine at the 5′-splice site. The P10 helix then facilitates a conformational change, in which the 3′-exon (upper dashed line) is positioned adjacent to the 5′-exon (lower dashed line), allowing the nucleophilic attack of the 3′-uridine in the 3′-splice site, joining the 5′-exon and 3′-exon.

Figure 7.2.7.13 shows an interactive iCn3D model from cryoEM of the full-length holo L-16 ScaI Tetrahymena ribozyme (7EZ2).



NCBI iCn3D Figure 7.2.7.13: Holo L-16 ScaI Tetrahymena ribozyme (7EZ2) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PhK3BSbFLDnDt8

Three sets of coplanar bases are found in the active site. These include C262, A263, and G312 (top layer, brown), G264, C311, and the ωG - labeled G3 in the iCn3D model- (cyan layer) and A261, A265, and U310 (bottom layer red). The ωG - labeled G3 - is the nucleophile.

The structure is nearly identical to the apo-form of the ribozyme, with just an internal guide RNA sequence undergoing a large change and the guanosine binding site undergoing a small shift on binding RNA substrates.

7.2.7.3.2: Group 2 Introns

Group II introns are found in mRNA of bacteria and some Archaea, in rRNA, tRNA, and mRNA of chloroplasts and mitochondria, and fungi, plants, and protists. No Class 2 introns appear to be found in eukaryotic genomes. Some of these introns are in gene-

encoding proteins, but most are in bacterial noncoding sequences. In Group II introns, an adenosine cofactor initiates the splicing reaction, as shown in Figure 7.2.7.9.

What's especially interesting about group II introns is that they can reinsert into DNA. Hence, they can be considered mobile genetic elements. A **maturase/reverse transcriptase** enzyme (562 amino acids) is associated with the intron, which helps stabilize the active site of the intron for the **reversible** scission and ligation of the intron. Reintegration of the excised branched lariat intron into DNA is called retrotransposition (copy/paste). Mitochondrial and chloroplast Group 2 introns have lost mobility and act as classic introns.

The structures of a group II intron from Thermosynechococcus vestitus (a cyanobacterium) before and after integration have been determined. A branch-site domain VI helix swings 90°, enabling DNA integration. The maturase/revere transcriptase protein assists excision of the intron through the interaction of domain VI of that intron that positions the key adenosine for branched lariat formation during forward splicing, as shown in Figure 7.2.7.9. The changes in the structure of the group II intron retroelement before (6ME0) and after DNA integration (6MEC) are shown in Figure 7.2.7.14.

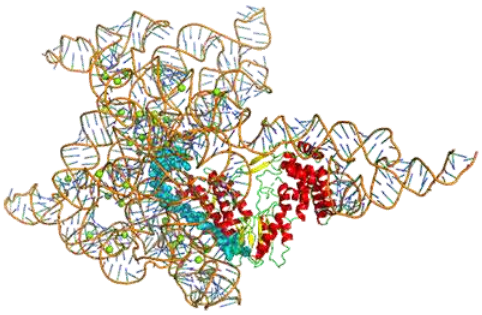
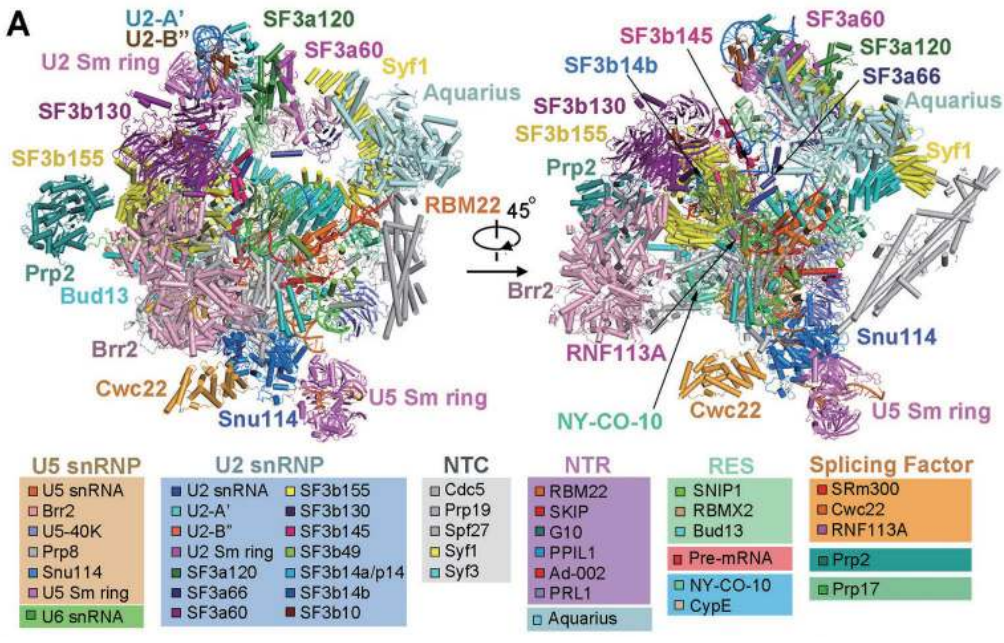


Figure 7.2.7.14: Changes in the structure of the group II intron retroelement before (6ME0) and after DNA integration (6MEC)

The target DNA before integration is shown in cyan spacefill and orange spacefill after integration. The protein is shown as a colored cartoon. The Group II intron is 867 nucleotides, and the sense target DNA is 47 nucleotides in length.

7.2.7.4: Spliceosomal Introns

A comparison of Figure 7.2.7.9 through Figure 7.2.7.11 shows the similarities between spliceosomal introns and group II self-splicing introns. Spliceosomal and group II self-splicing introns are structurally and mechanistically homologous, right down to the stereochemistry of the splicing reaction. In eukaryotes, introns in pre-mRNA are removed by splicing and subsequent exon ligation, releasing the intron as a branched lariat molecule. This reaction is performed by the spliceosome, a large nuclear ribonucleoprotein (RNP) complex. Spliceosomes remove introns and splice the exons of most **nuclear genes**. They are composed of 5 kinds of small nuclear RNA (**snRNA**) molecules and over 100 different protein molecules. The RNA — not the protein — catalyzes the splicing reactions. The molecular details of the reactions are similar to those of Group II introns, which has led to speculation that this splicing machinery evolved from them. Figure 7.2.7.15 shows two views of the cryo-EM structure of the human-activated spliceosome (the Bact complex)



C

Figure 7.2.7.15: Cryo-EM structure of the human-activated spliceosome (the Bact complex). Zhang et al. https://www.nature.com/articles/cr201814.pdf. Creative Commons Attribution 4.0 Unported License. http://creativecommons.org/licenses/by/4.0/

There are 52 proteins, three small nuclear RNAs (snRNAs), and one pre-mRNA. The total molecular mass is 1.8M. U2, U5, and U6 snRNAs are colored marine, orange, and green, respectively. Pre-mRNA is colored red. Figure 7.2.7.16 shows just the structural changes in RNA and protein components that occur between the early Bact complex (left panel) and the mature Bact complex (right panel).

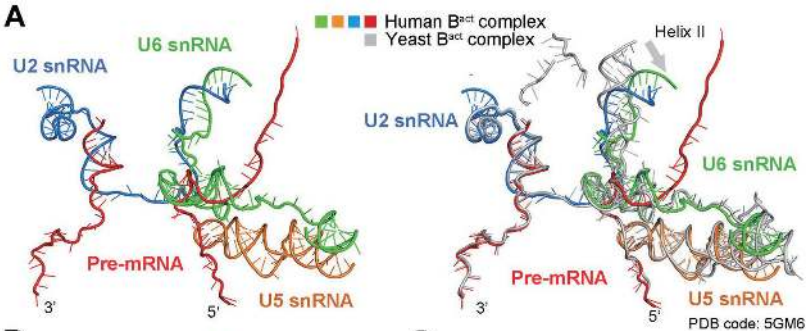


Figure 7.2.7.16: Structural changes in RNA and protein components that occur between the early Bact complex (left panel) and the mature Bact complex of the human spliceosome.

7.2.7.5: Ribonuclease P

This enzyme is a ribonucleoprotein that cleaves RNA through the catalytic action of one essential RNA subunit that displays ribozyme activity. It's found in most organisms. As with many ribozymes, the activity is increased 2-3 fold with bound proteins that stabilize the folded ribozyme and help bind the preferred substrate, pre-tRNA. Figure 7.2.7.17 shows bacterial RNase P ribozyme in complex with tRNA (3q1r).

Figure 7.2.7.17: Bacterial RNase P ribozyme in complex with tRNA (3q1r).

The tRNA is shown as a magenta cartoon/surface, and the protein is in cartoon form. The enzyme cleaves the 5' head end of the precursors of transfer RNA (tRNA) molecules. The enzyme is a heterodimer in bacteria with one RNA and protein subunit.

Figure 7.2.7.18 shows a possible transition state with key residues involved in binding two Mg^{2+} ions in the active site. These ions are essential for catalysis as they stabilize the pentavalent intermediate/transition state.

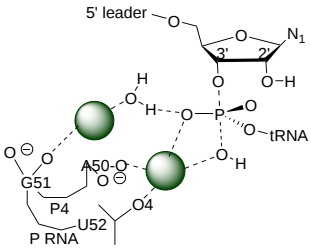
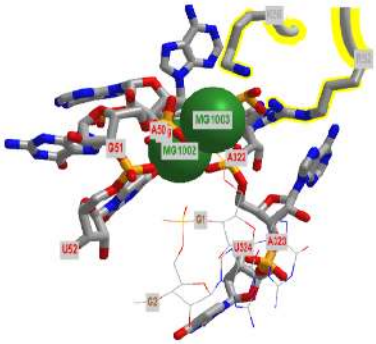


Figure 7.2.7.18: Transition state of RNase P with key RNA active site

Figure 7.2.7.19 shows an interactive iCn3D model of the active site with key residues labeled of bacterial RNase P ribozyme in complex with tRNA (3q1r).



NCBI iCn3D Figure 7.2.7.19: Bacterial RNase P holoenzyme in complex with tRNA (3q1r) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...6SjVMa8g6K47x5

7.2.7.6: RNA Polymerase Ribozyme

If primordial RNA acted as a metabolic enzyme catalyst and the holder of the genetic information, it would need RNA polymerase activity. Artificial ribozymes with class I RNA ligase activity have been made. Figure 7.2.7.20 shows an interactive iCn3D model of the active site region of the Class I ligase ribozyme-substrate preligation complex, C47U mutant, Mg^{2+} bound (3R1L).

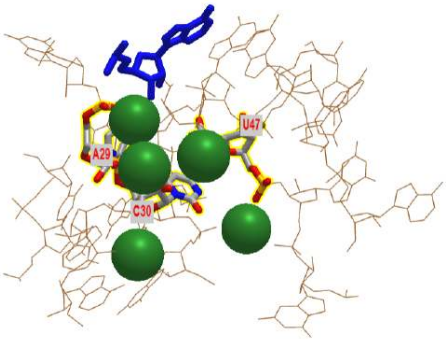


Figure 7.2.7.20: Active site of the Class I ligase ribozyme-substrate preligation complex, C47U mutant, Mg^{2+} bound (3R1L). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...EEQV3Mzc8Jzdv6

The blue stick represents just the 3' terminal adenosine end of the target substrate (5'UCCAGUA3') to which a new nucleoside would be added. The brown represents the active site region of the ribozyme. Three catalytic residues, A29, C30, and C47, have been identified in the actual ribozyme. In iCn3D model is of a mutant, C47U, which has no catalytic activity. The green sphere represents Mg^{2+} ions.

Figure 7.2.7.21 shows the active site residues and how they might facilitate stabilization of the pentavalent intermediate/transition state and the similarity of the active site to a protein RNA polymerase.

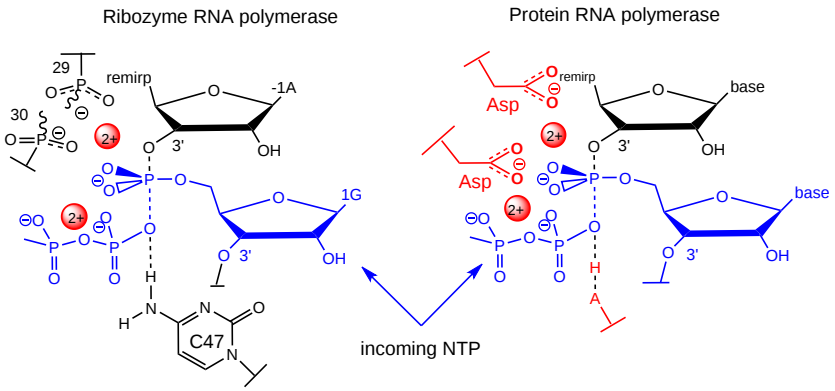


Figure 7.2.7.21: Comparison of Transition State Models of Ribozyme RNA ligase and a protein RNA Polymerases (after Schechner et al)

A divalent Mg^{2+} in the active site of the ribozyme enhances the nucleophilicity of the 3-OH on the primer, which attacks the terminal phosphate of the G(1)TP substrate to form a pentavalent intermediate. The Mg cation is stabilized by oxygens on P 29 and 30 of the ribozyme. The Mg ion also stabilizes the developing charge in the transition state and in the intermediate. Stabilization of analogous divalent cations in the protein polymerase occurs through Asp side changes in the protein.

7.2.7.7: Ribosome

Protein synthesis from mRNA templates occurs on a ribosome, a nanomachine composed of proteins and ribosomal RNAs (rRNA). The ribosome is composed of two very large structural units. The smaller unit (termed 30S and 40S in bacteria and eukaryotes, respectively) coordinates the correct base pairing of the triplet codon on the mRNA with another small adapter RNA, transfer or tRNA, that brings a covalently connected amino acid to the site. Peptide bond formation occurs when another tRNA-amino acid molecule binds to an adjacent codon on mRNA. The tRNA has a cloverleaf tertiary structure with some intrastrand H-bonded secondary structure. The last three nucleotides at the 3' end of the tRNA are CpCpA. The amino acid is esterified to the terminal 3'OH of the terminal A by a protein enzyme, aminoacyl-tRNA synthetase.

Covalent amide bond formation between the second amino acid and the first, forming a dipeptide, occurs at the peptidyl transferase center on the larger ribosomal subunit (50S and 60S in bacteria and eukaryotes, respectively). The ribosome ratchets down the mRNA, so the dipeptide-tRNA is now at the **P** or **Peptide site**. It awaits a new tRNA-amino acid at the **A** or **Amino site**. The figure below shows a schematic of the ribosome with bound mRNA on the 30S subunit and tRNAs covalently attached to amino

acid (or the growing peptide) at the A and P sites, respectively. Figure 7.2.7.22 shows a cartoon model of the prokaryotic ribosome with bound mRNA, tRNAs, and the P and A sites.

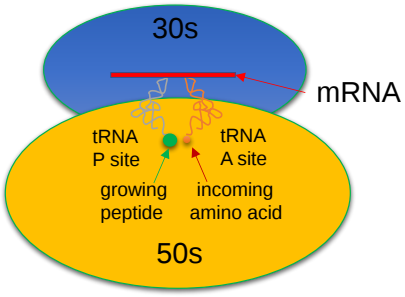


Figure 7.2.7.22: Cartoon model of the prokaryotic ribosome with bound mRNA, tRNAs, and the P and A sites.

We present another, more detailed model of the ribosome complex illustrating protein synthesis in Figure 7.2.7.23.

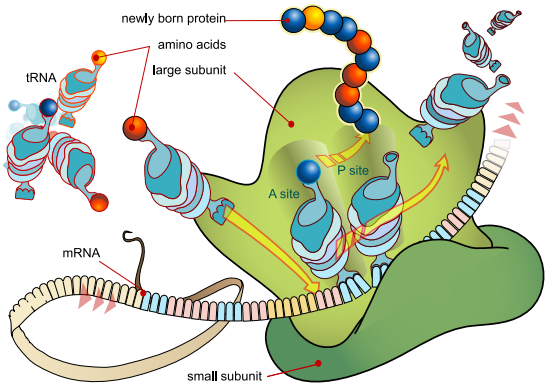


Figure 7.2.7.23: Ribosome mRNA translation. https://en.Wikipedia.org/wiki/File:R...slation_en.svg. public domain by its author, LadyofHats

A likely mechanism (derived from crystal structures with bound substrates and transition state analogs) for forming the amide bond between a growing peptide on the P-site tRNA and the amino acid on the A-site tRNA is shown below. Catalysis does not involve any ribosomal proteins (not shown) since none is close enough to the peptidyl transferase center to provide amino acids that could participate in general acid/base catalysis, for example. Hence, the rRNA must act as the enzyme (i.e., it is a ribozyme). Initially, it was thought that a proximal adenosine with a perturbed pKa could, at physiological pH, be protonated/deprotonated and hence act as a general acid/base in the reaction. However, none was found. The most likely mechanism to stabilize the oxyanion transition state at the electrophilic carbon attack site is precisely located water, which is positioned at the oxyanion hole by H-bonds to uracil 2584 on the rRNA. The cleavage mechanism involves the concerted proton shuffle shown below. In this mechanism, the substrate (Peptide-tRNA) assists its own cleavage in that the 2'OH is in position to initiate the protein shuttle mechanism. (A similar mechanism might occur to facilitate hydrolysis of the fully elongated protein from the P-site tRNA.) Of course, all of this requires perfect positioning of the substrates. Isn't that what enzymes do best? The main mechanisms for catalysis of peptide bond formation by the ribosome (as a ribozyme) are intramolecular catalysis and transition state stabilization by the appropriately positioned water molecule. These processes are illustrated in Figure 7.2.7.24.

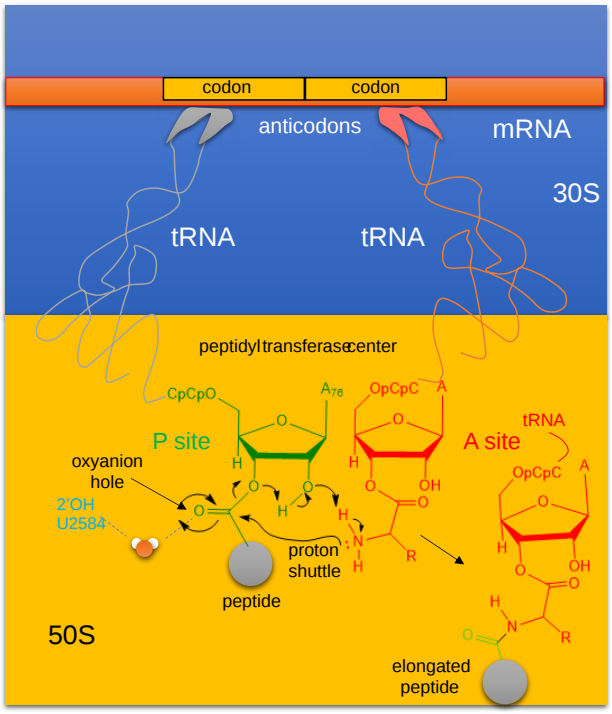


Figure 7.2.7.24: Mechanism Peptide Bond Formation by the Ribosome

The crystal structure of the eukaryotic ribosome has recently been published (Ben-Shem et al). It is significantly larger (40%) with a mass of around $3x10^6$ Daltons. The 40S subunit has one rRNA chain (18) and 33 associated proteins, while the larger 60S subunit has three rRNA chains (25S, 5.8S, and 5S) and 46 associated proteins. The larger size of the eukaryotic ribosome facilitates more interactions with cellular proteins and greater regulation of cellular events. Figure 7.2.7.25 shows the two copies of the 80S yeast ribosome (4v88) presented to humble readers and authors alike.Figure 7.2.7.25: The 80S yeast ribosome (4v88). Each subunit is given a different color.

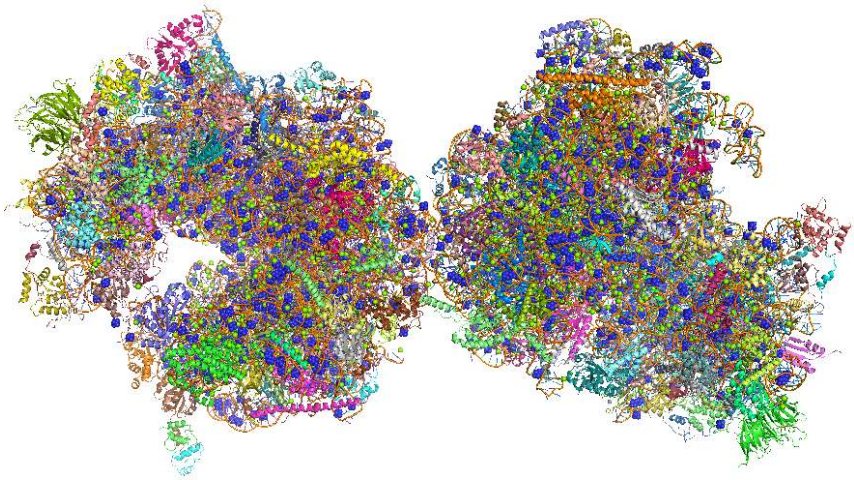


Figure 7.2.7.25: Structure of the yeast 80S ribosome.

7.2.7.8: Ribozyme methyltransferase

The ribozymes described above and generally found in nature catalyze phosphoryl transfer reactions and, for the ribosome, peptide bond formation. In vitro evolution can be used to drive new enzymatic functionalities, which would have been required in an RNA-only world that preceded the use of proteins as catalysts. RNA ribozymes have only four bases that can be employed in binding and catalytic steps, compared to 20 amino acids that can serve the same function in proteins. However, as in the case of protein, small-molecule cofactors that bind to a potential ribozyme might facilitate greater catalytic efficiency and an expanded repertoire of

reaction types. Indeed, we have seen above how small molecules can bind to riboswitches. Figure 7.2.7.26 shows the reaction and structure of a methyltransferase one ribozyme (MTR1) that acts as a methyltransferase. The small ligand, O^6-methylguanine, binds to the ribozymes and acts as a cofactor in the methylation of adenine 63 in the RNA.

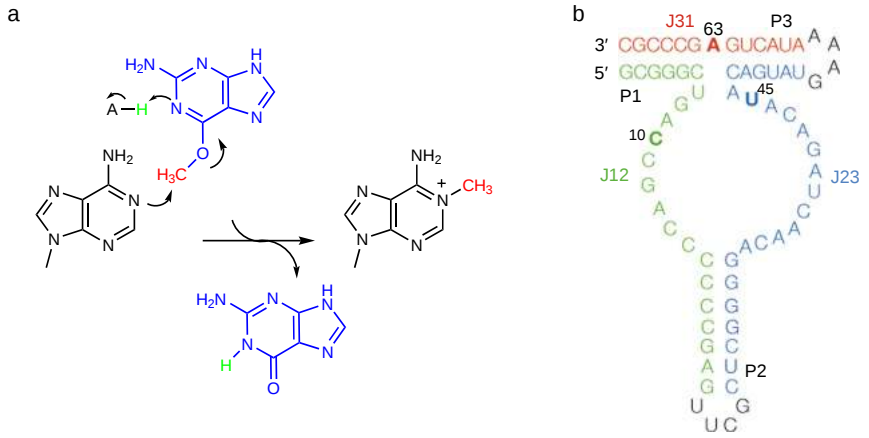


Figure 7.2.7.26: Deng, J., Wilson, T.J., Wang, J. *et al.* Structure and mechanism of a methyltransferase ribozyme. *Nat Chem Biol* **18**, 556–564 (2022). https://doi.org/10.1038/s41589-022-00982-z. Creative Commons Attribution 4.0 International License. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

Panel a shows the chemical reaction in which the methyl group of the small ligand O^6-methylguanine is transferred to N1 of Adenine 63 in the RNA.

Panel b shows the sequence of the MTR1 ribozyme as crystallized for the experiments. The RNA is a three-way junction composed of three arms, P1, P2, and P3. A GNRA tetraloop has been added to the end of the P3 helix so that the entire ribozyme comprises a single RNA strand. Subsections of the strands are named J12 (colored green), J23 (colored blue), and J31 (colored red).

X-ray crystal structures of the ribozyme in the presence of the cofactor O^6-methylguanine were determined. The final structure contained guanine and an A63 methylated adenosine (1MA), implying the methyl group of the O^6-methylguanine had transferred to A63, leaving guanine bound in the active site. The structure of the guanine bound to the ribozyme is shown in Figure 7.2.7.27.

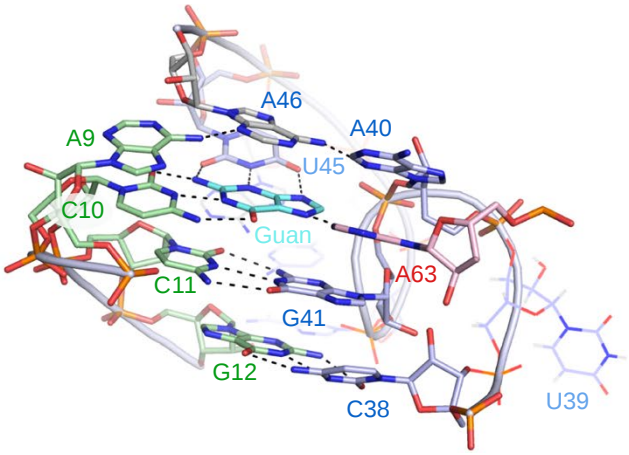


Figure 7.2.7.27: The four planes of nucleobase interactions in the core of the ribozyme. The four planes are composed of (from bottom to top) G12•C38, C11•G41, exogenous guanine hydrogen bonded to C10, U45, and A63, and the triple interaction A9•A46•A40. Deng, J et al., ibid.

Figure 7.2.7.28 shows a hypothetical reaction mechanism.

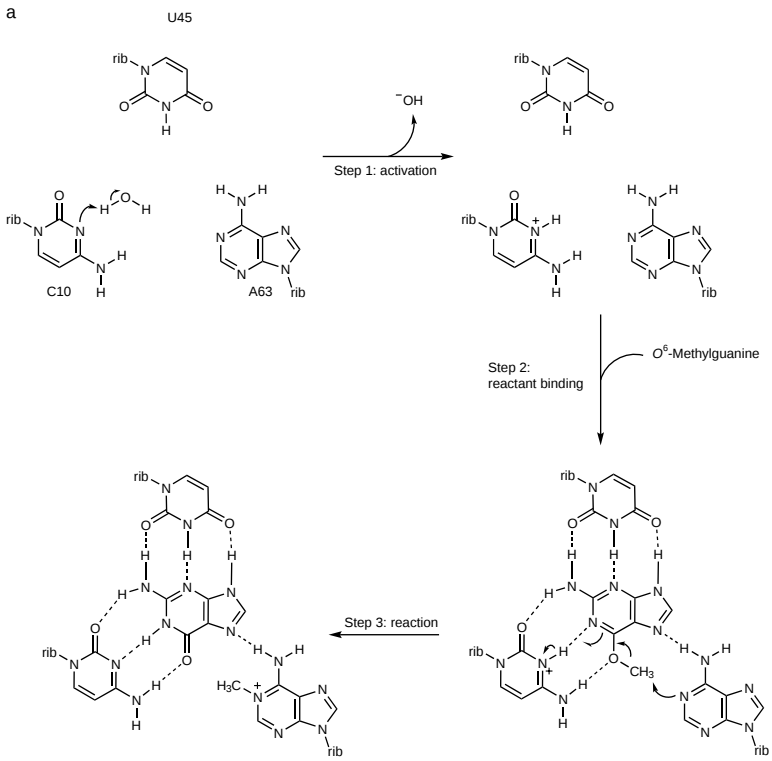


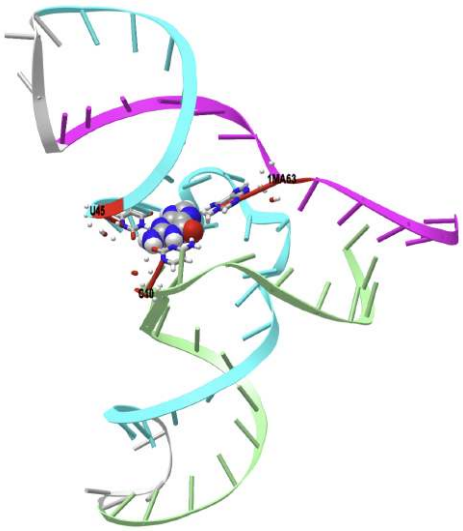
Figure 7.2.7.28: A proposed catalytic mechanism for the ribozyme methyltransferase. Deng, J et al., ibid.

In step 1, the nucleobase of C10 becomes protonated, and in step 2, the O^6-methylguanine becomes bound. However, these two steps are likely coordinated, as the binding will raise the pK_a of the cytosine.

The methyl transfer reaction occurs in step 3 by the nucleophilic attack of A63 N1 on the methyl group of O^6-methylguanine and the coordinated breakage of the guanine O6–C bond. This involves a train of electron transfers, the movement of the proton from C10 N3 to guanine N1, and a concomitant shift of the positive charge from C10 to the N^1-methyladenine at position 63. In principle, the guanine can now be released as a product, although there is no evidence that this occurs with the present form of the ribozyme. Regeneration of an active ribozyme would also require an exchange of the substrate strand to place an unmethylated adenine at position 63.

The proposed mechanism is fully consistent with the structure of the MTR1 riboswitch and the effect of the substitutions at C10 and U45 on activity. The complete loss of methylation activity of the C10U variant is fully consistent with the proposed role as a general acid in addition to ligand binding.

Figure 7.2.7.29 shows an interactive iCn3D model of a methyltransferase ribozyme (7V9E).



NCBI iCn3D Figure 7.2.7.19: A methyltransferase ribozyme (7V9E). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LtyVP4DBnkEgH6

The parts of the ribozyme are named J12 (light green), J23 (cyan), and J31 (magenta). The active site residues, C10, U45 and 1MA63 (1-methyladenosine) are shown in CPK-colored sticks and labeled (disregard small separated spheres). The Ba^{2+} ion in the crystal structure is not displayed. The guanine is shown in spacefill.

7.2.7.9: Summary

This chapter examines the fascinating world of ribozymes—RNA molecules that not only store genetic information but also function as catalysts. It emphasizes that the capacity for catalysis is not unique to proteins; RNA can fold into complex secondary and tertiary structures that support catalytic activity. The chapter is organized around several key themes:

1. **Introduction to Ribozymes and Catalytic Principles:**
   The chapter opens by discussing the fundamental concept that any molecule capable of forming defined three-dimensional structures and containing functional groups—such as those needed for general acid/base, electrostatic, nucleophilic, or transition state stabilization catalysis—can act as a catalyst. Ribozymes illustrate this principle by employing RNA's limited chemical repertoire to drive electron flow from donors (sources) to acceptors (sinks) during bond making and breaking.

2. **Classification and Diversity:**
   Ribozymes are classified based on size and function. The chapter highlights that there are twelve classes of ribozymes, including small self-cleaving RNAs (e.g., hammerhead, viroid, and hairpin ribozymes, as well as riboswitches like glmS) and large catalytic RNAs such as Group I and Group II introns, ribonuclease P, and the catalytic RNAs of the spliceosome and ribosome. These examples demonstrate both cis (self-cleavage) and trans catalytic activities, illustrating the diversity of mechanisms by which RNA enzymes can operate.

3. **Mechanisms of Catalysis:**
   Detailed mechanistic insights are provided into how ribozymes achieve catalysis. The chapter explains the SN2 transesterification mechanism typical of self-cleaving ribozymes, where a general base activates a 2'-OH nucleophile that attacks an adjacent phosphodiester bond, leading to bond cleavage and formation of a pentavalent transition state. Examples such as the hammerhead ribozyme illustrate how conserved nucleotides (e.g., G8 and G12) and even subtle conformational arrangements contribute to efficient catalysis.

4. **Functional and Structural Examples:**
   Using interactive models and structural diagrams, the chapter explores specific ribozyme examples:

   - **Hammerhead Ribozymes:** Discussed in terms of their conserved secondary structure and active site architecture, which enable efficient self-cleavage and even trans-cleavage of RNA substrates.

- **Viroids and Hairpin Ribozymes:** These small RNAs demonstrate the ability of ribozymes to function in viral and satellite RNA contexts.
- **Glucosamine-6-phosphate Ribozyme (glmS):** A unique example where the ribozyme is embedded in an mRNA and uses its own catalytic activity, modulated by binding of a small molecule cofactor, to regulate gene expression.
- **Group I and II Introns and Spliceosomal RNPs:** These larger ribozymes exemplify how RNA catalysis is integrated into essential processes such as RNA splicing and ribosome function, hinting at an evolutionary legacy from an RNA world.

5. **Evolutionary and Functional Significance:**
   The chapter considers the evolutionary implications of ribozymes, noting that RNA-based catalysis may have preceded protein enzymes. This dual functionality of RNA—as both genetic material and catalyst—supports the RNA world hypothesis and underscores the importance of ribozymes in early biochemistry.

6. **Integration with Modern Techniques:**
   Finally, the chapter highlights the methods used to study ribozymes, such as X-ray crystallography, cryo-electron microscopy, and interactive modeling tools like iCn3D. These techniques have revealed detailed insights into ribozyme structures, active sites, and catalytic mechanisms, bridging the gap between theoretical models and experimental observations.

Overall, this chapter provides a comprehensive overview of ribozyme catalysis, demonstrating how RNA can serve as an enzyme through diverse structural strategies and catalytic mechanisms. It sets the stage for deeper exploration of RNA's roles in modern biology and the evolutionary transition from RNA to protein catalysts.

This page titled 7.2.7: Ribozymes - RNA Enzymes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.7: Ribozymes - RNA Enzymes** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.2.8: Cofactors and Catalysis - A Little Help From My Friends

Learning Goals (DeepSeek, 1/30/25)

- **Understanding Cofactor Categories and Functions:**
  - Define and differentiate between metal cofactors and coenzymes, including the concept of prosthetic groups versus loosely bound coenzymes.
  - Explain how vitamin-derived coenzymes (e.g., TPP, FAD, NAD^+, PLP) and metal ions (e.g., iron in heme) contribute to enzyme catalysis by acting as electron sources or sinks.
- **Mechanistic Role in Electron Pushing:**
  - Describe the principles of electron flow in chemical reactions, identifying what constitutes an electron “source” and an electron “sink” in reaction mechanisms.
  - Illustrate, using arrow-pushing techniques, how cofactors facilitate electron transfer during bond-making and bond-breaking.
- **Specific Cofactor Mechanisms:**
  - Analyze the role of thiamine pyrophosphate (TPP) in the decarboxylation of α-keto acids and discuss its chemical mechanism in enzymes like pyruvate dehydrogenase.
  - Explain how FAD functions in redox reactions, particularly in the hydride transfer during the oxidation of succinate, and how NAD^+/$NADP^+$ participates as a dissociable electron acceptor in oxidation reactions.
  - Describe the versatile roles of pyridoxal phosphate (PLP) in catalyzing reactions such as decarboxylation, β-elimination, racemization, and transamination, emphasizing the importance of Schiff base formation and electron sink properties.
- **Integration of Cofactor Function with Enzyme Structure:**
  - Discuss how cofactors bind to specific residues (e.g., coordination of heme by histidines) and how this interaction is critical for proper enzyme function.
  - Evaluate how cofactor binding influences enzyme conformation, stability, and catalytic activity.
- **Clinical and Biological Implications:**
  - Relate vitamin deficiencies (e.g., lack of thiamine or niacin) to impaired coenzyme function and the resulting metabolic or disease states.
  - Discuss how alterations in cofactor binding or function can serve as regulatory mechanisms in cellular metabolism.
- **Analytical and Structural Approaches:**
  - Utilize interactive structural models (e.g., iCn3D) to examine cofactor binding sites within enzymes, and interpret how structural data supports mechanistic insights.
  - Compare reaction mechanisms that do and do not require cofactors (e.g., spontaneous decarboxylation vs. TPP-facilitated decarboxylation) to appreciate the catalytic advantages conferred by cofactors.

Achieving these goals will provide students with a comprehensive understanding of how cofactors serve as essential components in enzymatic catalysis by mediating electron flow, thereby linking fundamental chemical principles with the regulation of biological processes.

7.2.8.1: Cofactors and Electron Pushing: Sources and Sinks

To make and break bonds, electrons have to be moved. In drawing reaction mechanisms, we showed how electrons move from "sources" to "sinks." In many enzyme-catalyzed reactions, vitamin derivatives are used as substrates or "cofactors" or "coenzymes" to facilitate the flow of electrons in bond-making and breaking. The section focuses on cofactors, which facilitate the flow of electrons from the substrate to the product. We will see these enzymes in more detail in specific chapter sections.

Cofactors are molecules that bind to enzymes and are required for catalytic activity. They can be divided into two major categories: ***metals*** and ***coenzymes. Metal cofactors*** commonly found in human enzymes include iron, magnesium, manganese, cobalt, copper, zinc, and molybdenum. ***Coenzymes*** are small organic molecules that are often derived from vitamins. ***Coenzymes*** can bind loosely with the enzyme and be released from the active site. As such, they are also considered substrates for the reaction. Alternatively, they may be tightly bound and cannot dissociate easily from the enzyme. In this case, after their initial participation in an enzyme-catalyzed reaction, the enzyme would no longer be able to use the cofactor in another round of catalysis until the initial state of the cofactor is reformed, which takes another chemical reaction and often an additional substrate.

Tight-binding coenzymes are referred to as ***prosthetic groups***. Enzymes not yet associated with a required cofactor are called ***apoenzymes***, whereas enzymes bound with their required cofactors are called ***holoenzymes.*** Sometimes, organic molecules and metals combine to form coenzymes, such as in the case of the heme cofactor (Figure 7.15). Coordination of heme cofactors with their enzyme counterparts often involves interactions with histidine residues, as shown in the succinate dehydrogenase enzyme shown in Figure 7.2.8.1.

Figure 7.2.8.1: The Heme Cofactor. The family of heme cofactors contains an iron metal coordinated with a porphyrin ring structure, as shown in the *left-hand panel* within the structure of Heme B. In the right-hand panel, Heme B is shown complexed with the succinate dehydrogenase enzyme from the Kreb Cycle. The structure of Heme B shown in the left-hand panel is from Yikrazuul, and the crystal structure of Succinate Dehydrogenase complexed with Heme B is from Richard Wheeler.

Many biological cofactors are vitamin B derivatives, as shown below in Table 7.2.8.1. Many vitamin deficiencies cause disease states due to the inactivity of ***apoenzymes*** that can not function without the correctly bound ***coenzyme.***

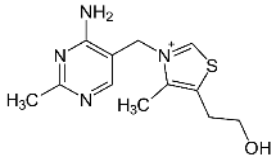
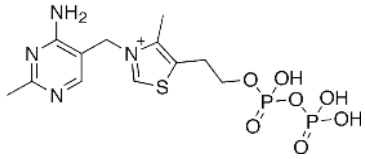
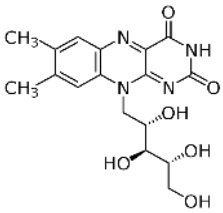
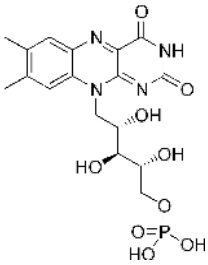
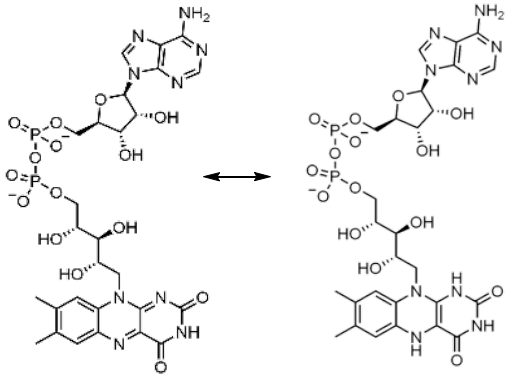
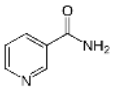
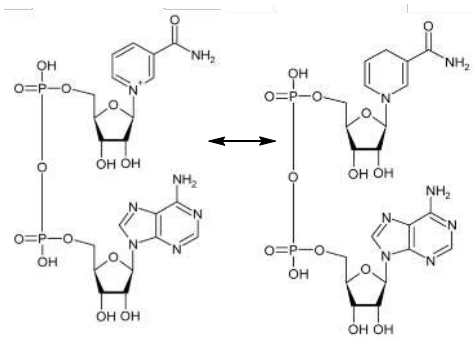
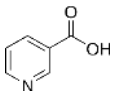
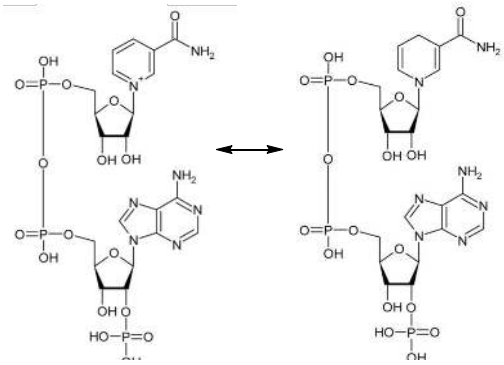
B Vitamins	Modified Enzyme Cofactors
Vitamin B1 (Thiamine)	Thiamine Diphosphate (TPP)
	
Vitamin B2 (Riboflavin)	Flavin Mononucleotide (FMN)
	
	Flavin Adenine Dinucleotide (FAD <-> $FADH_2$)
	
Vitamin B3 (Niacinamide) - amide form	Nicotinamide adenine dinucleotide (NAD <-> NADH)
	
Vitamin B3 (Niacin) - carboxylic acid form	Nicotinamide dinucleotide phosphate (NADP <-> NADPH)
	

Table 7.2.8.1: Essential B-Vitamins and their Modified Enzyme Cofactors

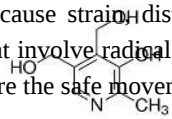
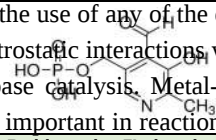
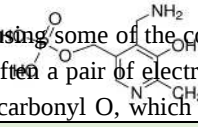
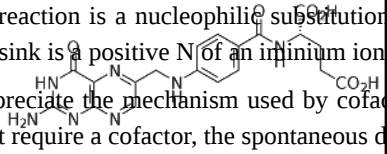
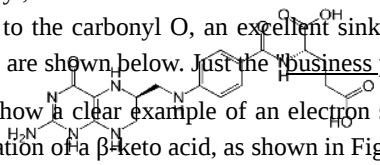
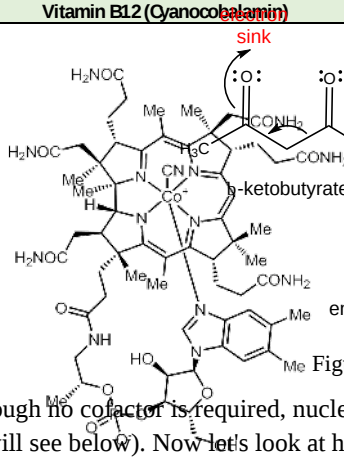
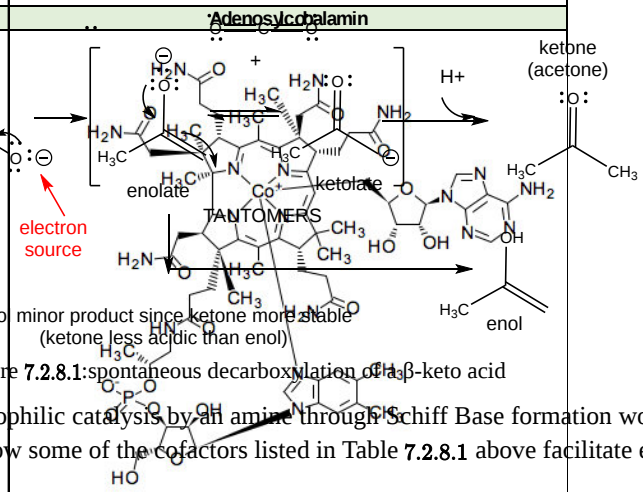
Vitamin B6 (Pyridoxine)	Pyridoxal5'-phosphate
	
Pyridoxamine 5'-phosphate	
	
Vitamin B9 (Folic Acid)	**Tetrahydrofolate**
	
Vitamin B12 (Cyanocobalamin)	**Adenosylcobalamin**
	

Figure 7.2.8.1:spontaneous decarboxylation of a β-keto acid

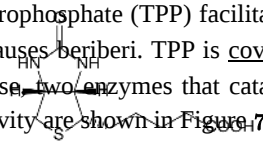
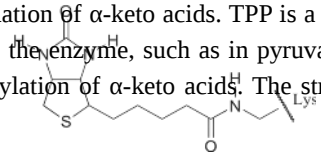
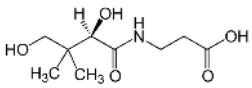
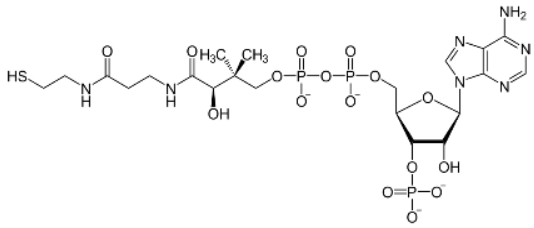
Cofactors can help to mediate enzymatic reactions through the use of any of the different catalytic strategies listed above. They can serve as nucleophiles, mediate covalent catalysis, form electrostatic interactions with the substrate, and stabilize the transition state. They can also cause strain, distortion, or facilitate acid-base catalysis. Metal-aided catalysis can often use homolytic reaction mechanisms that involve radical intermediates. This can be important in reactions such as those occurring in the electron transport chain that require the safe movement of single electrons.

We present plausible mechanisms for prototypical reactions using some of the cofactors shown in Table 7.2.8.1 above. Each shows the flow of electrons from a source to a sink. The source is often a pair of electrons on an anion, formed by the prior removal of a proton from the atom by a general base. A sink could be a carbonyl O, which receives a pair of electrons from one of the C=O bonds of the carbonyl. As a bond is made to the carbonyl, one of the double bonds must break with the electrons going (temporarily if the reaction is a nucleophilic substitution reaction) to the carbonyl O, an excellent sink since it is so electronegative. An even better sink is a positive N of an iminium ion; examples are shown below. Just the "business parts" of the cofactors are shown below.

To appreciate the mechanism used by cofactors and show a clear example of an electron source/sink, let's look at a reaction that doesn't require a cofactor, the spontaneous decarboxylation of a β-keto acid, as shown in Figure 7.2.8.1.

Even though no cofactor is required, nucleophilic catalysis by an amine through Schiff Base formation would speed up the reaction (as we will see below). Now let's look at how some of the cofactors listed in Table 7.2.8.1 above facilitate electron flow in reactions.

7.2.8.2: Thiamine pyrophosphate - decarboxylation of α-keto acids

Biotin	Biotin-Enzyme Complex
	
Pantothenic Acid	**Coenzyme A**
	

Thiamine pyrophosphate (TPP) facilitates the decarboxylation of α-keto acids. TPP is a derivative of thiamine, vitamin B1, whose deficiency causes beriberi. TPP is covalently attached to the enzyme, such as in pyruvate dehydrogenase and alpha-ketoglutarate dehydrogenase, two enzymes that catalyze the decarboxylation of α-keto acids. The structure and "business" end of TPP and its catalytic activity are shown in Figure 7.2.8.2.

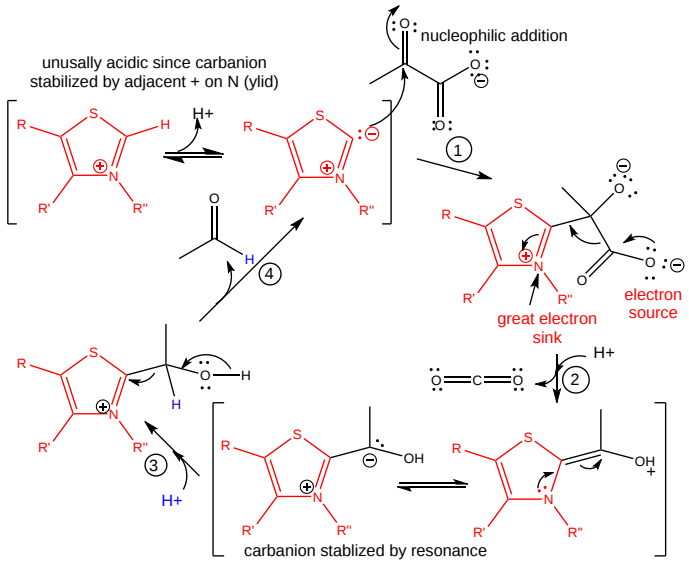


Figure 7.2.8.2: Role of TPP in the decarboxylation of pyruvate (step 1) and release of acetaldehyde (step 4)

The number of arrows leading to the product does not reflect the number of steps.

Figure 7.2.8.3 shows an interactive iCn3D model of the thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae (1pvd).

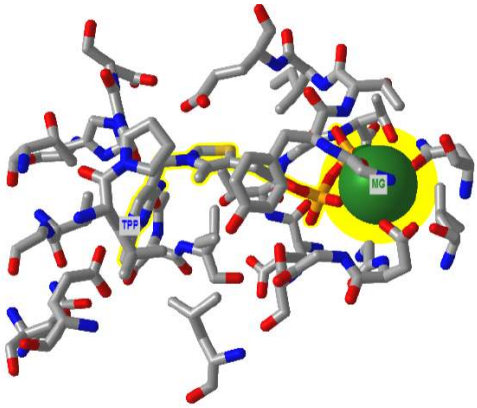


Figure 7.2.8.3: Thiamin diphosphate-dependent enzyme pyruvate decarboxylase from the yeast Saccharomyces cerevisiae (1pvd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...cCq9EJrquwggj8

7.2.8.3: Flavin Adenine Dinucleotide (FAD) - hydride transfer

FAD and its reduced form, $FADH_2$, are tightly or covalently attached to an enzyme, so FAD must be regenerated in each catalytic cycle. Figure 7.2.8.4 shows an example of how this cofactor facilitates the transfer of a $:H^-$ hydride ion to the "business end" of FAD. In contrast to a transfer of protons (H^+), an acid/base reaction, hydride transfer removes 2 electrons from the substrate (in this case, succinate) along with a proton in an oxidation reaction as FAD is reduced.

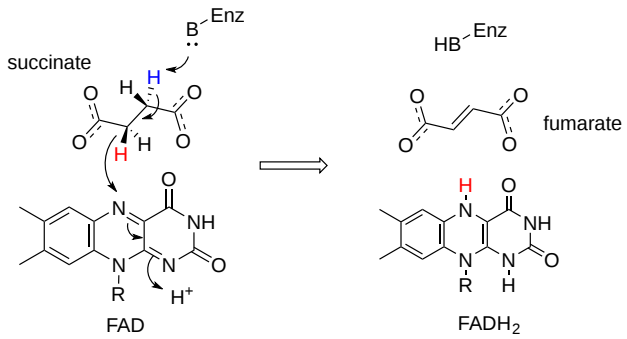


Figure 7.2.8.4: Oxidation of succinate by FAD

Figure 7.2.8.5 shows an interactive iCn3D model of the FAD-binding domain of cytochrome P450 BM3 from Priestia megaterium in complex with $NADP^+$ (4DQL)

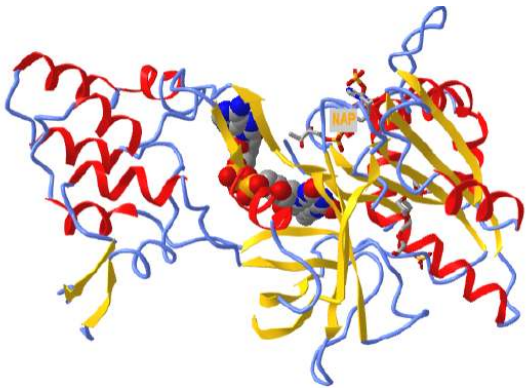


Figure 7.2.8.5: FAD binding domain of cytochrome P450 BM3 in complex with $NADP^+$ (4DQL). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?hoT1WDCUv1wZFMyRA

FAD is shown in spacefill. $NADP^+$, which reoxidizes the reduced $FADH_2$ back to FAD, is shown in sticks and labeled NAP.

7.2.8.4: Nicotinamide Adenine Dinucleotide (FAD) reactions

NAD^+ and a phosphorylated form, $NADP^+$, are one of nature's most widely used oxidizing agents and are used as dissociable substrates/cofactors for many different types of enzyme-catalyzed oxidation reactions. The free enzyme is continually active since the NAD^+ or $NADP^+$ cofactor binds (as a substrate) and dissociates (as a product) after each catalytic cycle. The biological synthesis of NAD^+ requires the vitamin nicotinic acid, also called niacin (nicotinic acid), an absence of which causes pellagra.

Oxidation of an alcohol to an aldehyde: The oxidation of ethanol to acetaldehyde by NAD^+, catalyzed by the enzyme alcohol dehydrogenase, is shown in Figure 7.2.8.6.

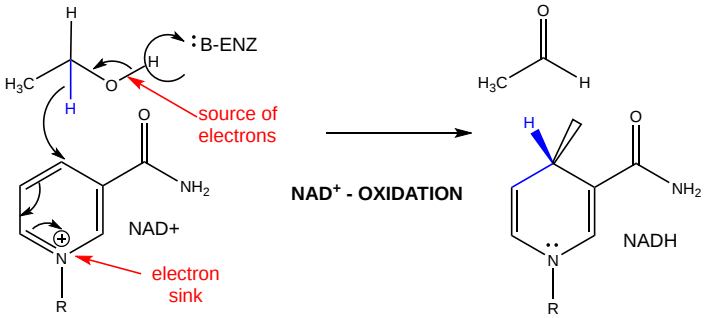


Figure 7.2.8.6: Oxidation of ethanol by NAD^+

The product acetaldehyde contributes to hangovers after ethanol consumption. Note that this reaction is a hydride transfer, which would not be expected to occur in the aqueous environment of a cell, given the extreme reactivity and basicity of a $:H^-$ hydride ion. This transfer happens in the enzyme's active site, which is anhydrous after binding substrates.

Oxidative decarboxylation of an alcohol: A two-step mechanism for this reaction is shown in Figure 7.2.8.7

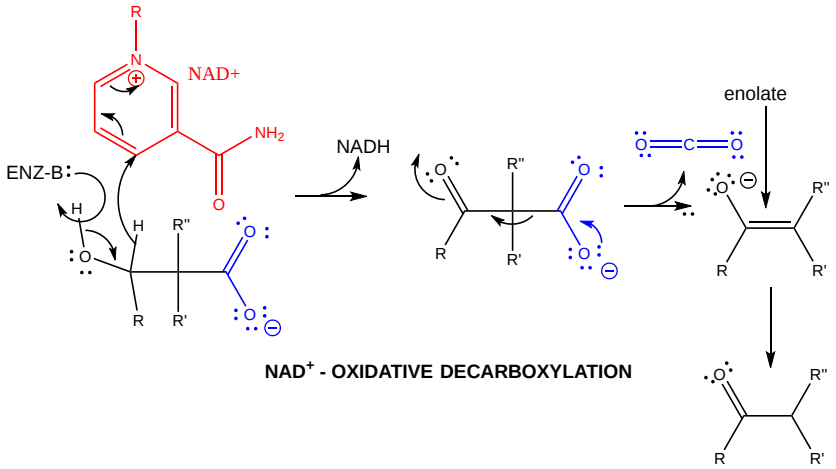


Figure 7.2.8.7: Oxidative decarboxylation of an alcohol

After the first step, an electron sink (the oxygen of the carbonyl) is present at the β-carbon, facilitating the decarboxylation step.

Oxidative deamination of an amine: A two-step reaction, a hydride transfer to form a Schiff base, followed by hydrolysis of the Schiff base, is shown in Figure 7.2.8.8.

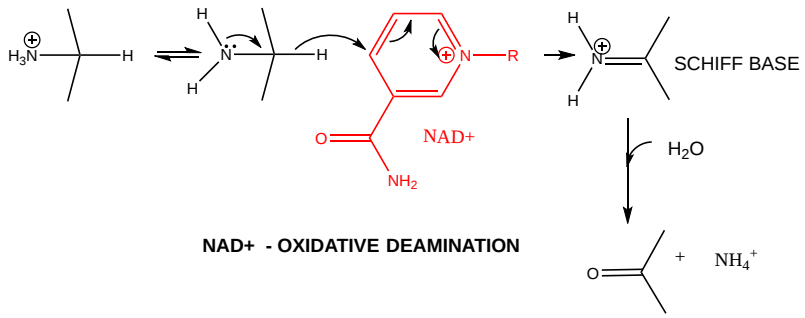


Figure 7.2.8.8: **Oxidative deamination of an amine**

We will discuss Schiff base chemistry in more detail below.

7.2.8.5: Pyridoxal Phosphate Enzymes

Pyridoxal phosphate (PLP) is a derivative of vitamin B6 or pyridoxal. Deficiencies cause convulsions, chronic anemia, and neuropathy. It assists in many reactions (catalyzed by PLP-dependent enzymes). The PLP is bound covalently to lysine residues in a Schiff base linkage (aldimine). This form reacts with many free amino acids (as substrates) to replace the Schiff base to Lys of the enzyme with a Schiff base to the amino acid substrate. First, we will review Schiff base (an imine) formation by the reaction of an aldehyde or ketone with an amine, as shown in Figure 7.2.8.9.

Schiff base (imine) formation

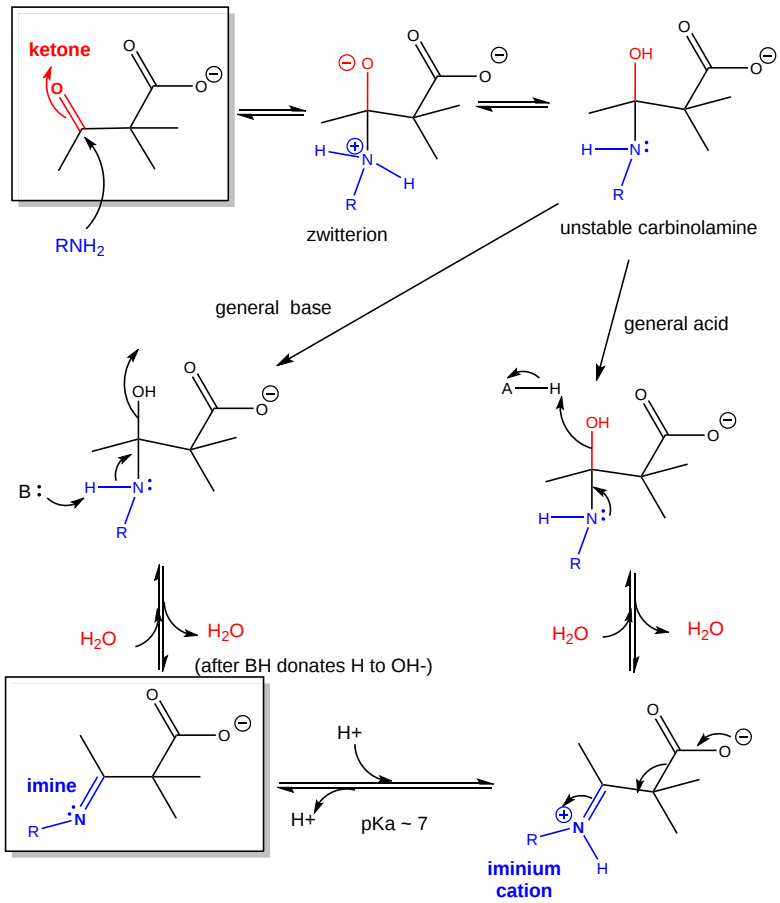


Figure **7.2.8.9**

The reaction is essentially a nucleophilic attack of a carbonyl carbon of an aldehyde or ketone by an amine, followed by a dehydration step. Note that the net effect is to replace one electron sink, a carbonyl (C=O), with an **imine** (C=NH ↔ C=NH_2^+), with pKa around 7.0. Hence, 50% of the imine is protonated at neutral pH to form the **iminium cation**, a much better electron sink than the starting carbonyl!

The structure of pyridox**al** phosphate, which contains a reactive **al**dehyde, is converted to an imine by reaction with the ε-amino side chain of a lysine in the active site of a PLP-dependent enzyme, as shown in Figure **7.2.8.10**.

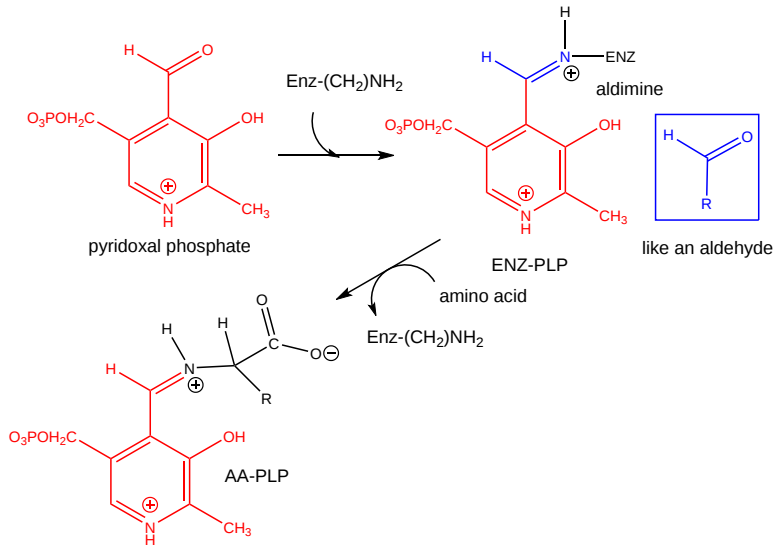


Figure 7.2.8.10

The figure also shows the replacement of the enzymes' lysine ε-NH_2-PLP bond to that of a free amino as an incoming substrate, a process which should proceed with a ΔG^0 of approximately 0. This occurs in PLP-dependent enzymes with free amino acids as substrates (we will discuss several examples below).

Figure 7.2.8.11 shows an interactive iCn3D model of the E. Coli Aspartate aminotransferase, W140H mutant, maleate complex (1ARI).

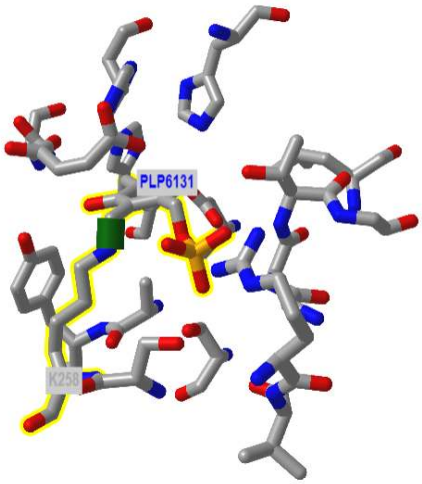


Figure 7.2.8.11: **Aspartate aminotransferase, W140H mutant, maleate complex** (1ARI). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZQixXhCwX5EEQ9

Note that the PLP is in Schiff base linkage with the ε-NH_2 group of a lysine in the enzyme's active site.

PLP is quite impressive!

From a chemistry perspective, PLP is an ideal molecule to facilitate electron flow in biochemical reactions. William Jencks noted this in his classic text, Catalysis in Chemistry, in which he wrote this elegant description of its properties:

"It has been said that God created an organism especially adapted to help the biologist find an answer to every question about the physiology of living systems; if this is so, it must be concluded that pyridoxal phosphate was created to provide satisfaction and enlightenment to those enzymologists and chemists who enjoy pushing electrons, for no other coenzyme is involved in

> such a wide variety of reactions, in both enzyme and model systems, which can be reasonably interpreted in terms of the chemical properties of the coenzyme. Most of these reactions are made possible by a common structural feature. That is, electron withdrawal toward the cationic nitrogen atom of the imine and into the electron sink of the pyridoxal ring from the alpha carbon atom of the attached amino acid activates all three of the substituents of this carbon for reactions that require electron withdrawal from this atom."

We'll present three examples of the reaction of an amino acid with a PLP-dependent enzyme. In each case, a different bond to the α-carbon of the amino acid substrate is broken.

α-decarboxylation of an amino acid: Figure 7.2.8.12 shows a plausible reaction mechanism.

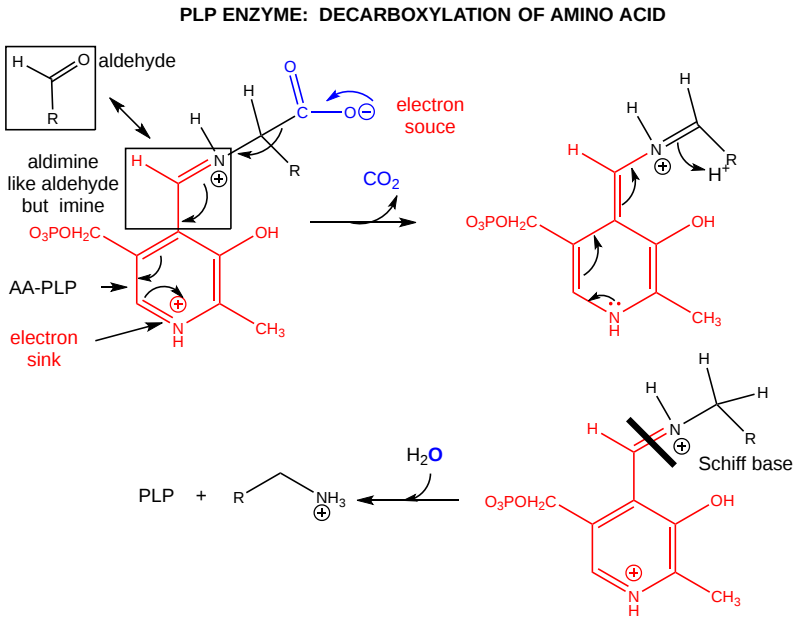


Figure 7.2.8.12: PLP-dependent decarboxylation of an amino acid.

β-elimination from serine: The enzyme serine dehydratase catalyzes the reaction shown in Figure 7.2.8.13.

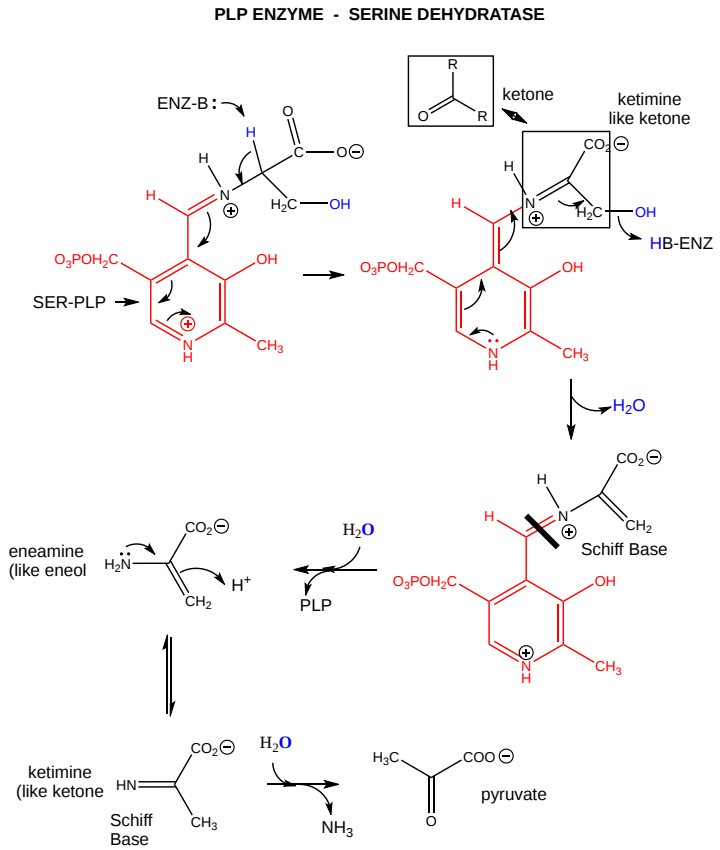


Figure 7.2.8.13: PLP-dependent β-elimination from serine

Racemization of amino acids: Amino acid racemases use PLP as a cofactor, using a mechanism shown in Figure 7.2.8.14.

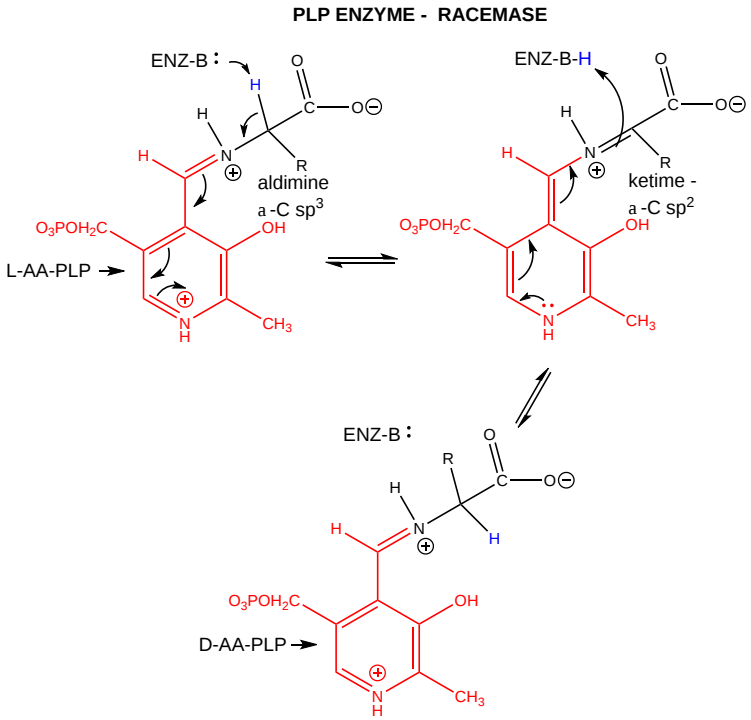


Figure 7.2.8.14: PLP-dependent racemization of amino acids

Why do racemases exist since the biological world consists of only L-amino acids? There are two possible reasons. Some D-amino acids are found, such as in bacterial cell walls. In addition, amino acids spontaneously racemize on their own, albeit slowly.

Racemases with oxygen atoms in the beta-carbon racemize at a much higher rate since they can stabilize the carbanion intermediate formed when the alpha proton is removed during racemization. The concentration of D-Asp and D-Asn can also be used to date biological material.

Transamination reactions: PLP enzymes also catalyze the transamination reaction, which is shown in Figure 7.2.8.15

Amino Acid 1 + α-keto acid 1 ↔ α-keto acid 2 + Amino Acid 2 For example: Figure 7.2.8.x

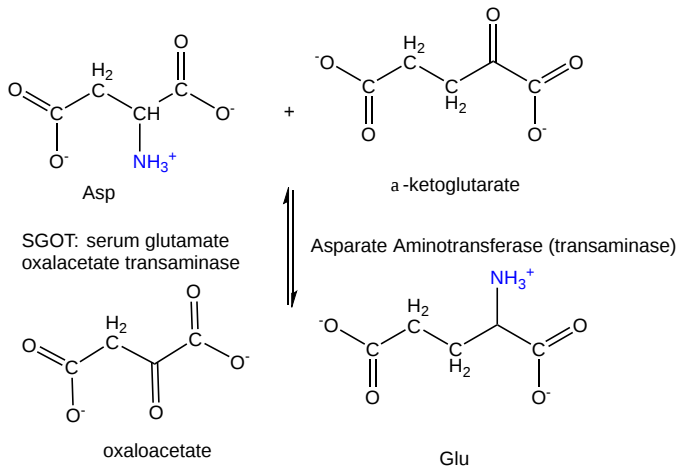


Figure 7.2.8.15: PLP-dependent transamination reaction

First, Asp, bound to PLP through a Schiff base link, loses the α-H and forms a ketimine through a tautomerization reaction, which ultimately hydrolyzes to form the released oxaloacetate and pyridoxamine. The pyridoxamine reacts with α-ketoglutarate in the reverse of the first three reactions to form Glu.

See Chapter 18.2 for a great description of the role of PLP in transamination reactions. Here is a link to more mechanistic details of the reactions of PLP.

We will explore other cofactors in future chapters.

7.2.8.6: Summary

This chapter explores how enzymes harness cofactors to facilitate the movement of electrons—a crucial process in the making and breaking of chemical bonds during catalysis. The discussion begins by establishing the fundamental concept of electron flow, where electrons are transferred from “sources” (electron-rich groups) to “sinks” (electron-deficient centers). This electron-pushing mechanism underpins many enzyme-catalyzed reactions.

A major focus is placed on the two primary classes of cofactors: metal ions and coenzymes. Metal cofactors, such as iron, magnesium, copper, and zinc, are essential for stabilizing charged intermediates and often participate directly in redox reactions. In contrast, coenzymes are small organic molecules—many of which are derived from vitamins—that can bind transiently to enzymes. Depending on their binding characteristics, these coenzymes may act as loosely bound substrates or as prosthetic groups that remain tightly attached throughout the catalytic cycle.

The chapter details several prototypical cofactors and their roles in catalysis:

- **Thiamine Pyrophosphate (TPP):** Derived from vitamin B1, TPP is critical for the decarboxylation of α-keto acids in enzymes such as pyruvate dehydrogenase. Its unique structure allows it to stabilize carbanion intermediates during bond cleavage.
- **Flavin Adenine Dinucleotide (FAD) and Nicotinamide Adenine Dinucleotide (NAD^+/$NADP^+$):** These coenzymes play central roles in oxidation-reduction reactions. FAD facilitates hydride transfer, as seen in the oxidation of succinate, while NAD^+/$NADP^+$ acts as a reversible electron carrier in numerous dehydrogenase reactions.
- **Pyridoxal Phosphate (PLP):** A derivative of vitamin B6, PLP is perhaps the most versatile coenzyme, participating in a range of reactions including decarboxylation, β-elimination, racemization, and transamination. Through the formation of Schiff base intermediates with amino acid substrates, PLP acts as an efficient electron sink, thereby lowering the activation energy for these transformations.

Throughout the chapter, the emphasis is on how cofactors enable precise control over electron movement, ensuring that reactions proceed efficiently under physiological conditions. Structural examples—illustrated with interactive models and diagrams—demonstrate how specific amino acid residues, such as histidines coordinating a heme group, facilitate these interactions.

Ultimately, this chapter underscores the critical role of cofactors in enzyme catalysis. By providing both the chemical functionality and the structural framework for electron transfer, cofactors bridge the gap between the inherent reactivity of small molecules and the exquisite specificity required for biological regulation. This understanding is fundamental for appreciating how enzymes achieve their remarkable catalytic efficiencies and for designing therapeutic strategies that target these processes.

This page titled 7.2.8: Cofactors and Catalysis - A Little Help From My Friends is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **6.8: Cofactors and Catalysis - A Little Help From My Friends** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

7.3: Basic Principles of Catalysis

A printable version of this section is here: BiochemFFA_4_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

If there is a magical component to life, an argument can surely be made for it being catalysis. Thanks to catalysis, reactions that can take hundreds of years to complete in the uncatalyzed "real world," occur in seconds in the presence of a catalyst. Chemical catalysts, such as platinum, can speed reactions, but enzymes (which are simply super-catalysts with a "twist," as we shall see) put chemical catalysts to shame (Figure 4.1). To understand enzymatic catalysis, it is necessary first to understand energy. Chemical reactions follow the universal trend of moving towards lower energy, but they often have a barrier in place that must be overcome. The secret to catalytic action is reducing the magnitude of that barrier.

Enzyme	Nonenzymatic Half-Life	Uncatalyzed Rate (k_{un} s^{-1})	Catalyzed Rate (k_{cat} s^{-1})	Rate Enhancement (k_{cat} s^{-1}/k_{un} s^{-1})
OMP decarboxylase	78,000,000 years	28×10^{-16}	39	1.4×10^{17}
Staphylococcal nuclease	130,000 years	1.7×10^{-13}	95	5.6×10^{14}
Carboxypeptidase A	7.3 years	3.0×10^{-9}	578	1.9×10^{11}
Ketosteroid isomerase	7 weeks	1.7×10^{-7}	66,000	3.9×10^{11}
Triose phosphate isomerase	1.9 days	4.3×10^{-6}	4,300	1.0×10^{9}
Chorismate mutase	7.4 hours	2.6×10^{-5}	50	1.9×10^{6}
Carbonic anhydrase	5 seconds	1.3×10^{-1}	1×10^{6}	7.7×10^{6}

Abbreviations:
OMP - Orotine monophosphate
AMP - Adenosine monophosphate

Figure 4.1 - Rate enhancement for several enzymes Image by Aleia Kim

Before discussing enzymes, it is appropriate to pause and discuss an important concept relating to chemical/biochemical reactions. That concept is *equilibrium* and it is very often misunderstood. The "equi" part of the word relates to equal, as one might expect, but it does not relate to absolute concentrations. What happens when a biochemical reaction is at equilibrium is that the concentrations of reactants and products do not change over time. This does not mean that the reactions have stopped. Remember that reactions are reversible, so there is a forward reaction and a reverse reaction: if you had 8 molecules of A, and 4 of B at the beginning, and 2 molecules of A were converted to B, while 2 molecules of B were simultaneously converted back to A, the number of molecules of A and B remain unchanged, i.e., the reaction is at equilibrium. However, you will notice that this does not mean that there are equal numbers of A and B molecules.

Concentration Matters

So, contrary to the perceptions of many students, the concentrations of products and reactants are not equal at equilibrium, unless the ΔG°' for a reaction is zero, because when this is the case,

$$\Delta G = \ln\left(\frac{[\text{Products}]}{[\text{Reactants}]}\right) \tag{7.3.1}$$

since the ΔG°' is zero. Because ΔG itself is zero at equilibrium, then

$$[Products] = [Reactants]. \tag{7.3.2}$$

This is the only circumstance where

$$[Products] = [Reactants] \tag{7.3.3}$$

at equilibrium. Reiterating, at equilibrium, the concentrations of reactant and product do not change over time. That is, for a reaction $A \rightleftharpoons B[A]$ at time zero when equilibrium is reached, $[A]_{T_0}$, will be the same 5 minutes later (assuming A and B are chemically stable). Thus,

$$[A]_{T_0} = [A]_{T+5} \tag{7.3.4}$$

Similarly,

$$[B]_{T_0} = [B]_{T+5} \tag{7.3.5}$$

For that matter, at any amount of time X after equilibrium has been reached,

$$[A]T0 = [A]T + 5 = [A]TX \tag{7.3.6}$$

and

$$[B]T0 = [B]T + 5 = [B]TX \tag{7.3.7}$$

However, unless ΔG°' = 0, it is wrong to say [A]T0 = [B]T0 As we study biochemical reactions and reaction rates, it is important to remember that 1) reactions do not generally start at equilibrium; 2) all reactions move in the direction of equilibrium; and 3) reactions in cells behave just like those in test tubes - they do not begin at equilibrium, but they move towards it. Dynamic reactions The reactions occurring in cells, though, are very dynamic and complex. In a test tube, they can be studied one at a time. In cells, the product of one reaction is often the substrate for another one. Reactions in cells are interconnected in this way, giving rise to what are called metabolic pathways. There are, in fact, thousands of different interconnected reactions going on continuously in cells.

Attempts to study a single reaction in the chaos of a cell is daunting to say the least. For this reason, biochemists isolate enzymes from cells and study reactions individually. It is with this in mind that we begin our consideration of the phenomenon of catalysis by describing, first, the way in which enzymes work.

Activation energy

Figure 4.2 schematically depicts the energy changes that occur during the progression of a simple reaction. In order for the reaction to proceed, an activation energy must be overcome in order for the reaction to occur.

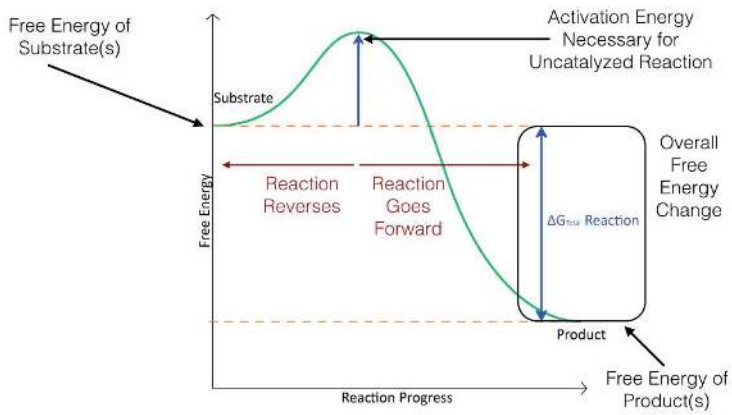


Figure 4.2 - Energy changes during the course of an uncatalyzed reaction. Image by Aleia Kim

In Figure 4.3, the activation energy for a catalyzed reaction is overlaid. As you can see, the reactants start at the same energy level for both catalyzed and uncatalyzed reactions and that the products end at the same energy for both as well. The catalyzed reaction, however, has a lower energy of activation (dotted line) than the uncatalyzed reaction. This is the secret to catalysis - overall ΔG for a reaction does NOT change with catalysis, but the activation energy is lowered.

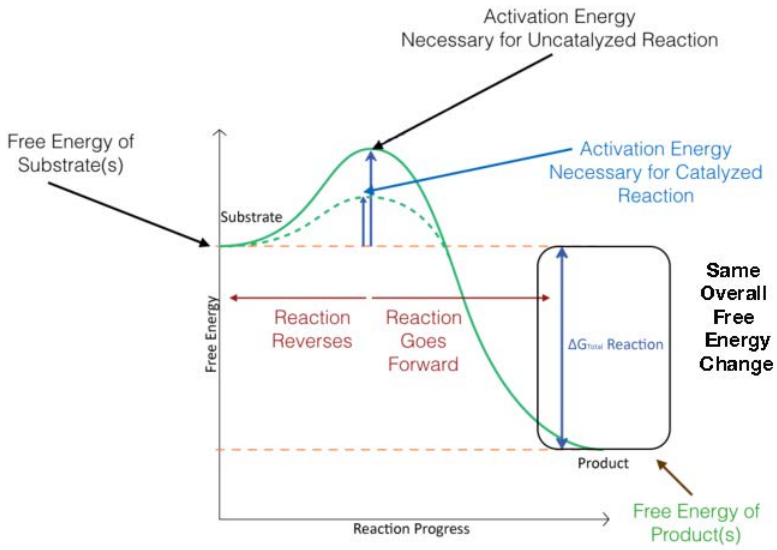


Figure 4.3 - Energy changes during the course of an uncatalyzed reaction (solid green line) and a catalyzed reaction (dotted green line). Image by Aleia Kim

Reversibility

The extent to which reactions will proceed forward is a function of the size of the energy difference between the product and reactant states. The lower the energy of the products compared to the reactants, the larger the percentage of molecules that will be present as products at equilibrium. It is worth noting that since an enzyme lowers the activation energy for a reaction that it can speed the reversal of a reaction just as it speeds a reaction in the forward direction. At equilibrium, of course, no change in concentration of reactants and products occurs. Thus, enzymes speed the time required to reach equilibrium, but do not affect the balance of products and reactants at equilibrium.

Exceptions

The reversibility of enzymatic reactions is an important consideration for equilibrium, the measurement of enzyme kinetics, for Gibbs free energy, for metabolic pathways, and for physiology. There are some minor exceptions to the reversibility of reactions, though. They are related to the disappearance of a substrate or product of a reaction. Consider the first reaction below which is catalyzed by the enzyme carbonic anhydrase:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^- + H^+ \tag{7.3.8}$$

In the forward direction, carbonic acid is produced from water and carbon dioxide. It can either remain intact in the solution or ionize to produce bicarbonate ion and a proton. In the reverse direction, water and carbon dioxide are produced. Carbon dioxide, of course, is a gas and can leave the solution and escape.

When reaction molecules are removed, as they would be if carbon dioxide escaped, the reaction is pulled in the direction of the molecule being lost and reversal cannot occur unless the missing molecule is replaced. In the second reaction occurring on the right, carbonic acid (H2CO3) is “removed” by ionization. This too would limit the reaction going back to carbon dioxide in water. This last type of “removal” is what occurs in metabolic pathways. In this case, the product of one reaction (carbonic acid) is the substrate for the next (formation of bicarbonate and a proton).

In the metabolic pathway of glycolysis, ten reactions are connected in this manner and reversing the process is much more complicated than if just one reaction was being considered.

General mechanisms of action

As noted above, enzymatically catalyzed reactions are orders of magnitude faster than uncatalyzed and chemical-catalyzed reactions. The secret of their success lies in a fundamental difference in their mechanisms of action.

Every chemistry student has been taught that a catalyst speeds a reaction without being consumed by it. In other words, the catalyst ends up after a reaction just the way it started so it can catalyze other reactions, as well. Enzymes share this property, but in the middle, during the catalytic action, an enzyme is transiently changed. In fact, it is the ability of an enzyme to change that leads to its incredible efficiency as a catalyst.

Changes

These changes may be subtle electronic ones, more significant covalent modifications, or structural changes arising from the flexibility inherent in enzymes, but not present in chemical catalysts. Flexibility allows movement and movement facilitates alteration of electronic environments necessary for catalysis. Enzymes are, thus, much more efficient than rigid chemical catalysts as a result of their abilities to facilitate the changes necessary to optimize the catalytic process.

Substrate binding

Another important difference between the mechanism of action of an enzyme and a chemical catalyst is that an enzyme has binding sites that not only ‘grab’ the substrate (molecule involved in the reaction being catalyzed), but also place it in a position to be electronically induced to react, either within itself or with another substrate.

The enzyme itself may play a role in the electronic induction or the induction may occur as a result of substrates being placed in very close proximity to each other. Chemical catalysts have no such ability to bind substrates and are dependent upon them colliding in the right orientation at or near their surfaces.

Active site

Reactions in an enzyme are catalyzed at a specific location within it known as the ‘active site’. Substrates bind at the active site and are oriented to provide access for the relevant portion of the molecule to the electronic environment of the enzyme where catalysis occurs.

Enzyme flexibility

As mentioned earlier, a difference between an enzyme and a chemical catalyst is that an enzyme is flexible. Its slight changes in shape (often arising from the binding of the substrate itself) help to optimally position substrates for reaction after they bind.

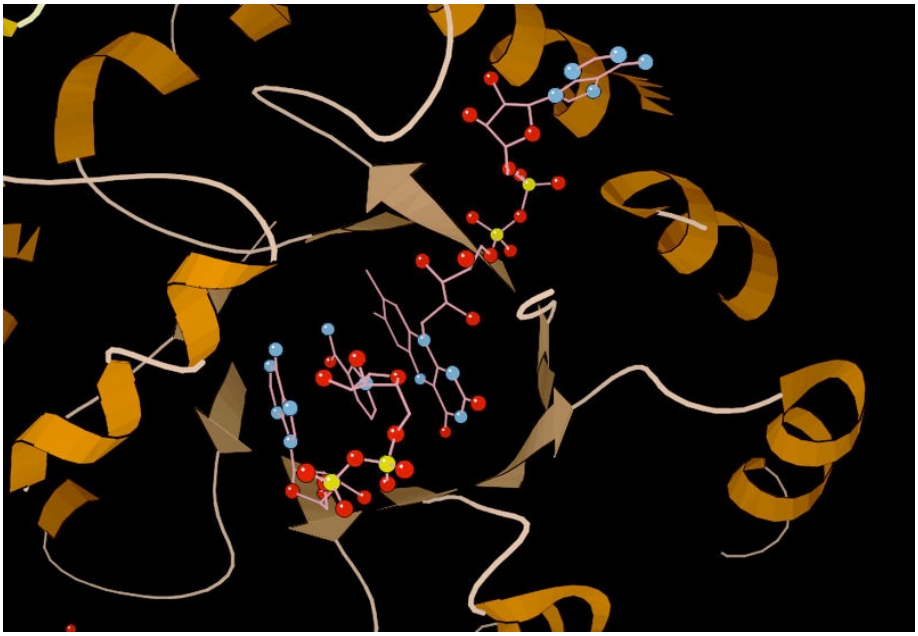


Figure 4.4 - Substrate binding by methylenetetrahydrofolate reductase

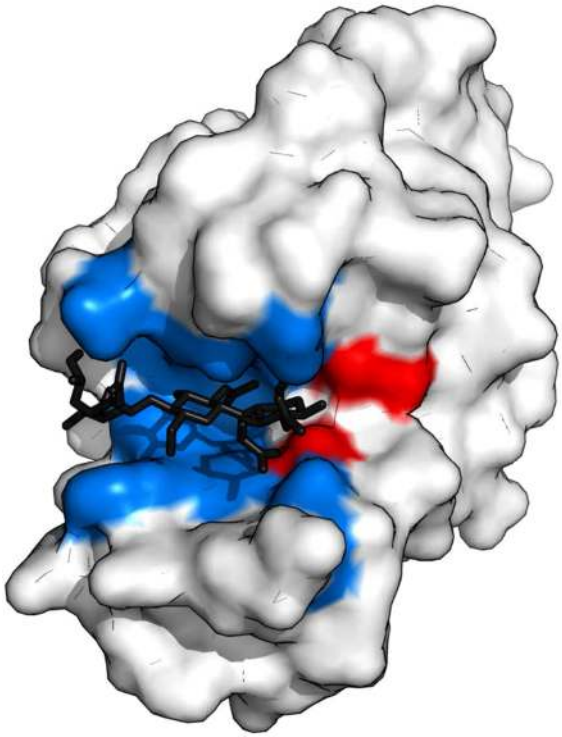


Figure 4.5 - Lysozyme with substrate binding site (blue), active site (red) and bound substrate (black). Wikipedia

Induced fit

These changes in shape are explained, in part, by Koshland's Induced Fit Model of Catalysis (Figure 4.6), which illustrates that not only do enzymes change substrates, but that substrates also transiently change enzyme structure. At the end of the catalysis, the enzyme is returned to its original state. Koshland's model is in contrast to the Fischer Lock and Key model, which says simply that an enzyme has a fixed shape that is perfectly matched for binding its substrate(s). Enzyme flexibility also is important for control of enzyme activity. Enzymes alternate between the T (tight) state, which is a lower activity state and the R (relaxed) state, which has greater activity.

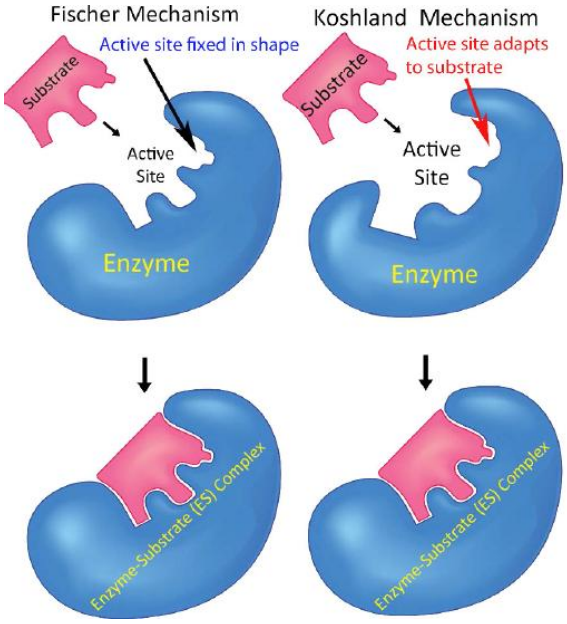


Figure 4.6 - Fischer's lock and key model (left) Vs. Koshland's induced fit model (right). Image by Aleia Kim

Induced Fit

The Koshland Induced Fit model of catalysis postulates that enzymes are flexible and change shape on binding substrate. Changes in shape help to 1) aid binding of additional substrates in reactions involving more than one substrate and/or 2) facilitate formation of an electronic environment in the enzyme that favors catalysis. This model is in contrast to the Fischer Lock and Key Model of catalysis which considers enzymes as having pre-formed substrate binding sites.

Ordered binding

The Koshland model is consistent with multi-substrate binding enzymes that exhibit ordered binding of substrates. For these systems, binding of the first substrate induces structural changes in the enzyme necessary for binding the second substrate.

There is considerable experimental evidence supporting the Koshland model. Hexokinase, for example, is one of many enzymes known to undergo significant structural alteration after binding of substrate. In this case, the two substrates are brought into very close proximity by the induced fit and catalysis is made possible as a result.

Reaction types

Enzymes that catalyze reactions involving more than one substrate, such as

$$A + B \rightleftharpoons C + D \tag{7.3.9}$$

can act in two different ways. Enzymatic reactions can be of several types, as shown in Figure 4.7. In one mechanism, called sequential reactions, at some point in the reaction, both substrates will be bound to the enzyme. There are, in turn, two different ways in which this can occur - random and ordered.

Figure 4.7 - Categories of enzymatic reactions

Types of Reactions	
Single Substrate - Single Product	A ⇄ B
Single Substrate - Multiple Products	A ⇄ B + C
Multiple Substrates - Single Products	A + B ⇄ C
Multiple Substrates - Multiple Products	A + B ⇄ C + D

Consider lactate dehydrogenase, which catalyzes the reaction below:

$$NADH + Pyruvate \rightarrow Lactate + NAD^{+} \tag{7.3.10}$$

This enzyme requires that NADH must bind prior to the binding of pyruvate. As noted earlier, this is consistent with an induced fit model of catalysis. In this case, binding of the NADH changes the enzyme shape/environment so that pyruvate can bind and without binding of NADH, the substrate cannot access the pyruvate binding site. This type of multiple substrate reaction is called **sequential ordered binding**, because the binding of substrates must occur in the right order for the reaction to proceed.

Random binding

A second mechanism of binding/catalysis is exhibited by creatine kinase which catalyzes the following reaction:

$$Creatine + ATP \rightarrow \text{Creatine phosphate} + ADP \tag{7.3.11}$$

For this enzyme, substrates can bind to it in any order. Creatine kinase displays sequential random binding. It is worth mentioning that random binding is not inconsistent with Koshland's induced fit model. Rather, random binding simply means that the enzyme's induced fit doesn't affect substrate binding sites and involves other parts of the enzyme. In summary, sequential binding can occur as ordered binding or as random binding.

Double displacement reactions

Not all enzymes that catalyze multi-substrate reactions, though, bind A and B by the sequential mechanisms above. This other category of enzyme includes those that exhibit what are called "ping-pong" (or double displacement) mechanisms. In these

enzymes, the enzyme functions as both a catalyst and a carrier of a group between individually bound substrates. Examples of this type of enzyme include the class of enzymes known as transaminases. A general form of the reactions catalyzed by these enzymes is shown in Figure 4.8.

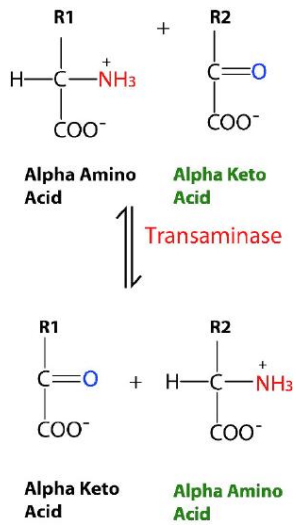


Figure 4.8 - Double displacement reaction mechanism of a transaminase

In reversible transaminase reactions, an oxygen and an amine are swapping between the molecules. It can be represented as follows (where N is the amine and O is the oxygen).

A=O + C=N ⇄ B-N + D=O

This reaction occurs not as one transfer reaction swapping the N and the O, but rather as a set of two half-reactions. In this case, the enzyme is a both donor and a carrier of the group being swapped. The first half-reaction goes as follows

A=O + Enz-N ⇄ B-N + Enz=O

Next a second half-reaction goes as

C-N + Enz=O ⇄ D=O + Enz=N

The sum of these half-reactions then is

A=O + C=N ⇄ B-N + D=O

Note that at no time did the enzyme bind both A and C simultaneously. It is also important to recognize that the enzyme existed in two states - Enz=O and Enz-N. The shuffling of the enzyme between these two states is what gives rise to the ping-pong name of this mechanism - it literally goes back and forth like a ping-pong ball in a table tennis match.

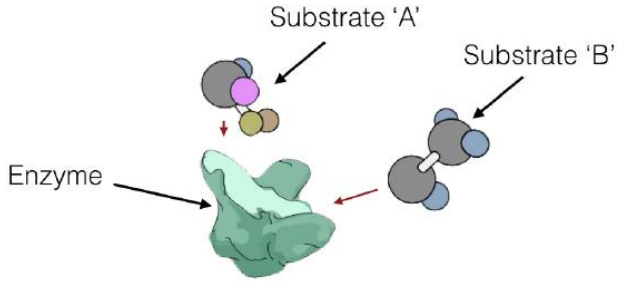


Figure 4.9 - Binding of substrates (A+B) to free enzyme

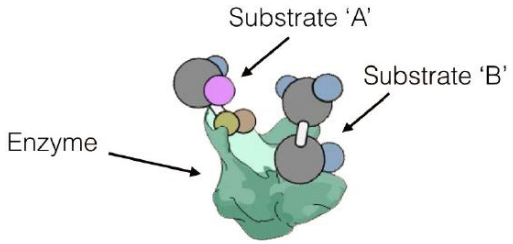


Figure 4.10 ES complex formation

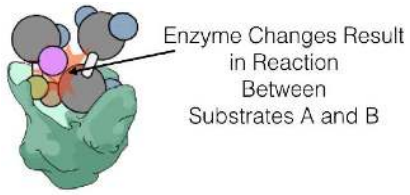


Figure 4.11 - ES* complex - The reaction occurs.

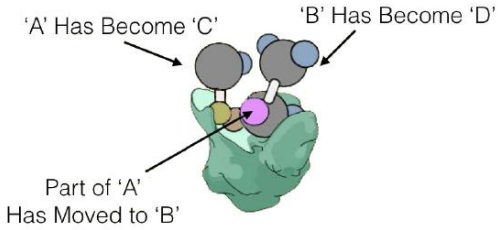


Figure 4.12 - EP complex - the reaction is complete.

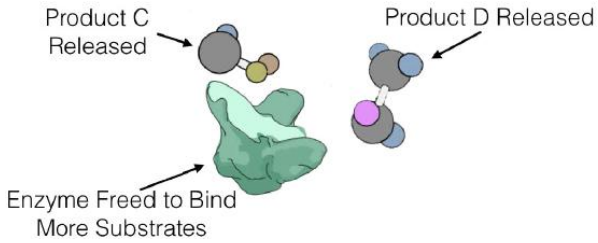


Figure 4.13 - E + P - the products are released

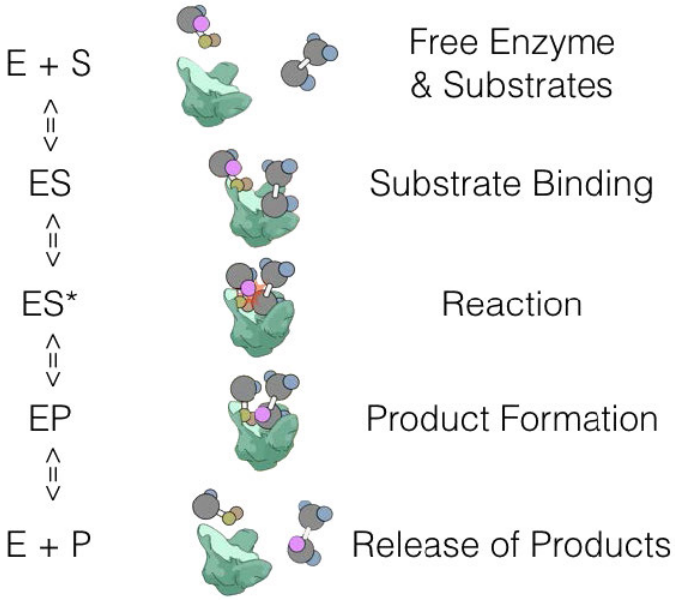


Figure 4.14 - Summary

Enzyme kinetics

To understand how an enzyme enhances the rate of a reaction, we must understand enzyme kinetics. We present a model here proposed by Leonor Michaelis and Maud Menten. In order to understand the model, it is necessary to understand a few parameters.

First, we describe a reaction in simple terms proceeding as follows

E + S ⇄ ES -> E + P

where E is enzyme, S is substrate, and P is product. In this scheme, ES is the Enzyme-Substrate complex, which is simply the enzyme bound to its substrate.

We could define the ES state a bit further with

E + S ⇄ ES -> ES* -> EP -> E + P

where ES* is the activated state and EP is the enzyme-product complex before release of the product.

The first consideration we have is velocity. The velocity of a reaction is the rate of creation of product over time, measured as the concentration of product per time. The time is a critical consideration when measuring velocity. In a closed system (in which an enzyme operates), all reactions will advance towards equilibrium. Enzymatically catalyzed reactions are no different in the end result from non-enzymatic reactions, except that they get to equilibrium faster.

Equilibrium

At equilibrium, the ratio of product to reactant does not change. That is a property of equilibrium. Since the system is closed, the concentration of product over time will not change. The velocity will thus be zero under these conditions and we will have learned nothing about the reaction if we wait too long to study it.

Velocity

Enzyme	Turnover Number (per second)
Carbonic anhydrase	600,000
3-Ketoesteroid isomerase	280,000
Acetylcholinesterase	25,000
Penicillinase	2,000
Lactate dehydrogenase	1,000
Chymostrypsin	100
DNA Polymerase I	15
Tryptophan synthetase	2
Lysozyme	0.5

Figure 4.22 - Kcat/Km values for perfect enzymes. Image by Aleia Kim

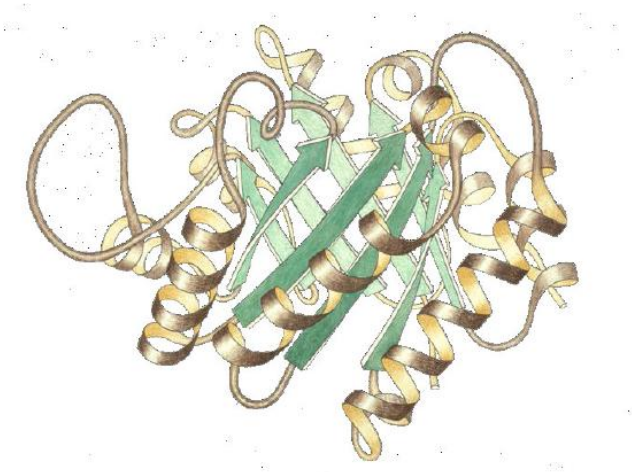


Figure 4.23 - Triose phosphate isomerase - A perfect enzyme. Wikipedia

Consequently, in Michaelis-Menten kinetics, velocity is measured as initial velocity (V0). This is accomplished by measuring the rate of formation of product early in the reaction before equilibrium is established and under these conditions, there is very little if any of the reverse reaction occurring.

The other two assumptions are related. First, we use conditions where there is much more substrate than enzyme. This makes sense. If the substrate is not in great excess, then the enzyme's conversion of substrate to product will occur much faster than the enzyme can bind substrate.

Waiting for substrate

Thus, the enzyme would "wait" for substrate to bind if there were not sufficient amounts of it to bind to the enzyme in a timely fashion (when substrate concentration is low). This would not give an accurate measure of velocity, since the enzyme would be inactive a good deal of the time. Because of this, we assume saturation of the enzyme with substrate will give a maximal velocity of the reaction.

Steady state

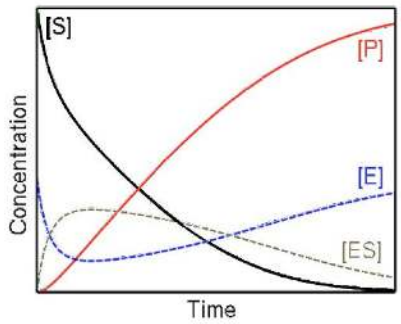


Figure 4.15 - Concentration of product (P), substrate (S), enzyme (E), and enzyme-substrate complex (ES) versus time for an enzymatic reaction

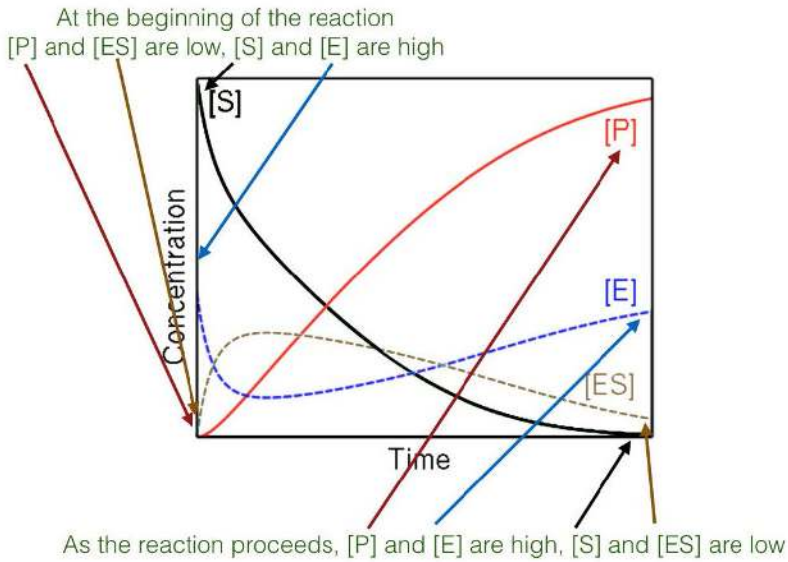


Figure 4.16 - Change in concentration of reaction materials over time

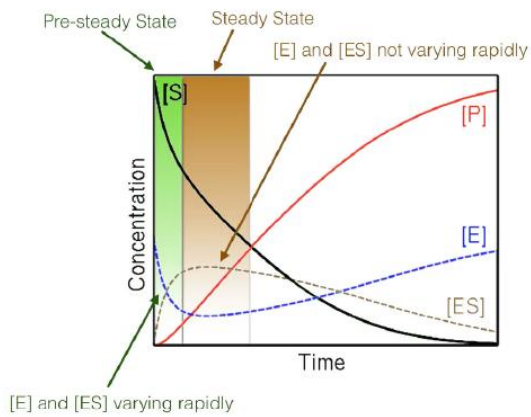


Figure 4.17 - Steady state versus non-steady state conditions

Last, the high concentration of substrate combined with measuring initial conditions results in studying reactions that are under so-called steady state conditions (Figure 4.17). When steady state occurs, the concentration of the ES complex over time is not significantly changing during the period of analysis.

Reiterating, the three assumptions for Michaelis-Menten kinetics are

1. Measurement of initial velocity of a reaction
2. Substrate in great excess compared to enzyme
3. Reaction conditions occurring under steady state

Experimental considerations

Now we turn our attention to how studies of the kinetic properties of an enzyme are conducted. To perform an analysis, one would do the following experiment - 20 different tubes would be set up with enzyme buffer (to keep the enzyme stable), the same amount of enzyme, and then a different amount of substrate in each tube, ranging from tiny amounts in the first tubes to very large amounts in the last tubes. The reaction would be allowed to proceed for a fixed, short amount of time and then the reaction would be stopped and the amount of product contained in each tube would be determined.

The initial velocity (V0) of the reaction then would be the concentration of product found in each tube divided by the time that the reaction was allowed to run. Data from the experiment would be plotted on a graph using initial velocity (V0) on the Y-axis and the concentration of substrate on the X-axis, each tube, of course having a unique reaction velocity corresponding to a unique substrate concentration.

For an enzyme following Michaelis-Menten kinetics, a curve like that shown in Figure 4.18 or 4.19 would result. At low concentration of substrate, it is limiting and the enzyme converts it into product as soon as it can bind it. Consequently, at low concentrations of substrate, the rate of increase of [P] is almost linear with [S] (Figure 4.19).

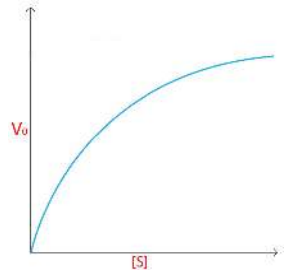


Figure 4.18 - Kinetics of an enzyme obeying Michaelis-Menten kinetics. Image by Aleia Kim

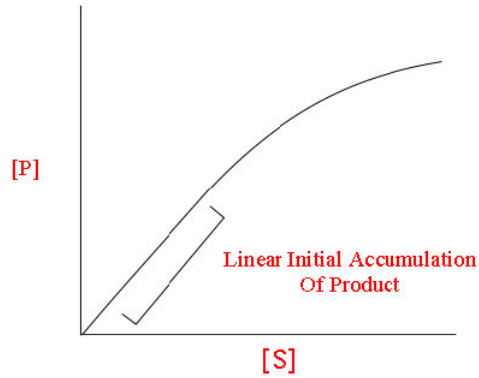


Figure 4.19 - Linear relationship between [P] and [S] at low [S]

Non-linear increase

As the substrate concentration increases, however, the velocity of the reaction in tubes with higher substrate concentration ceases to increase linearly and instead begins to flatten out, indicating that as the substrate concentration gets higher and higher, the enzyme has a harder time keeping up to convert the substrate to product.

Saturation

Not surprisingly, when the enzyme becomes completely saturated with substrate, it will not have to wait for substrate to diffuse to it and will therefore be operating at maximum velocity.

For an enzyme following Michaelis-Menten kinetics will have its velocity (v) at any given substrate concentration given by the following equation:

$$v = \frac{V_{\max}[S]}{K_m + [S]}$$

V_{max}

Two terms in the equation above require explanation. The first is V_{max}. It refers to the maximum velocity of an enzymatic reaction. Maximum velocity for a reaction occurs when an enzyme is saturated with substrate. Saturation is important because it means (per the assumption above) that none of the enzyme molecules are “waiting” for substrate after a product is released. Saturation ensures that another substrate is always instantly available. The unit of V_{max} is concentration of product per time = [P]/time.

On a plot of initial velocity versus substrate concentration (V0 vs. [S]), V_{max} is the value on the Y axis that the curve asymptotically approaches (dotted line in Figure 4.20). It should be noted that the value of V_{max} depends on the amount of enzyme used in a reaction. If you double the amount of enzyme used, you will double the V_{max}. If one wanted to compare the velocities of two different enzymes, it would be necessary to use the same amounts of enzyme in the reaction each one catalyzes.

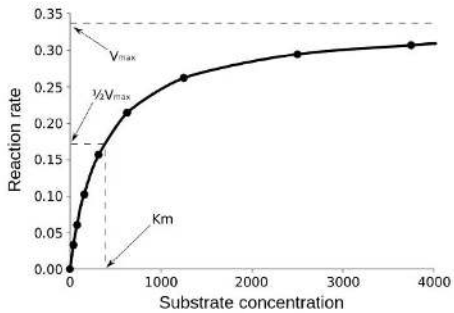


Figure 4.20 - V_{max}, $V_{max}/2$, and Km

Km

The second term is Km (also known as Ks). Referred to as the Michaelis constant, Km is the substrate concentration that causes the enzyme to work at half of maximum velocity (Vmax/2). What it measures, in simple terms, is the affinity an enzyme has for its substrate. The value of Km is inversely related to the affinity of the enzyme for its substrate. Enzymes with a high Km value will have a lower affinity for their substrate (will take more substrate to get to Vmax/2) whereas those with a low Km will have high affinity and take less substrate to get to Vmax/2. The unit of Km is concentration.

Affinities of enzymes for substrates vary considerably, so knowing Km helps us to understand how well an enzyme is suited to the substrate being used. Measurement of Km depends on the measurement of Vmax.

Common mistake

A common mistake students make in describing Vmax is saying that Km = Vmax /2. This is, of course, not true. Km is a substrate concentration and is the amount of substrate it takes for an enzyme to reach Vmax /2. On the other hand Vmax /2 is a velocity and a velocity certainly cannot equal a concentration.

Kcat

It is desirable to have a measure of velocity that is independent of enzyme concentration. Remember that Vmax depended on the amount of enzyme used. For this, we use the Kcat, also known as the turnover number. Kcat is a number that requires one to first determine Vmax for an enzyme and then divide the Vmax by the concentration of enzyme used to determine Vmax. Thus,

Kcat = Vmax /[Enzyme]

Since Vmax has units of concentration per time and [Enzyme] has units of concentration, the units on Kcat are time-1. While that might seem unintuitive, it means that the value of Kcat is the number of molecules of product made by each molecule of enzyme in the time given. So, a Kcat value of 1000/sec means each enzyme molecule in the reaction at Vmax is producing 1000 molecules of product per second. Note that since Kcat is a calculated value, it cannot be read from a V vs [S] graph as Vmax and Km can.

Enzyme	Turnover Number (per second)
Carbonic anhydrase	600,000
3-Ketoesteroid isomerase	280,000
Acetylcholinesterase	25,000
Penicillinase	2,000
Lactate dehydrogenase	1,000
Chymostrypsin	100
DNA Polymerase I	15
Tryptophan synthetase	2
Lysozyme	0.5

Figure 4.21 - Kcat (turnover number) values for several enzymes. Image by Aleia Kim

Amazing Kcat values

A Kcat value of 1000 molecules of product per enzyme per second might seem like a high value, but there are enzymes known (carbonic anhydrase, for example) that have a Kcat value of over 600,000/second (Figure 4.21). This astonishing value illustrates clearly why enzymes seem almost magical in their action. In contrast to V_{max}, which varies with the amount of enzyme used, Kcat is a constant for an enzyme under given conditions.

As seen earlier, enzymes that follow Michaelis-Menten kinetics produce hyperbolic plots of Velocity (V0) versus Substrate Concentration [S] (Figure 4.18). Not all enzymes, though, follow Michaelis-Menten kinetics. Many enzymes have multiple protein subunits and these sometimes interact differently upon binding of a substrate or an external molecule. See ATCase (HERE) for an example.

Perfect enzymes

Now, if we think about what an ideal enzyme might be, it would be one that has a very high velocity and a very high affinity for its substrate. That is, it wouldn't take much substrate to get to Vmax/2 and the Kcat would be very high. Such enzymes would have values of Kcat / Km that are maximum. Interestingly, there are several enzymes that have this property and their maximal Kcat / Km values are all approximately the same. Such enzymes are referred to as being "perfect" because they have reached the maximum possible value.

Enzyme	K_{Cat} / K_M ($s^{-1}M^{-1}$)
Acetylcholinesterase	1.6×10^8
Carbonic anhydrase	8.3×10^7
Catalase	4.0×10^7
Crotonase	2.8×10^8
Fumarase	1.6×10^8
Triose phosphate isomerase	2.4×10^8
β-Lactamase	1.0×10^8
Superoxide dismutase	7.0×10^9

Figure 4.22 - Kcat/Km values for perfect enzymes. Image by Aleia Kim

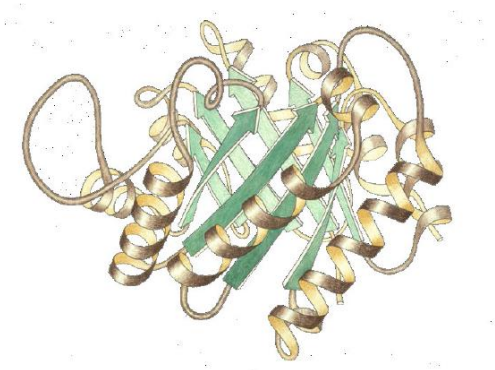


Figure 4.23 - Triose phosphate isomerase - A perfect enzyme. Wikipedia

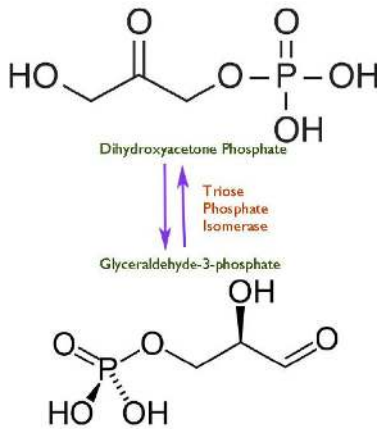


Figure 4.24 - Triose phosphate isomerase-catalyzed reaction

Diffusion limitation

Why should there be a maximum possible value of Kcat / Km? The answer is that movement of substrate to the enzyme becomes the limiting factor for perfect enzymes. Movement of substrate by diffusion in water has a fixed rate at any temperature and that limitation ultimately determines the maximum speed an enzyme can catalyze at. In a macroscopic world analogy, factories can't make products faster than suppliers can deliver materials. It is safe to say for a perfect enzyme that the only speed limit it has is the rate of substrate diffusion in water.

Given the "magic" of enzymes alluded to earlier, it might seem that all enzymes should have evolved to be "perfect." There are very good reasons why most of them have not.

Speed

Speed is a dangerous thing. The faster a reaction proceeds in catalysis by an enzyme, the harder it is to control. As we all know from learning to drive, speeding causes accidents. Just as drivers need to have speed limits for operating automobiles, so too must cells exert some control on the 'throttle' of their enzymes. In view of this, one might wonder then why any cells have evolved any enzymes to perfection. There is no single answer to the question, but a common one is illustrated by triose phosphate isomerase, which catalyzes a reaction in glycolysis shown in Figure 4.24.

The enzyme appears to have evolved this ability because at lower velocities, there is breakdown of an unstable enediol intermediate that then readily forms methyl glyoxal, a cytotoxic compound (Figure 4.25). Speeding up the reaction provides less opportunity for the unstable intermediate to accumulate and fewer undesirable byproducts to be made.

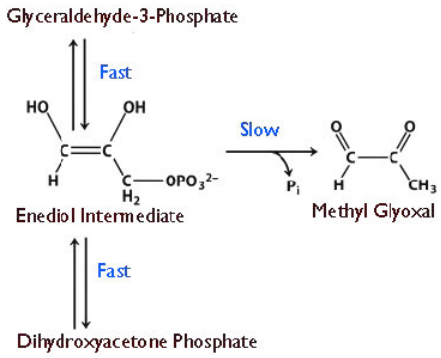


Figure 4.25 - A high speed reaction avoids production of methyl glyoxal. Image by Aleia Kim

Dissociation constant

In studying proteins and ligands, it is important to understand the "tightness" with which a protein (P) "holds onto" a ligand (L). This is measured with the dissociation constant (K_d). The formation of a ligand-protein complex LP occurs as

$$L + P \rightleftharpoons LP \tag{7.3.12}$$

The dissociation of the complex, therefore, is the reverse of this reaction, or

$$LP \rightleftharpoons L + P \tag{7.3.13}$$

so the corresponding dissociation constant is defined as

$$K_d = \frac{[L][P]}{[LP]} \tag{7.3.14}$$

where $[L]$, $[P]$, and $[LP]$ are the molar concentrations of the protein, ligand and the complex when they are joined together.

Smaller values of K_d indicate tight binding, whereas larger values indicate loose binding. The dissociation constant is the inverse of the association constant.

$$K_a = \frac{1}{K_d} \tag{7.3.15}$$

Where multiple molecules bond together, such as

$$J_xK_y \rightleftharpoons xJ + yK \tag{7.3.16}$$

The complex J_xK_y is breaking down into x subunits of J and y subunits of K. The dissociation constant is then defined as

$$K_d = \frac{[J]^x[K]^y}{[J_xK_y]} \tag{7.3.17}$$

where $[J]$, $[K]$, and $[J_xK_y]$ are the concentrations of J, K, and the complex J_xK_y, respectively.

Lineweaver-Burk plots

The study of enzyme kinetics is typically the most math intensive component of biochemistry and one of the most daunting aspects of the subject for many students. Although attempts are made to simplify the mathematical considerations, sometimes they only serve to confuse or frustrate students. Such is the case with modified enzyme plots, such as Lineweaver-Burk (Figure 4.26).

Indeed, when presented by professors as simply another thing to memorize, who can blame students? In reality, both of these plots are aimed at simplifying the determination of parameters, such as K_m and V_{max}. In making either of these modified plots, it is important to recognize that the same data is used as in making a V0 vs. [S] plot. The data are simply manipulated to make the plotting easier.

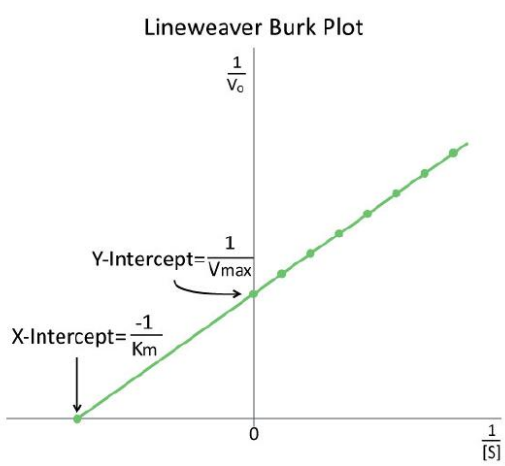


Figure 4.26 - A Lineweaver-Burk plot of $1/V_0$ vs $1/[S]$. Image by Aleia Kim

Double reciprocal

For a LineWeaver-Burk plot, the manipulation is using the reciprocal of the values of both the velocity and the substrate concentration. The inverted values are then plotted on a graph as 1/V0 vs. 1/[S]. Because of these inversions, Lineweaver-Burk

plots are commonly referred to as ‘double-reciprocal’ plots. As can be seen in Figure 4.26, the value of Km on a Lineweaver Burk plot is easily determined as the negative reciprocal of the x-intercept , whereas the Vmax is the inverse of the y-intercept. Other related manipulation of kinetic data include Eadie-Hofstee diagrams, which plots V0 vs V0/[S] and gives Vmax as the Y-axis intercept with the slope of the line being -Km.

Molecularity of reactions

The term molecularity refers to the number of molecules that must come together in order for a reaction to take place. Reactions of the sort of A -> B (where ‘A’ is the reactant and ‘B’ is the product) are unimolecular, since A is directly changed into B. The rate of the reaction is related only to the concentration of reactant A. For a bimolecular reaction where A + B ⇄ C the reaction depends on the concentration of both A and B and its rate will be related to the product of the concentration of A and of B.

Coenzymes

Organic molecules that assist enzymes and facilitate catalysis are co-factors called coenzymes. The term co-factor is a broad category usually subdivided into inorganic ions and coenzymes. If the coenzyme is very tightly or covalently bound to the enzyme, it is referred to as a prosthetic group. Enzymes without their co-factors are inactive and referred to as apoenzymes. Enzymes containing all of their co-factors are called holoenzymes.

Cofactor	Enzyme
Coenzyme	
5′-Deoxyadenosyl cobalamin	Methylmalonyl mutase
Biotin	Pyruvate carboxylase
Coenzyme A (CoA)	Acetyl CoA carboxylase
Flavin adenine nucleotide	Monoamine oxidase
Nicotinamide adenine dinucleotide	Lactate dehydrogenase
Pyridoxal phosphate	Glycogen phosphorylase
Tetrahydrofolate	Thymidylate synthase
Thiamine pyrophosphate	Pyruvate dehydrogenase
Metal	
K^+	Propionyl CoA carboxylase
Mg^{2+}	Restriction Endonucleases; Hexokinase
Mn	Superoxide dismutase
Mo	Nitrate reductase
Ni^{2+}	Urease
Se	Glutathione peroxidase
Zn^{2+}	Carbonic anhydrase; Carboxypeptidase

Figure 4.27 - Enzyme cofactors. Image by Aleia Kim

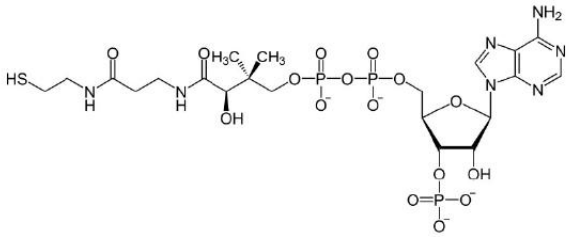


Figure 4.28 - Coenzyme A (CoA)

Pre-steady state kinetic studies

In the study of kinetic rates of enzymatic reactions, time zero is a very critical point. It establishes when the mixing of substrate with enzyme and measurement of formation of product begins. At time zero, there is no product. As shown in Figure 4.29, the appearance of product (on a short time scale) goes through an early burst phase with a steep slope for [product]/time and then changes.

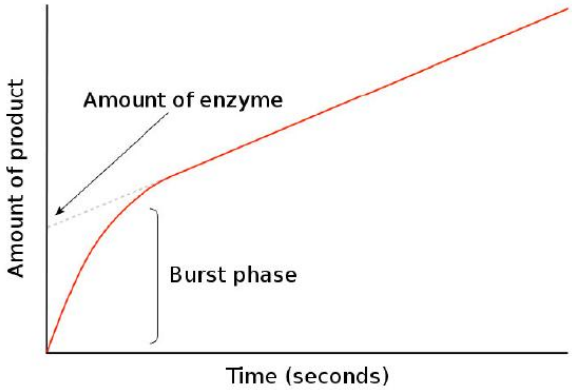


Figure 4.29 - Burst phase of product formation

This change occurs during a critical period in an enzymatic reaction and gives information about the rate of reaction cycles. The duration of the burst phase tells how long a single reaction turnover occurs, whereas the slow of the line post-burst phase tells the amount of "functional" enzyme performing the reaction.

After the burst phase, the slope of the line of the amount of product versus time decreases. This is due to the reaction entering conditions of steady state, used to study Michaelis-Menten kinetics. In steady state conditions, the amount of the enzyme-substrate complex (ES) is relatively constant over time. In simple terms, this occurs when the rate of formation of the ES complex equals the rate of conversion of the substrate to product by the enzyme with release of the product.

Earlier events

Events occurring prior to the conditions of steady state are referred to as pre-steady state. Depending on the enzyme, in as short as a few milliseconds, steady state conditions can be present meaning that if one hopes to study formation of reaction intermediates in pre-steady state, tools for this analysis must work very rapidly. One instrument commonly used for studying pre-steady state kinetics is called a stopped flow instrument.

It loads an enzyme solution and a substrate into separate syringes whose output is pointed into a mixing chamber. The solutions are rapidly mixed and measurements of product concentration begin. With a stopped flow instrument, dead times (time between mixing and detection) can be achieved of as small as 0.3 msec.

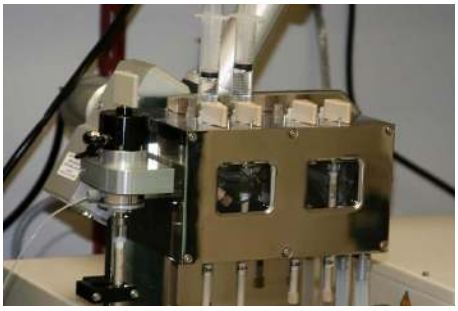
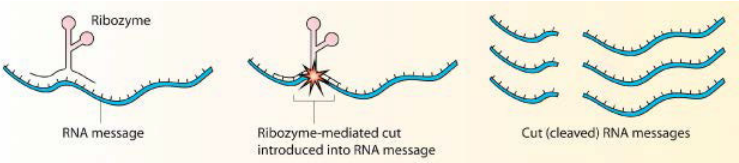


Figure 4.30 - Stopped flow instrument for studying pre-steady state kinetics. Wikipedia

Ribozymes

Proteins do not have a monopoly on acting as biological catalysts. Some RNA molecules are also capable of speeding reactions. The most famous of these molecules was discovered by Tom Cech in the early 1980s Studying excision of an intron in Tetrahymena, Cech was puzzled at his inability to find any proteins catalyzing the process. Ultimately, the catalysis was recognized as coming from the intron itself. It was a self-splicing RNA and since then, many other examples of catalytic RNAs have been found. Catalytic RNA molecules are known as ribozymes.



359

Figure 4.31 - Cleavage of an RNA by a ribozyme. Wikipedia

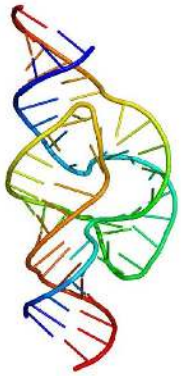


Figure 4.32 - Hammerhead ribozyme. Wikipedia

Not unusual

Ribozymes, however, are not rarities of nature. The protein-making ribosomes of cells are essentially giant ribozymes. The 23S rRNA of the prokaryotic ribosome and the 28S rRNA of the eukaryotic ribosome catalyze the formation of peptide bonds.

Ribozymes are also important in our understanding of the evolution of life on Earth. They have been shown to be capable via selection to evolve self-replication. Indeed, ribozymes actually answer a chicken/egg dilemma - which came first, enzymes that do the work of the cell or nucleic acids that carry the information required to produce the enzymes. As both carriers of genetic information and catalysts, ribozymes are likely both the chicken and the egg in the origin of life.

This page titled 7.3: Basic Principles of Catalysis is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **4.1: Basic Principles of Catalysis** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **7.1: Basic Principles of Catalysis** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

7.4: Derivation of Michaelis-Menten equation

Enzymes are **protein catalysts,** they influence the **kinetics** but **not the thermodynamics** of a reaction

- Increase the rate of a chemical reaction
- Do not alter the equilibrium

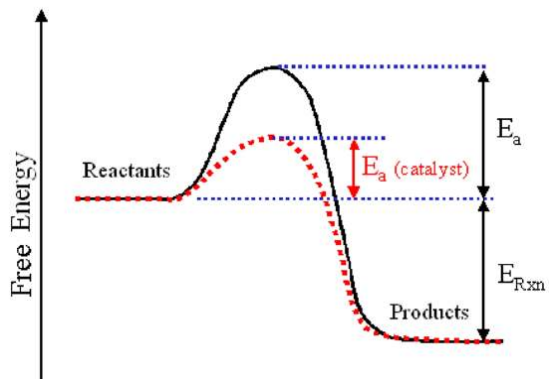
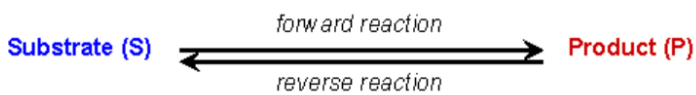


Figure 6.2.1: *Catalyst activity*

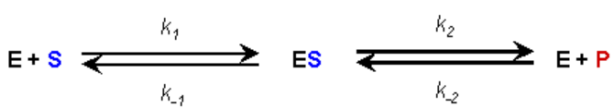
- They increase the rate by stabilizing the transition state (i.e. lowering the energy barrier to forming the transition state (they do not affect the energetics of either the reactant(s) or product(s)

Michaelis-Menten derivation for simple steady-state kinetics

The Michaelis-Menten equation is a **mathematical model** that is **used to analyze simple kinetic data**. The model has certain **assumptions**, and as long as these assumptions are correct, it will accurately model your experimental data. The derivation of the model will highlight these assumptions.

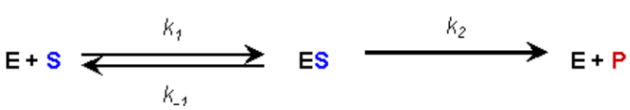


In an enzyme catalyzed reaction the substrate initially forms a reversible complex with the enzyme (i.e. the **enzyme and substrate have to interact** for the enzyme to be able to perform its catalytic function). The standard expression to show this is the following:



ASSUMPTION #1:

- There is no product present at the start of the kinetic analysis
- Therefore, as long as we monitor **initial reaction rates** we can **ignore** the reverse reaction of **E+P going to ES**



ASSUMPTION #2:

- During the reaction an **equilibrium** condition is established for the binding and dissociation of the Enzyme and Substrate (Briggs-Haldane assumption)
- Thus, the **rate of formation** of the ES complex is equal to the **rate of dissociation plus breakdown**

ASSUMPTION #3:

- **[E] << [S]**
- The enzyme is a catalyst, it is not destroyed and can be recycled, thus, only small amounts are required
- The amount of S bound to E at any given moment is **small** compared to the amount of free S
- It follows that **[ES] << [S]** and therefore **[S] is constant** during the course of the analysis (NOTE: this assumption requires that the reaction is monitored for a **short period**, so that not much S is consumed and [S] does not effectively change - see next assumption)

ASSUMPTION #4:

- Only the **initial velocity** of the reaction is measured
- [P] = 0 (reverse E + P reaction can be ignored)
- [S] » $[S]_{initial}$

ASSUMPTION #5:

- The enzyme is either present as free enzyme or as the ES complex
- $\mathbf{[E]_{total} = [E] + [ES]}$

Michaelis-Menten derivation using above assumptions:

Rate of ES formation = $\boldsymbol{k_1[E][S] + k_{-2}[E][P]}$

Assumption #1 says we can ignore the k_{-2} reaction, therefore:

Rate of ES formation = $\boldsymbol{k_1[E][S]}$

Assumption #5 says $[E] = [E]_{total} - [ES]$, therefore:

Rate of ES formation = $\boldsymbol{k_1([E]_{total} - [ES])[S]}$

The rate of ES breakdown is a combination of the dissociation and the conversion to product:

Rate of ES breakdown = $\boldsymbol{k_{-1}[ES] + k_2[ES]}$

Rate of ES breakdown = $\boldsymbol{(k_{-1} + k_2)[ES]}$

Assumption #2 says the rate of ES formation equals the rate of breakdown:

$\boldsymbol{k_1([E]_{total} - [ES])[S] = (k_{-1} + k_2)[ES]}$

Rearrange to define in terms of rate constants:

$\boldsymbol{([E]_{total} - [ES])[S] / [ES] = (k_{-1} + k_2) / k_1}$

$\boldsymbol{([E]_{total}\,[S] / [ES]) - [S] = (k_{-1} + k_2) / k_1}$

Define a new constant, $\boldsymbol{K_m = (k_{-1} + k_2) / k_1}$

$\boldsymbol{([E]_{total}\,[S] / [ES]) - [S] = K_m}$

Solve for the [ES] term (for reasons that will be given in the next step):

$\boldsymbol{[ES] = [E]_{total}\,[S] / (K_m + [S])}$

The actual reaction velocity measured at any given moment is given by:

$\boldsymbol{V = k_2[ES]}$

Multiple both sides of the above equation by k_2:

$\boldsymbol{k_2[ES] = k_2[E]_{total}\,[S] / (K_m + [S])}$

thus

$\boldsymbol{V = k_2[E]_{total}\,[S] / (K_m + [S])}$

The maximum possible velocity (Vmax) occurs when all the enzyme molecules are bound with substrate [ES] = $[E]_{total}$, thus:

$\mathbf{V_{max} = k_2[E]_{total}}$

Substituting this into the prior expression gives:

$\mathbf{V = V_{max}\,[S] / (K_m + [S])}$

This is the mathematical expression that is used to model your experimental kinetic data

It is known as the Michaelis-Menten equation

Experimental approach

The general approach is to add a **known concentration of substrate** to the enzyme and to determine the **initial reaction rate** for that concentration of substrate

- Reaction rates are typically given as moles (or micromole) of product produced per unit of time (sec or min) per mole (or micromole) of enzyme
- The experiment is repeated for a wide range of substrate concentrations
- A table of [S] versus V datapoints are collected
- These datapoints are plotted (V versus S) and should fit a curve that agrees with the Michaelis-Menten equation

The V_{max} and K_m terms are **intrinsic** properties of the particular enzyme/substrate combination that you are studying

- They will be determined from the features of the V versus S plot

V_{max}

There are a limited number of enzyme molecules and they can only perform a single reaction at a time. Thus, at **high [S]** the enzymes can be **saturated**

- Under saturating conditions the reaction is going as fast as it can, and additional increases in [S] do not increase the reaction rate.
- The maximum observable rate is V_{max} and the data will **asymptote to this value** at high [S]
- At low [S] the reaction rate is generally linearly proportional to the [S] (i.e. at low [S] if you double [S] the V will double)

K_m

$\mathbf{K_m = (k_{-1} + k_2) / k_1}$ **= (rate of breakdown of ES)/(rate of formation of ES)**

- K_m is **similar**, but not exactly equal to, a dissociation constant (K_d) for the ES complex
- If $k_{-1} >> k_2$, then $K_m \gg K_d$
- Due to this similarity to the expression for K_d, a **low value of K_m** is often interpreted as a **high affinity of the enzyme for the substrate**, and a large value for K_m is often interpreted as a weak affinity of the enzyme for the substrate
- K_m has units of molar concentration (just like the units for [S])

There is a mathematical treatment that allows for the determination of Km from the experimental V versus [S] data

- Consider the situation when the [S] being evaluated results in a value of V that is exactly $^1/_2$ of the maximum reaction velocity:

$\mathbf{V = V_{max}\,[S] / (K_m + [S])}$

$\mathbf{^1/_2 V_{max} = V_{max}\,[S] / (K_m + [S])}$

$\mathbf{^1/_2 = [S] / (K_m + [S])}$

$\mathbf{K_m + [S] = 2[S]}$

$\mathbf{K_m = [S]}$

Thus, K_m equals the substrate concentration that results in exactly one half the maximum possible reaction velocity

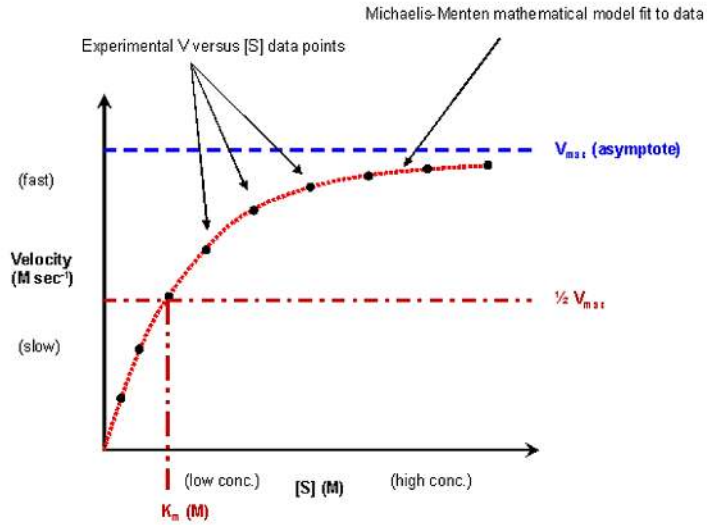


Figure 6.2.2: *K_m*

Lineweaver-Burke (the "*double reciprocal*" plot)

- The Michaelis-Menten equation can be rearranged by taking the reciprocal, to yield:

$$\frac{1}{V} = \frac{K_m}{V_{max}[S]} + \frac{1}{V_{max}}$$

- If X = 1/[S] and Y=1/***V*** then this is a linear equation with a slope of $K_m/\boldsymbol{V}_{max}$ and a Y intercept of $1/\boldsymbol{V}_{max}$

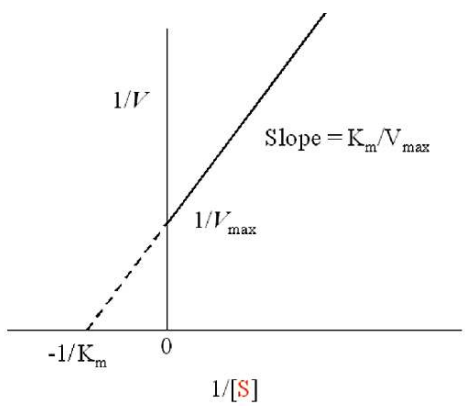


Figure 6.2.3: *1/S and1/V*

- Since the plot of 1/[S] versus 1/v data should be a straight line, it is easier to fit a linear function to the data in this form, and Vmax and K_m can be readily determined from the plot

Reversible Inhibition

There are two major categories of reversible inhibitors: **competitive** reversible inhibitors, and **noncompetitive** reversible inhibitors:

Competitive inhibitors

The inhibitor (**I**) competes with the substrate (**S**) for the enzyme active site (also known as the **S-binding site**). Binding of either of these molecules in the active site is a **mutually exclusive** event

- The substrate and inhibitor share a high degree of structural similarity. However, the inhibitor cannot proceed through the reaction to produce product.
- Increasing the concentration of substrate will **outcompete** the inhibitor for binding to the enzyme active site
- A competitive reversible inhibitor can be identified by its characteristic effects upon kinetic data

The expression for the Michaelis-Menten expression in the presence of a reversible competitive inhibitor is:

$V = V_{max}$ [S] / (K_m(1+[I]/K_i) + [S])

Where K_i is the actual EI complex dissociation constant

The effects of the reversible competitive inhibitor on the kinetics are as follows:

- If no inhibitor is present (i.e. if [**I**] = 0) then the equations are the same
- As inhibitor is added, the effect is to modify the apparent value of K_m. In particular, **the apparent K_m will be increased** by a value equal to **(1 + [I]/K_I)**. If K_m is increased, **the reaction velocity v will decrease**.
- Note that as [**S**] gets very large the value of the denominator is essentially equal to [**S**] and **v @ v_{max}**. Thus, the reaction velocity can be driven to vmax with a high enough substrate concentration

The diagnostic criteria for reversible competitive inhibition is that while the *apparent* Km is affected by addition of the inhibitor, the value of v_{max} does not change

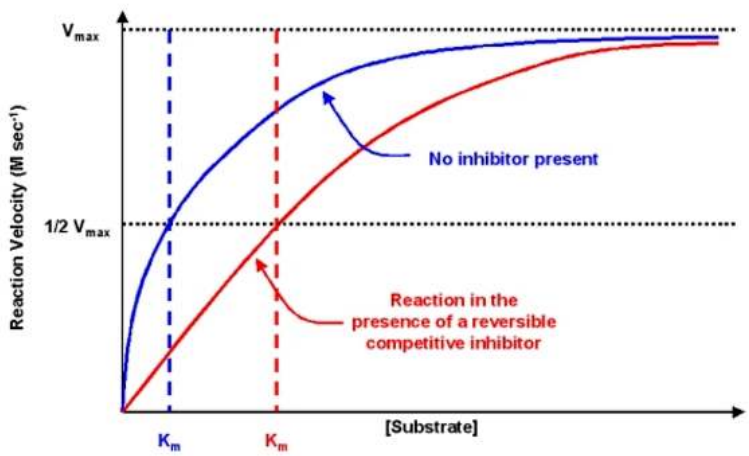


Figure 6.2.4: *Effect of reversible competitive inhibitor*

How is the Lineweaver-Burke double reciprocal plot affected by the presence of a reversible competitive inhibitor?

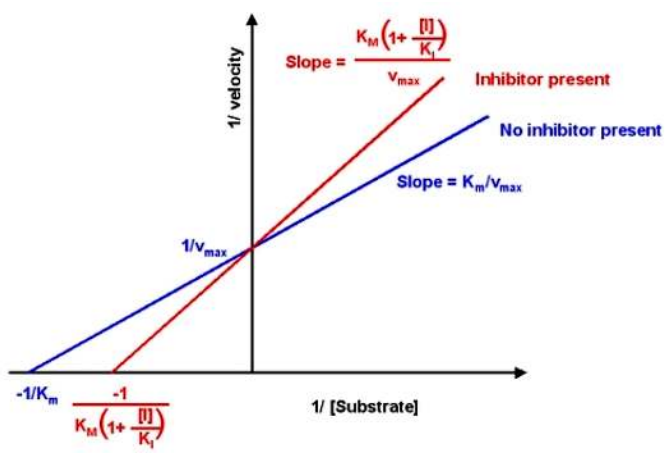


Figure 6.2.5: *Double reciprocal plot with reversible competitive inhibitor*

Noncompetitive Inhibitors

Noncompetitive inhibitors react with both E **and ES** (this is because the noncompetitive inhibitor does not bind at the same site in the enzyme as the substrate)

- Inhibition **cannot** be overcome by increasing the concentration of S
- The effect on kinetics is as if the enzyme were less active (**v_{max} is reduced**), but that the affinity for substrate is unaffected (**K_m remains the same**) since the substrate binding site is not occupied by the noncompetitive inhibitor.

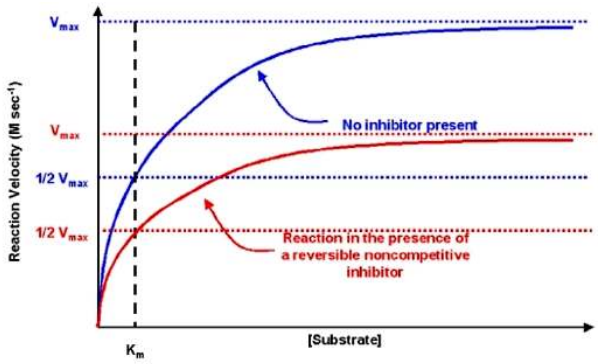


Figure 6.2.6: *Effect of reversible noncompetitive inhibitor*

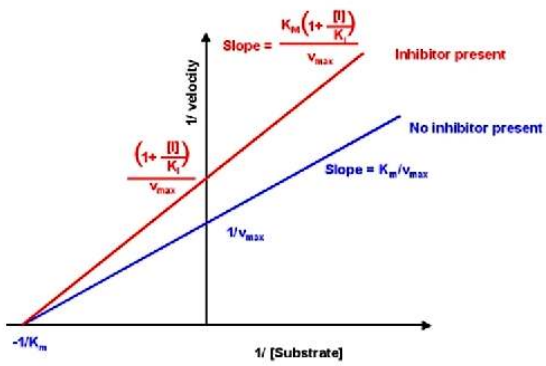


Figure 6.2.7: *Double reciprocal plot with noncompetitive inhibitor*

This page titled 7.4: Derivation of Michaelis-Menten equation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Michael Blaber.

- **6.2: Enzyme kinetics** by Michael Blaber has no license indicated.
- **7.2: Derivation of Michaelis-Menten equation** by Michael Blaber has no license indicated.

7.5: Binding - The First Step Towards Protein Function

Learning Goals (ChatGPT o1, 1/27/25)

Below is a series of learning goals designed to help junior and senior biochemistry majors master the concepts and mathematics behind reversible ligand binding to macromolecules:

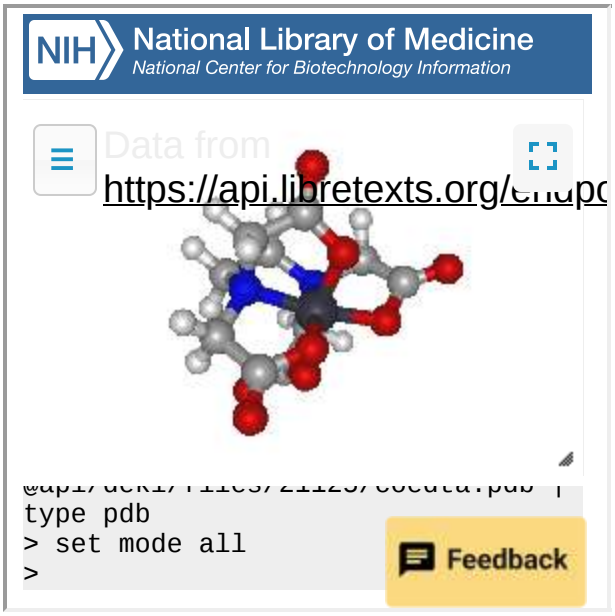
- **Conceptual Foundations of Binding**
  - Define reversible binding and distinguish noncovalent interactions from covalent bonds.
  - Identify various types of ligands (e.g., substrates, hormones, metal ions) and their roles in biological processes.
  - Explain the significance of ligand binding in eliciting the biological properties of macromolecules.
- **Equilibrium and the Dissociation Constant (K_D)**
  - Derive and interpret the dissociation constant (K_D) as a measure of binding strength.
  - Understand how KD is defined as the reciprocal of the equilibrium constant (K_{eq}) and its units (molarity).
  - Compare scenarios of "tight" versus "loose" binding by relating K_D values to biological function.
- **Mathematical Modeling of Binding**
  - Apply mass balance equations for both macromolecule (M) and ligand (L) to develop expressions for the concentration of the bound complex (ML).
  - Derive the hyperbolic binding equation $ML=M_0L/(K_D+L)$ and explain its assumptions and limitations.
  - Develop a quadratic expression for ML when total ligand concentration (L_o) is known and free ligand concentration is not directly measured.
- **Fractional Saturation and Graphical Analyses**
  - Define fractional saturation (Y) and show how it can be calculated from binding data.
  - Interpret hyperbolic binding curves and understand how the ligand concentration at half-maximal binding relates to K_D.
  - Compare different plotting methods (e.g., hyperbolic, double reciprocal, Scatchard, and semilog plots) and describe what each reveals about the binding interaction.
- **Kinetic Considerations in Binding**
  - Relate the equilibrium binding parameters to kinetic rate constants (k_{on} and k_{off}).
  - Explain how the ratio of k_{on} to k_{off} yields the equilibrium constant and how changes in these rates affect binding affinity.
  - Use first-order kinetics to calculate the half-life ($t_{½}$) of a bound complex and discuss experimental approaches to measure k_{off}.
- **Complex Binding Scenarios**
  - Analyze binding interactions in systems involving dimerization or oligomerization, including how mass balance changes for multimeric species.
  - Derive and interpret equations for a macromolecule binding a ligand at two independent sites, and assess how differences in K_D for each site affect overall saturation.
  - Discuss the continuum of binding affinities from nonspecific (weak) to specific (tight) interactions, and link these to thermodynamic parameters like $\Delta G°$.

By achieving these learning goals, students will integrate both the conceptual understanding and the quantitative skills necessary to analyze binding equilibria in diverse biochemical systems.

7.5.1: Reversible Binding of a Ligand to a Macromolecule

Reversible, noncovalent binding of two or molecules is the first step in expressing the biological properties of almost all biomacromolecules. If one of the molecules is small, it's often called a ligand. Ligands are often called other names. Substrates are the reactants that bind to the active sites of enzymes. Hormones and neurotransmitters bind to solution phase or membrane-bound receptor proteins. Metal ions (monatomic, like Ca^{2+}, or molecular, like $CH_3CO_2^-$) are also considered ligands when bound to proteins or nucleic acids.

You might be more familiar with the term ligand when it's applied to the coordination of a transition metal complex by electron pair donors (Lewis acids) on single or multidentate molecules, which for transition metal complexes are called ligands. Here is an interactive molecule model of a cobalt ion binding to EDTA, a multidentate ligand.



The cobalt ion (dark grey ball) is octahedrally coordinated to the multidentate ligand EDTA.

Whether a macromolecule M and a ligand L bind to each other depends on their relative concentrations and how tightly they bind. Compare this to an acid. Its pKa and the pH of the medium determine if it deprotonates.

Biochemists rarely discuss equilibrium constants to describe the strength of a binding interaction, but rather their reciprocals - the dissociation constants, K_D. For the reactions $M + L \leftrightarrow ML$, where M is a free macromolecule, L is a free ligand, and ML is a macromolecule-ligand complex (which is held together by intermolecular forces, not covalent forces), the K_D is given by

$$K_D = [M]_{eq}\,[L]_{eq}/[ML]_{eq} \tag{7.5.1}$$

Figure 7.5.1 shows free and bound M and L.

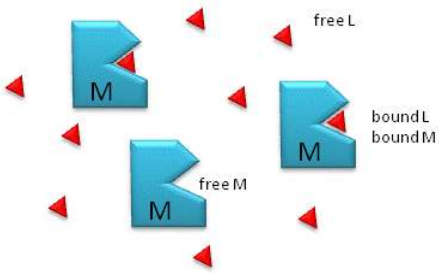


Figure 7.5.1: Free macromolecule (M), free ligand (red triangle) and bound complex ML

Notice the unit of K_D is molarity, M.

- The lower the K_D (i.e., the higher the [ML] at any given M and L), the tighter the binding.

- The higher the K_D, the looser the binding. K_Ds for biological molecules are finely tuned to their environments.

KD values vary from about 1 mM (weak interactions) for some enzyme-substrate complexes to pM-fM levels. Examples of very tight, non-covalent interactions include the avidin (an egg protein)-biotin (a vitamin) and thrombin (enzyme initiating clotting)-hirudin (a leech salivary protein) complexes. The values are "tuned" so that the relative concentrations of free and bound M and L are appropriate for a biological setting.

To understand binding, it is important not only to know the non-covalent interactions that lead to binding but also to ask the simple question: is the ligand bound to the macromolecule? If so, to what extent? To know if M or L is bound, we must use simple mathematics that you would have learned in Introductory or Analytical Chemistry courses. We'll start with the mathematical description, which is harder for students to understand than the noncovalent interactions.

We will start with three basic equations:

For the Dissociation constant:

$$K_D = ([M]eq[L]eq)/[ML]eq = ([M][L])/[ML] \tag{7.5.2}$$

(Note that K_D has units of molarity.);

For Mass Balance of M:

$$M_0 = M + ML \tag{7.5.3}$$

where M_0 is the total amount of macromolecule. (note: brackets and the eq subscript will be left off if the resulting equation is non-ambiguous)

For Mass Balance of L:

$$L_0 = L + ML \tag{7.5.4}$$

where L_0 is the total amount of ligand

We want to derive equations that give ML as a function of known or measurable values. The K_D equation (5.1) shows that ML depends on free M and free L. From the equations above, we can derive two fundamental and equally valid equations that are useful under different experimental conditions.

7.5.1.1: Case 1:

This applies when you can readily measure **free** L OR when experimental conditions are that Lo >> Mo (so L= Lo), often encountered in a lab setting. You don't have to measure free L since, in this case, it is approximately the total ligand that was added to the system.

Substitute 5.1.3 into 5.1.1 gives

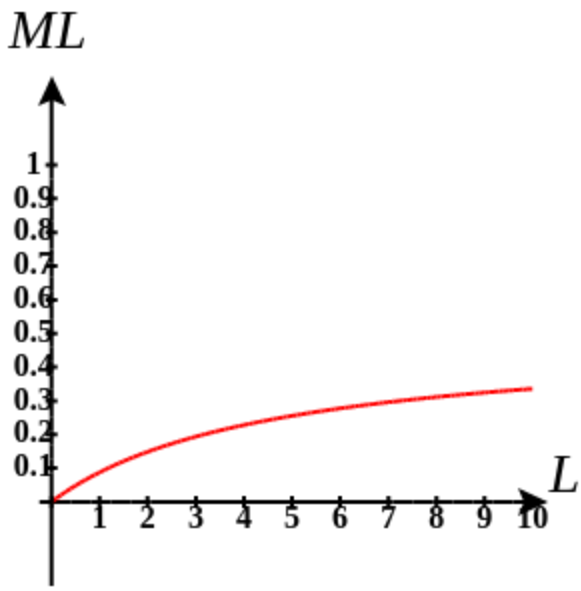
$$\begin{gathered} K_D = ([M][L])/[ML] = [Mo - ML][L]) /[ML \\ (ML)K_D = (M_o)\,L - (ML)L \\ (ML)K_D + (ML)L = (M_o)\,L \\ (ML)\,(K_D + L) = (M_0)\,L \end{gathered} \tag{7.5.5}$$

or

$$(ML) = \frac{(M_0)\,L}{K_D + L} \tag{7.5.6}$$

This equation is ALWAYS TRUE for the chemical equation written above. L is the free ligand concentration at equilibrium.

An interactive plot of the concentration of the ML complex (ML) vs free L (L) is shown below. Vary the sliders and note the changes in the graph.



If L_0 >> M_0, then the equations simplify to:

$$ML = \frac{(M_0)\,(L_0)}{K_D + L} \tag{7.5.7}$$

Dividing this equation by Mo gives the fractional saturation Y of the macromolecule M.

$$Y = [ML]/M_0 = \frac{L}{K_D + L} \tag{7.5.8}$$

where Y can vary from 0 (when L = 0) to 1 (when L >> K_D)

Note that the interactive graph above and graphs of ML vs L (equation 5.1.10) and Y vs L (equation 5.1.11) are all HYPERBOLAS

To get a "gut" level understanding of the graphs of $(ML) = (M_0)(L)/(K_D + L)$ and $Y = L/(K_D + L)$, let's consider 3 different values or sets of values of free ligand:

1. L = 0: This gives ML = 0
2. L = K_D: $(ML) = (M_0)(L)/(L + L) = (M_0)(L)/(2L) = Mo/2$ which indicates that M is half saturated. The operational definition of K_D is the ligand concentration at which the M is half-saturated.
3. L >> K_D: ML = M_0, and the macromolecule is saturated with ligand.

7.5.1.2: Case 2 (more general):

This applies when you know K_D, but don't know free L or haven't measured it, and you wish to calculate how much ML is present at equilibrium, given a K_D value. L_0 does not have to be much greater than M_0 in this case. If it were, like it is often in an experimental system, you would know that free L = L_0, and you could use Case 1.

In this case, we will substitute mass balance equations for both M_0 (Eq 5.1.2) and L_0 (Eq 5.1.3) into the equation for K_D (Eq. 5.1.1). This gives:

$$\begin{gathered} K_D = ([M][L])/[ML] = [M_0 - ML]\,[L_0 - ML]\,/[ML] \\ (ML)K_D = (M_0 - ML)\,(L_0 - ML) \\ (ML)K_D = (M_0)\,(L_0) - (ML)\,(L_0) - (ML)\,(M_0) + (ML)^2 \end{gathered} \tag{7.5.9}$$

or

$$(ML)^2 - (L_0 + M_0 + K_D)\,(ML) + (M_0)\,(L_0) = 0 \tag{7.5.10}$$

This can be rearranged into the form $ax^2 + bx + c = 0$ where

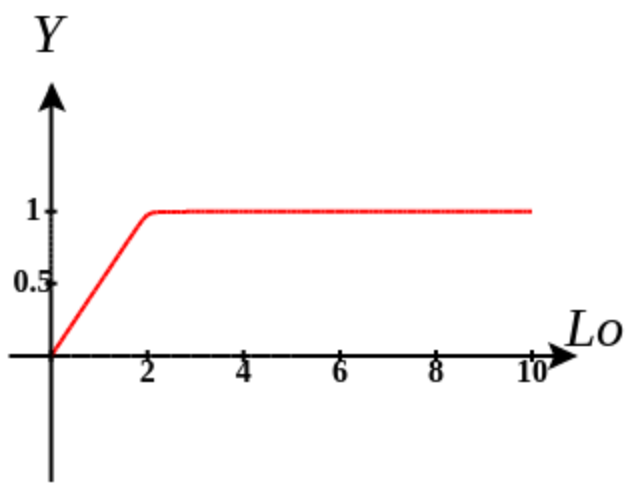
- a = 1
- b = - (L_0 + M_0 +K_D)
- c = (M_0)(L_0)

with the well known solution $x = [(-b) - (b^2 - 4(a)(c))^{1/2}]/2a$. Therefore,

$$(ML) = \left[(L_0 + M_0 + K_D) - \left((L_0 + M_0 + K_D)^2 - 4\,(M_0)\,(L_0)\right)^{1/2}\right]/2 \tag{7.5.11}$$

An interactive plot of the Y, fractional saturation, vs total L (L_0) is shown below. Vary the sliders and note the changes in the graph.

Fx Saturation vs Lo



In the derivations, we came up with two equations for ML, Eq 5.1.10, which gives ML vs. L, and Eq 5.1.16, which gives ML vs. L_0.

Both equations are valid. In the first, you must know free L, which is often L_0 if $M_0 << L_0$. In the second, you don't need to know free M or L at all. At a given Lo, Mo, and K_D, you can calculate ML, which should be the same ML you get from the first equation

if you know free L.

Equations 5.1.10 and 5.1.16 are useful in several circumstances. They can be used to

- calculate the concentration of ML if K_D, M_0, and L (for Eq. 5.1.10) or if K_D, M_0, and L_0 (for Eq. 5.1.16) are known. This is analogous to using the Henderson-Hasselbalch equation to calculate the protonation state (HA) and, hence, the charge state of an acid at various pH values. In the former binding case, we measure the concentration of a reversibly bound ligand (ML), and in the latter case, we measure the concentration of covalently bound protons (HA).
- calculate K_D if ML, M_0, and L (for Eq. 5.1.10) or if ML, M_0, and L_0 (for Eq. 5.1.16) are known. A separate chapter will discuss techniques for extracting the K_D from binding data.

7.5.2: Interpretation of Binding Analyses

It is important to get a mathematical understanding of the binding equations and graphs. It is equally important to get an intuitive understanding of their properties. Just as we used the +/- 2 pH rule in determining at a glance the charge state of an acid, you need to be able to determine the extent of binding (how much of M is bound with L) given their relative concentrations and the K_D. The usual situation is that $[M_0]$ is << $[L_0]$. What happens to the binding curves for M + L ↔ ML if the K_D gets progressively lower? Intuitively, you should expect binding to increase, especially as L increases. The curves below should help you develop the intuition you need with respect to binding equilibria. Figure **7.5.2** show Y vs L_0 at Varying K_Ds.

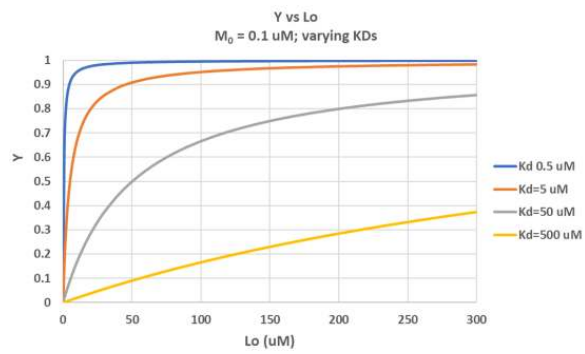


Figure **7.5.2**: Fractional saturation Y vs L_0 at Varying K_Ds

Figure **7.5.3** shows Y vs L_0 at a very low K_D (0.001 uM = 1 nM, resulting in a sharp "titration" curve. Any increment of L added is bound, so effectively, none is present in the free form. The graph abruptly changes to a horizontal line when all the macromolecules are bound. This curve could be used to determine $[M_0]$!

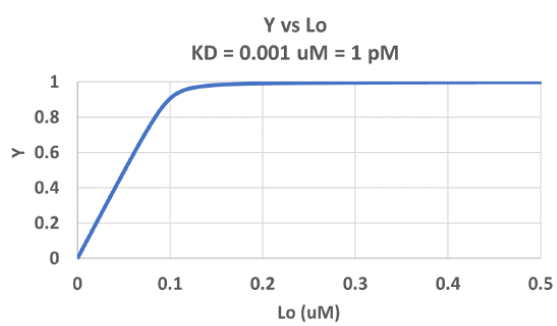
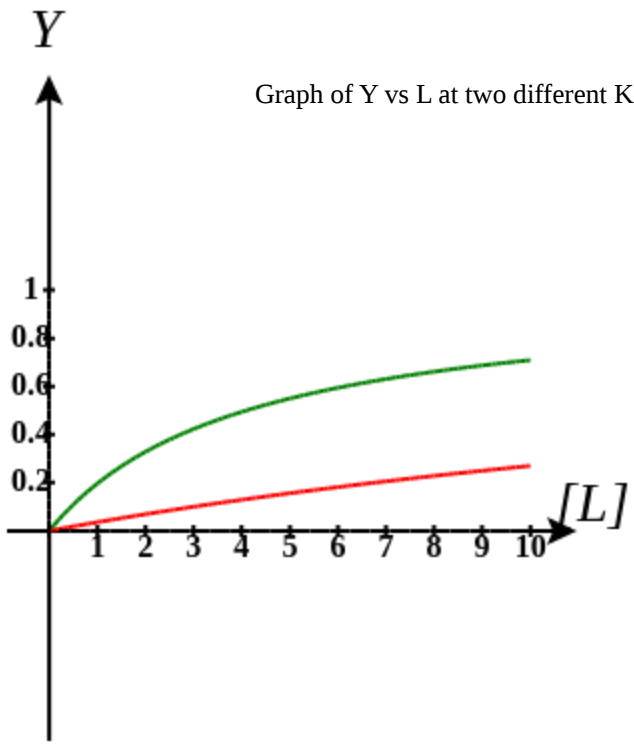


Figure **7.5.3**: Fractional saturation Y vs L_0 for very low K_D binding interactions

Note that in the last graph, given the same M_0 and L_0 concentrations, the "titration curves" for a binding equilibrium characterized by even tighter binding (for example, a K_D = 0.1 pM or 0.01 pM) would be indistinguishable from the graph when K_D = 1 pM. It should be apparent that for all of these K_D values, all of the added ligand is bound until $[L_0] > [M_0]$. To differentiate these cases, much lower ligand concentrations would be required so that all is not bound upon the addition of ligand. Also note that this curve is NOT hyperbolic, which makes sense since the graph is of Y vs L_0, not Y vs L, and since L_0 is not >> M_0.

The interactive graph below shows fractional saturation Y vs L at two different K_D values



It is quite interesting to compare graphs of Y (fractional saturation) vs L (free) and Y vs Lo (total L) in the special case when L_0 is not >> M_0. Figure 7.5.4 when M_0 = 4 μM, Kd = 0.19 μM . Under the ligand concentration used, it should be apparent that L can't be approximated by L_0

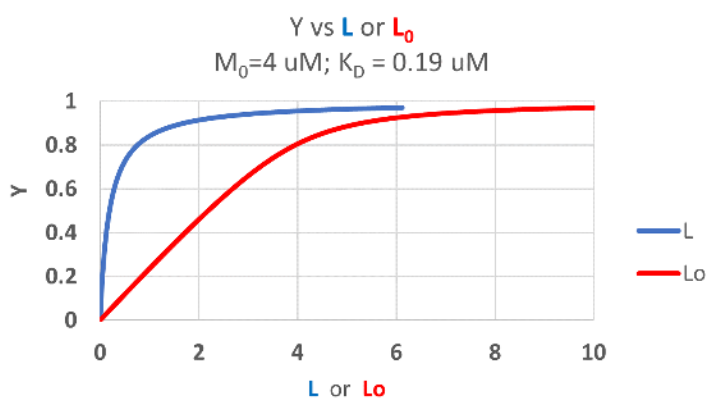


Figure 7.5.4: Fractional saturation Y vs L or L_0 when L_0 is not >> M_0

Two points should be evident from these graphs when L is not approximated by Lo:

- a graph of Y vs L_0 is not truly hyperbolic, but it does saturate
- a K_D value (ligand concentration at half-maximal binding) can not be estimated by inspection from the Y vs L_0, but it can be from the Y vs. L graph.

Figure 7.5.5 shows a comparison of the extent of covalent binding of a proton to an acid at pH values around the pKa and, by analogy, the extent of noncovalent binding of a ligand at log[L] values around the log K_D.

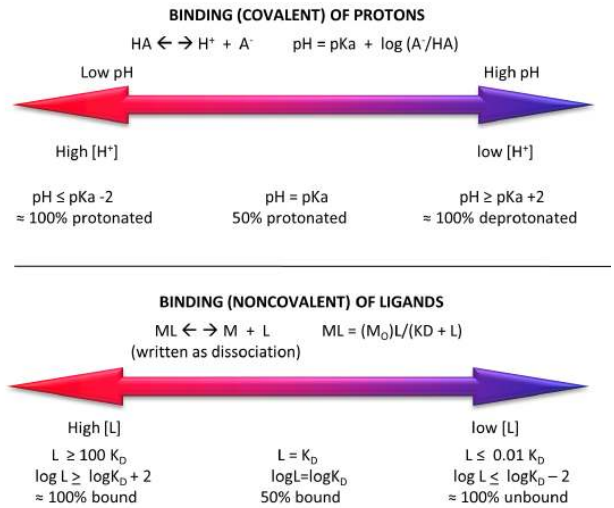


Figure 7.5.5: Analogy between covalent binding of protons and noncovalent binding of ligands to a molecule

7.5.3: Different Graphical Analyzes of Binding

In addition to the hyperbolic plots of [ML] vs [L] and fractional saturation Y vs [L], a variety of derivative plots are often encountered. The equations and their graphs (for two different K_D values) are shown below. The graphs are in the form of Y vs L_0, when L_0 is approximately equal to free L.

Hyperbolic saturation plot:

$$\mathrm{Y} = \frac{\mathrm{L}}{\mathrm{K_D} + \mathrm{L}} \tag{7.5.12}$$

Double reciprocal plot:

$$\frac{1}{\mathrm{Y}} = \frac{\mathrm{K_D} + \mathrm{L}}{\mathrm{L}} = \frac{\mathrm{K_D}}{\mathrm{L}} + 1 = \mathrm{K_D}\left(\frac{1}{\mathrm{L}}\right) + 1 \tag{7.5.13}$$

A plot of 1/Y vs 1/L has a slope of K_D and a y-intercept of 1 (which is the number of binding sites for this simple mechanism)

The Scatchard plot:

$$\begin{aligned} \mathrm{Y}\,(\mathrm{K_D} + \mathrm{L}) &= \mathrm{L} \\ Y\,(\,\mathrm{K_D}) + YL &= L \\ Y\,(\,\mathrm{K_D}) = L - \mathrm{YL} &= \mathrm{L}(1 - \mathrm{Y}) \end{aligned} \tag{7.5.14}$$

which gives the final Scatchard plot equation:

$$\frac{Y}{\mathrm{L}} = \frac{1 - \mathrm{Y}}{\mathrm{K_D}} = -\frac{\mathrm{Y}}{\mathrm{K_D}} + \frac{1}{\mathrm{K_D}} \tag{7.5.15}$$

A plot of Y/L vs Y has a slope of $-1/K_D$ and a y-intercept of $1/K_D$.

Y vs logL

Plotting Y vs **L** gives a hyperbola, but a plot of Y vs **log L** gives a **sigmoidal** plot. Plots of Y vs log L are often used in the research literature instead of traditional hyperbolic plots of Y vs L. There are several reasons for this:

- The log [L] is more fundamentally related to the thermodynamic expression that relates ΔG^0 and K_{eq} or K_D, namely

$$\Delta G^0 = -RT \ln K_{eq} = R\,\mathrm{T} \ln K_D \tag{7.5.16}$$

- plots of Y vs. L plateau over a very large range of [L], but given the compression of the X-axis values in a semilog plot, the plots reach a saturation plateau over a much narrower range of log [L]. A range from 1-100 on the [L] scale becomes 0-2 on the log [L] scale. Since it takes a very high ligand concentration to truly reach saturation (100xK_D), it's much easier to see if saturation is achieved in semilog plots.
- multiple plots of binding data for different K_D values have the same shape on a semilog plot. Binding data for a ligand to wild-type and mutant proteins, all with different K_Ds, will give identical plots with curves for higher K_D values shifted to the right. Semilog plots are also used routinely to display multiple plots of binding in the absence and presence of a binding inhibitor.

Figure **7.5.6** show different graphs for ligand binding to a macromolecule

7.5.4:

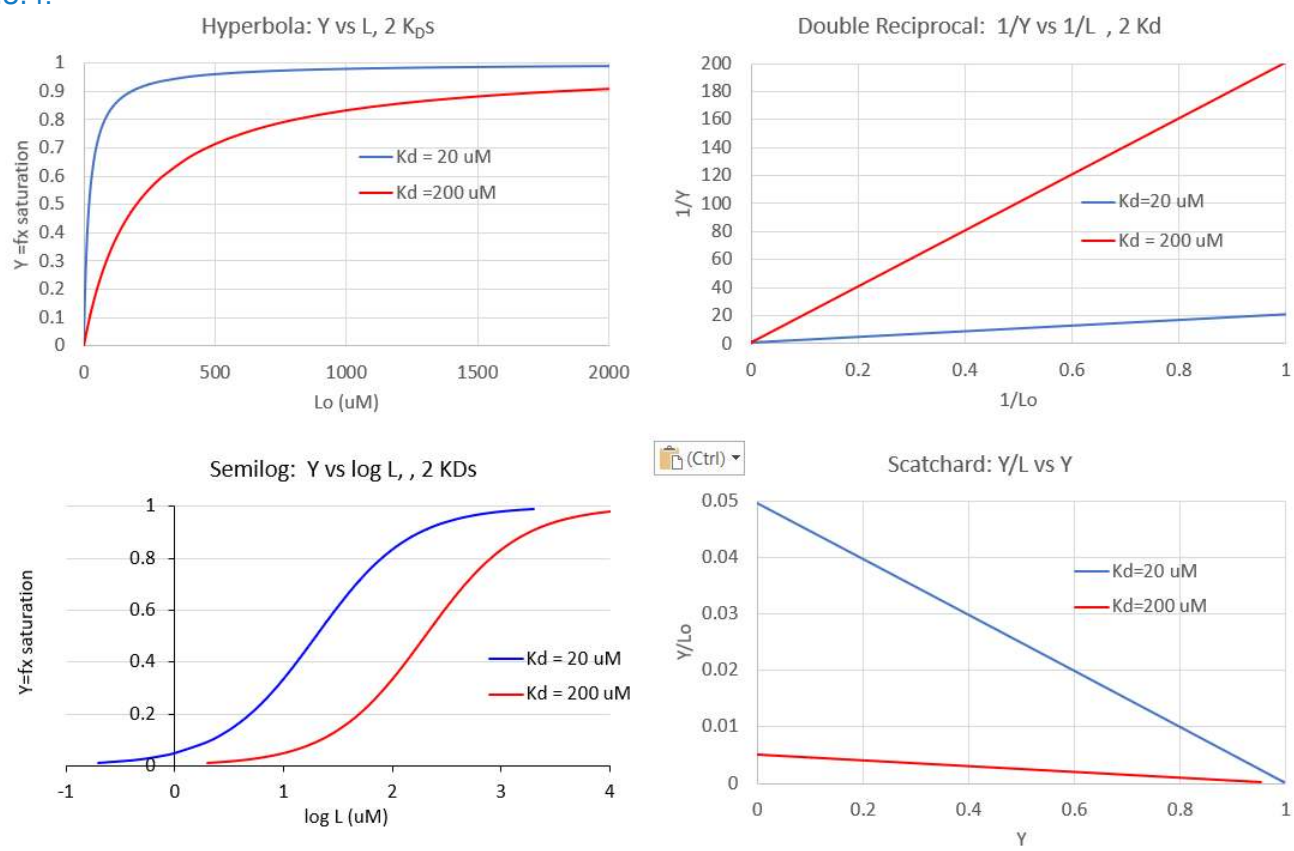


Figure **7.5.6**: Different graphs for ligand binding to a macromolecule

Sigmoidal binding curves: A note of caution

The graph of Y vs. log [L] is sigmoidal, but the same data would give a hyperbola if plotted as Y vs. [L]. However, as we will see in section 5.3, there are some occasions when the graph of Y vs. **[L]** is sigmoidal. For example, this can occur when the binding of a ligand to a multimeric binding protein affects the binding of the ligand to additional sites on the protein. This is an example of allosteric binding, which we will explore in detail in section 5.3. So, if you see a sigmoidal plot, be careful to examine the graph to see if it is a regular or semilog plot.

Straight line transformations of the hyperbolic binding equations are useful to get approximate values of K_D, but linear regression analysis to get slopes and intercepts is not statistically optimal, as the errors in the y variable (Y) and the y and x variables in the Scatchard plot are not identical across values. To determine K_D, it is best to fit experimental data to the nonlinear function for the hyperbola.

7.5.5: Dimerization and Multiple Binding Sites

In the previous examples, we considered the case of a macromolecule M binding a ligand L at a single site, as described in the equation below:

M + L ↔ ML

where K_D = [M][L]/[ML]

We saw that the binding curves (ML vs L or Y vs L are hyperbolic, with a K_D = L at half-maximal binding. But many other chemical equilibria can mechanistically explain binding data. We'll consider just two cases here.

Dimerization

A special yet common example of this equilibrium occurs when a macromolecule binds itself to form a dimer (D), as shown below:

M + M ↔ M_2 or D

where D is the dimer, and where

$$K_D = [M][M]/[D] = [M]^2/[D] \tag{7.5.17}$$

At first glance, you would expect a graph of [D] vs [M] to be hyperbolic, with the K_D again equaling the [M] at half-maximal dimer concentration. This turns out to be true, but a simple derivation is in order. In the case of dimer formation, Mo, which superficially represents both M and L in the earlier derived expression, is changing. So we have to invoke mass balance of M again: $[Mo] = [M] + 2[D]$, where the coefficient 2 is necessary since there are 2 M in each dimer.

More generally, for the case of the formation of trimers (Tri), tetramers (Tetra), and other oligomers, $[Mo] = [M] + 2[D] + 3[Tri] + 4[Tetra]$+....

Rearranging (12) and solving for D gives $D = ([M_0] - [M])/2$. Substituting this into the K_D expression (1) gives

$$K_D = \frac{M^2}{\frac{[M_0]-[M]}{2}} = \frac{2M^2}{[M_0] - [M]} = \tag{7.5.18}$$

This can be rearranged into quadratic form for M (not D):

$$2M^2 + K_D(M) - K_D\,(M_0) = 0 \tag{7.5.19}$$

which is of the form y = ax^2+bx+c.

Solving the quadratic equation gives [M] and with M_0 , D can be calculated from $D = ([M_0] - [M])/2$.

A value Y, similar to fractional saturation, can be calculated, where Y is the fraction of total possible D, which can vary from 0-1: $Y = 2D/M_0$

A graph of Y vs Mo with a dimerization dissociation constant K_D = 25 uM is shown in Figure 7.5.7.

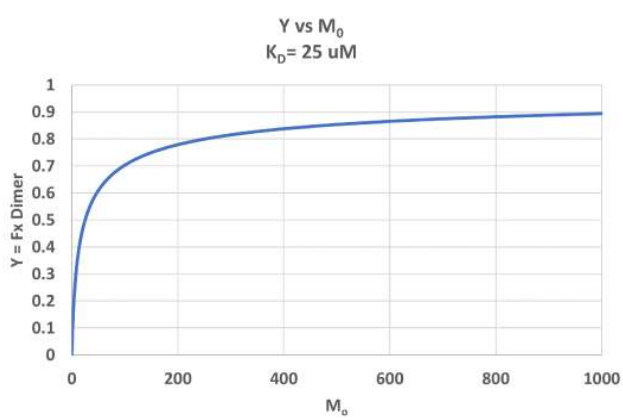
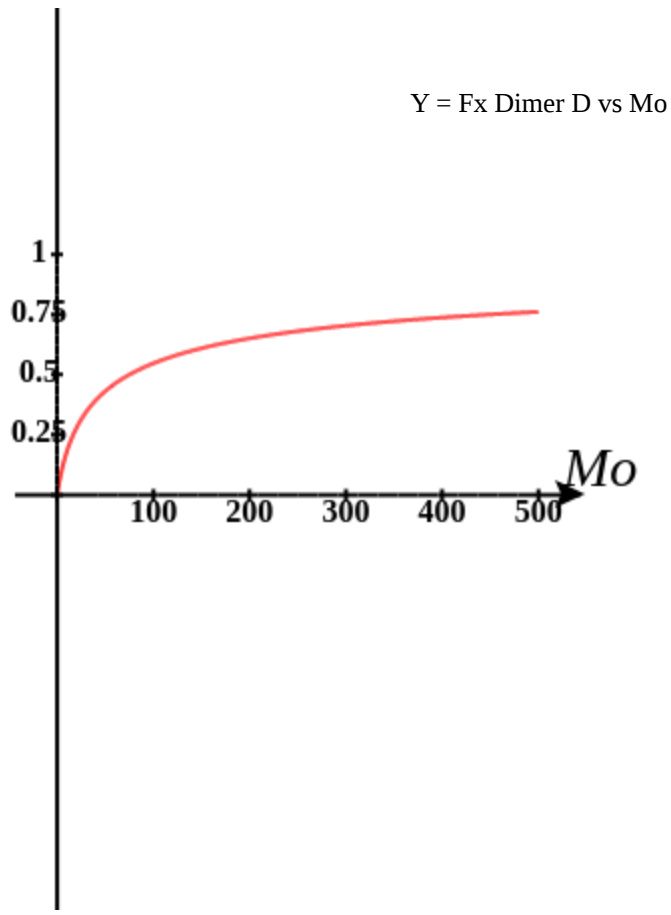


Figure 7.5.7: Fractional saturation Y vs M_0 for protein dimerization

Note that the curve appears somewhat hyperbolic. Half-maximal dimer formation does occur at a total M concentration $M_0 = K_D$. Also note, however, that even at M_0 = 1000 uM, which is 40x K_D, only 90% of the total possible D is formed (Y = 0.90). For the simple M + L ↔ ML equilibrium, if L_0 = 40x the K_D and $M_0 << L_0$, $Y = L/(K_D + L) = L/[(L/40) + L] = 0.976$

An interactive graph showing Y (the fraction of dimers) vs M_0 is shown. Move the sliders to show how changes in "K_D" affect the dimerization.



The aggregation state of a protein monomer is closely linked to its biological activity. Some proteins that can form dimers are active in the monomeric state, while others are active as dimers. High concentrations, found under conditions when proteins are crystallized for X-ray structure analysis, can drive proteins into the dimeric state. This may lead to the false conclusion that the active protein is a dimer. Determination of the actual physiological concentration of [Mo] and K_D gives investigators knowledge of the Y value, which can be correlated with biological activity. For example, interleukin 8, a chemokine that binds certain immune cells, exists as a dimer in X-ray and NMR structural determinations but as a monomer at physiological concentrations. Hence, the monomer, not the dimer, binds its receptors on immune cells. Viral proteases (herpes viral protease, HIV protease) are active in a dimeric form, in which the active site is formed at the dimer interface.

Binding of a ligand to two independent sites

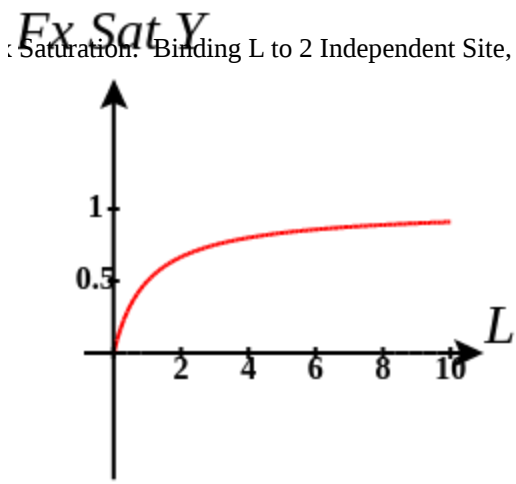
What if a ligand L binds to two different sites on the same biomacromolecule? Assuming that the binding of ligand L to one site does not affect the binding of the ligand to the other site (and vice versa), the following equation can be derived:

$$Y = \left[\frac{L}{(K_{D1} + L)} + \frac{L}{(K_{D2} + L)}\right]/2 \tag{7.5.20}$$

The numerator of the equation has a term for the fractional saturation of site 1 characterized by K_{D1} and a term for the fractional saturation of site 2 characterized by K_{D2}. These two terms are divided by 2 so that the fractional saturation of all sites is 1 at

saturating values of free ligand. Note that there can only be one free ligand concentration in solution, so the difference in the two terms in the numerator is the K_D value.

The interactive graph below shows such binding to two independent sites with different K_D values. Again, we'll assume the binding of one ligand does NOT influence the binding of the other.



7.5.6: The Binding Continuum

Binding affinities allow us to measure the relative strength of binding between two substances. But how "tight" is tight binding? Weak binding? Let us examine that issue by considering a binding continuum. Consider two substances, A and B, that might interact. Over what range of strengths can they actually bind to each other? It would be helpful to set up the extremes of the binding continuum. At one end, there is no binding at all. At the other end, consider two things that bind covalently. We have discussed how Kd reflects binding strength. Remember, $K_D = 1/K_{eq}$. Also, we know that K_{eq} is related to ΔG^o by the equations:

$$\Delta G^0 = -R\mathrm{T}\ln K_{eq} = R\mathrm{T}\ln K_D \tag{7.5.21}$$

Given these simple equations, you should be able to interconvert between K_{eq}, K_D, and ΔG^0. (Keep your units straight.).

No interaction: One end of the binding continuum represents no interaction. Let's assume that K_{eq} is tiny (K_D large) - for example, $K_{eq} \sim 2.4 \times 10^{-72}$. Plugging this into the equation \(ΔG^0 = - RTlnK_{eq}\), where R = 2.00 cal/mol·K, and T is about 300K, the $\Delta G^0 \sim +100$ kcal/mol (418 kJ/mo). If we add A + B, there is no drive to form AB. If AB did form, then it would immediately fall apart.

Covalent interaction: At the other end of the continuum, consider the interaction of 1H atom with another to form H_2. From a general chemistry book, we can get ΔG^0_{form}. Using simple thermodynamics, we can calculate ΔG^o for H-H formation. ($\Delta G^o = \Sigma\Delta G^0_{form}$ prod. - $\Sigma\Delta G^0_{form}$ react.) Doing this gives -97 kcal/mol (-406 kJ/mol).

Specific and Nonspecific Binding: Consider the interaction of a protein, the lambda repressor (R), with a small oligonucleotide to which it binds tightly (called the operator DNA, O). This is an example of a biologically tight but reversible interaction. Due to electrostatic interactions and H bonds between the positively charged protein and the negatively charged nucleic acid backbone, R can bind to many short oligonucleotides. The tight binding interaction, however, involves oligonucleotides of a specific base sequence. Hence, we can distinguish between tight binding, which usually involves specific DNA sequences, and weak binding, which involves nonspecific sequences. Likewise, we will speak of specific and nonspecific binding. R and O, which bind with a K_D of 1 pM, represent an example of specific binding.

In contrast, R and nonspecific DNA (D), which bind mostly through electrostatic interactions with a K_D of 1 mM, are an example of nonspecific binding. You might expect any positively charged protein, like mitochondrial cytochrome C, would bind negatively charged DNA. This nonspecific interaction would presumably have no biological significance since the two are localized in different cell compartments. In contrast, the interaction between positively charged histone proteins bound to DNA in the nucleus would be specific.

Rate constants for association and dissociation: When the reaction
M + L ↔ ML is at equilibrium, and the forward reaction rate is equal to the reverse reaction rate. As you learned in introductory chemistry, the forward reaction is biomolecular and in second order. Hence, the v_f, the rate in the forward direction, is proportional to [M][L] or
$v_f = k_f[M][L]$, where k_f is the rate constant in the forward direction. The rate of the reverse reaction, v_r is first order, proportional to [ML], and is given by $v_r = k_r[ML]$, where k_r is the rate constant for the reverse reaction. Notice that the units of k_f are $M^{-1}s^{-1}$, while the unit of k_r are s^{-1}. At equilibrium, $v_f = v_r$, or

$$k_f[M][L] = k_r[ML] \tag{7.5.22}$$

Rearranging the equation gives

$$[ML]/[M][L] = k_f/k_r = K_{eq} \tag{7.5.23}$$

Hence, K_{eq} is given by the ratio of rate constants. For tight binding interactions, $K_{eq} >> 1$, $K_D << 1$, and k_f is very large (in the order of 10^{8-9}), and k_r must be very small (10^{-2} - 10^{-4} s^{-1}).

To get a more intuitive understanding of K_Ds, it is often easier to think about the rate constants that contribute to binding and dissociation. Let us assume that k_r is the rate constant that describes the dissociation reaction. It is often called k_{off}. Likewise, k_f is often called the on rate (k_{on}). It can be shown mathematically that the rate at which two simple molecules associate depends on their radius and effective molecular weight. The maximal rate at which they will associate is the maximal rate at which diffusion will lead them together. Let us assume that the rate at which M and L associate is diffusion-limited. The theoretical k_{on} is about 10^8 $M^{-1}s^{-1}$. Knowing this, the K_D and the fact that $k_{on}/ k_{off} = K_{eq} = 1/K_D$, we can calculate k_{off}, a first-order rate constant.

We can also determine k_{off} experimentally. Imagine the following example. Adjust the concentrations of M and L such that Mo << Lo and Lo>> Kd. Under these conditions of ligand excess, M is entirely in the bound form ML. Now, at t = 0, dilute the solution so that Lo << Kd. The only process here is dissociation, since negligible association can occur given the new condition. If you can measure the biological activity of ML, then you could measure the rate of disappearance of ML with time and get k_{off}. Alternatively, if you could measure M's biological activity, the rate at which activity returns will give you k_{off}.

You will remember from Introductory Chemistry that for a first-order rate constant, the half-life ($t_{1/2}$) of the reaction can be calculated by the expression $k = 0.693/t_{1/2}$. Hence, given k_{off}, you can determine the $t_{1/2}$ for the associated species' existence. How long will a complex of ML last before it dissociates? Given ΔG^o or K_D, and assuming a k_{on} (10^8 $M^{-1}s^{-1}$), you should be able to calculate k_{off} and $t_{1/2}$. Or, you could determine k_{off} experimentally and then calculate $t_{1/2}$. Applying these principles,

you can calculate the binding parameters. Table **7.5.1** below shows calculated k_{off} and $t_{1/2}$ for binary complexes assuming diffusion-controlled k_{on}.

Complex	K_D (M)	k_{off} (s^{-1})	t½
H_2	1×10^{-71}	1×10^{-63}	2×10^{55} yr
RtV3 : Rt'L3(a)	10^{-17}	1×10^{-9}	2 yr
Avidin:biotin	10^{-15}	1×10^{-7}	80 days
thrombin:hirudin(b)	5×10^{-14}	5 x 10-6	2 days
lacrep:DNAoper(c)	1×10^{-13}	1×10^{-5}	0.8 days
Zif268:DNA(d)	10^{-11}	1×10^{-3}	700 s
GroEL:r-lactalbumin(e)	10^{-9}	0.1	7 s
TBP:TATA(f)	2×10^{-9}	2 x 10-1	3 s
TBP:TBP	4×10^{-9}	4×10^{-1}	2 s
LDH (pig): NADH(g)	7.1×10^{-7}(j)	7.1×10^{1}	10 ms
profilin: CaATP-G-actin	1.2 x 10-6	1.2×10^{2}	6 ms
TBP: DNAnonspec(h)	5×10^{-6}	5×10^{2}	1 ms
TCR(i): cyto C peptide	$7X10^{-5}$	$7X10^{3}$	100 us
lacrep:DNAnonspec(h)	1×10^{-4}	$1 X10^{4}$	70 us
uridine-3P: RNase	1.4×10^{-4} (j)	1.4X104	50 us
Creatine Kinase: ADP	8.2x10-4 (j)	$8.2X10^{4}$	10 us
Acetylcholine:Esterase	1.2×10^{-3}	1.2×10^{5}	6 us
no interaction	4×10^{73}	4×10^{81}	-

Table **7.5.1**: Calculated k_{off} and $t_{1/2}$ for binary complexes assuming diffusion-controlled k_{on}

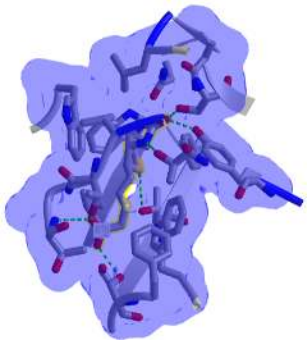
a. Trivalent Vancomycin derivative RtV3 + Trivalent D-Ala-D-Ala deriv, Rt'L3'
b. Hirudin is a potent thrombin inhibitor from leach saliva
c. lac rep is the E. Coli lac operon repressor protein, and DNAoper is the specific DNA binding region in the E. Coli genome that binds to the repressor
d. Zif268 is a mouse zinc-finger binding protein
e. GroEL is a chaperone protein; r-lactalbumin is the reduced form of lactalbumin
f. TBP is the TATA Binding Protein that binds to the TATA box consensus sequence
g. LDH is lactate dehydrogenase
h. DNAnonspec is DNA that does not contain the specific DNA sequence region involved in specific binding to a DNA-binding protein

i. TCR is the T-cell receptor

j. calculated from equation: KD = k_{off}/k_{on}.

What is usually measured is K_D and/or k_{off} (if the k_{off} is reasonable). This analysis is very simplified. Electrostatic forces and other orientation factors may significantly change k_{on}, while conformational changes in the complex may prevent ready unbinding of the bound ligand, dramatically altering k_{off}.

Figure 7.5.8 shows an interactive iCn3D model of one of the tightest binding complexes, avidin and biotin (2avi).



NCBI iCn3D Figure 7.5.8: Avidin-Biotin Complex (2avi) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZssSe5GoUfr9Q6

The blue shows the surface around the complex, which shows that the avidin is completely buried. Hydrogen bonds are shown to biotin (labeled as BTN) in green dashes.

It is important to note that even reactions characterized by high K_D (weak binding) can be specific. Specificity is ultimately defined as a binding interaction between a macromolecule and a ligand that can be co-localized in the same environment and for which a biological function is elaborated upon binding.

Recently added: 9/18/25

7.5.7: Modulators of Enzyme-Ligand Dissociation

Most medical drugs are designed to prevent the binding of a biological ligand to a target protein. The drug often has a structure similar to an endogenous ligand, so both bind to the same site on the protein. It hence decreases the amount of ligand bound to the protein. As mentioned above, it's important that once bound, a ligand or drug dissociates from the protein so that they are not permanently bound, which would leave the system in an unregulated state. When binding to the same site as an endogenous ligand, the drug acts as a "competitive" inhibitor, which decreases the "effective" k_{on} for the normal ligand. The drugs are designed to have a structure similar to the biological ligand or to be designed to be a complementary fit to the binding site for the drug on the protein. But what if the protein does not have a well-defined binding site or pocket? How could you find or design a drug to affect the function of a protein?

An alternative approach would be equally effective: design a drug that binds elsewhere on a protein but which alters the k_{off} of a bound ligand for the protein. This new drug could keep a protein:ligand complex bound longer than normal, by decreasing the effective k_{off} for the bound ligand. These new types of drugs have been called **LOCKTACs,** or "**load and lock targeting chimeras**". These drugs "lock together" the protein and ligand complex or promote their dissociation. That name is similar to another type of "drug" called PROTACs (Proteolysis-Targeting Chimeras), which modify a protein, leading to its proteolysis and loss of the protein's function.

By slowing the dissociation of the ligand, these LOCKTACs could alter the function of the protein in two ways.

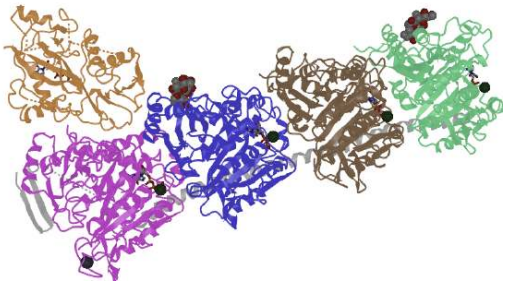
- If the bound complex is active, the drug (an activator, LOCKTAC) could enhance or activate the activity of the protein by keeping the ligand bound longer to the protein.

- If the bound complex is inactive, the drug (an inhibitory LOCKTAC) could inhibit the activity of the protein by keeping the ligand bound longer to the protein.

Here are two examples of activator LOCKTACs.

- Risdiplam: This drug stabilizes the interaction between SMN2 RNA and a splice factor, allowing better inclusion of exon 7 when splicing the gene transcript for the protein SMN2 (the Survival Motor Neuron protein critical for motor neurons that control muscle movement).
- Peloruside: This drug enhances interactions between two proteins (β-catenin with a ligase) and also stabilizes tubulin polymers against dissociation (depolymerization).

Figure 7.5.9 shows an interactive iCn3D models of the Tubulin-Peloruside A complex (4O4J). Peloruside A, an activator LOCKTAC, promotes microtubule (MT) assembly by interacting at the interface between protofilaments and decreasing mobility, both of which maintain function.



NCBI iCn3D Figure 7.5.9: Tubulin-Peloruside A complex (4O4J). (Copyright; author via source).
Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...820c7acf44b5dd

The tubulin monomers are shown in magenta, blue, brown, and green. Two Peloruside A molecules are shown in spacefill.

Here are two examples of inhibitory LOCKTACs (that inhibit function by preventing dissociation).

- Sovilnesib: This drug traps the kinesin motor KIF18A protein on microtubules, preventing it from moving along the microtubules.
- Trametinib: This drug acts to inhibit the activity of the enzyme MEK and a complex of the enzyme, MEK:KSR, where KSR is a "scaffold" protein that presents MEK to a signaling pathway critical for cell growth. This enzyme (a MAP Kinase) is hyperactive in many cancers. It is a protein kinase that uses a bound ATP to phosphorylate and activate downstream proteins in its signaling pathway. An upstream kinase must first phosphorylate MEK before it can phosphorylate other enzymes downstream in the pathway. Trametinib binds at a different (allosteric) site than ATP, so it does not compete with ATP for its binding site, even if ATP levels are high. Trametinib locks MEK in an inactive conformation even as it remains associated with KSR, but does not block MEK:KSR binding. Hence, MEK is locked into an inhibitory form that can not be activated by upstream phosphorylation.

Figure 7.5.10 shows an interactive iCn3D model of the KSR2:MEK1 in complex with AMP-PNP (a non-hydrolyzable ATP analog), and the allosteric MEK inhibitor, Trametinib (7JUR). KSR2 is in magenta and MEKI is in cyan. The ATP analog is shown in sticks, and Trametinib is in spacefill.

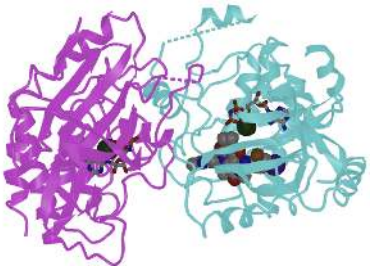


Figure 7.5.10: KSR2:MEK1 in complex with AMP-PNP, and allosteric MEK inhibitor Trametinib (7JUR). (Copyright; author via source).
Click the image for a popup or use this external link:

Trametinib can also be considered an allosteric inhibitor. We will discuss allosterism in the next section.

7.5.8: Summary

Chapter Summary: Reversible Binding of a Ligand to a Macromolecule

This chapter delves into the fundamental principles underlying the reversible, noncovalent binding of ligands to biomacromolecules—a key process that governs the biological activity of proteins, nucleic acids, and other macromolecules.

1. **Fundamentals of Ligand Binding**
   The chapter begins by defining reversible binding and distinguishing it from covalent interactions. Ligands, which can be small molecules, substrates, hormones, neurotransmitters, or metal ions, interact with macromolecules to trigger specific biological functions. The binding process is driven by noncovalent forces such as hydrogen bonds, electrostatic interactions, and van der Waals forces.

2. **Equilibrium and the Dissociation Constant (K_D)**
   A central concept is the dissociation constant, K_D, which quantitatively describes binding affinity. Defined as

   $K_D=[M][L]/[ML]$

   K_D serves as an inverse measure of binding strength—lower K_D values indicate tighter binding. The chapter explains that for many biological systems, K_D values are finely tuned to achieve the appropriate balance between free and bound forms under physiological conditions.

3. **Mathematical Modeling and Mass Balance**
   Using mass balance principles for both the macromolecule (M) and the ligand (L), the text derives key equations. For instance, when the free ligand concentration approximates the total ligand concentration ($L_0 >> M_0$), the bound complex concentration (ML) can be described by:

   $ML=M_0L/(K_D+L)$.

   This hyperbolic relationship illustrates how binding increases with ligand concentration, reaching half-maximal saturation when $L=K_D$. A more general quadratic form is also derived for cases where free ligand concentration is not easily approximated by L_0.

4. **Graphical Analyses of Binding Data**
   The chapter reviews various graphical methods to analyze binding interactions:

   - **Hyperbolic Plots:** Represent ML or fractional saturation (Y) versus free ligand concentration.
   - **Double Reciprocal and Scatchard Plots:** Linear transformations that facilitate approximate determination of K_D.
   - **Semilog Plots (Y vs. log L):** These yield sigmoidal curves that can help reveal binding characteristics over a broad range of ligand concentrations. These plots not only provide a visual representation of binding but also help in extracting binding parameters from experimental data.

5. **Complex Binding Scenarios**
   Beyond simple one-to-one interactions, the chapter discusses:

   - **Dimerization and Oligomerization:** Explaining how macromolecules can self-associate (e.g., forming dimers), which requires modified mass balance and equilibrium expressions.
   - **Multiple Binding Sites:** Addressing situations where a single macromolecule has more than one independent binding site, each characterized by its own K_D. These scenarios illustrate how binding complexity increases in multi-site or self-associating systems.

6. **Kinetics of Binding: On and Off Rates**
   The kinetic aspects of binding are also covered. The forward (association) rate constant (k_{on}) and the reverse (dissociation) rate

constant (k_{off}) combine to define the equilibrium through the relationship

$K_D=k_{off}/k_{on}$.

The discussion includes how experimental measurements of k_{off} and the calculation of the half-life of the complex contribute to our understanding of binding dynamics.

7. **The Binding Continuum and Specificity**
   Finally, the chapter places binding interactions along a continuum from no interaction to covalent bonding, emphasizing that even interactions with high K_D values (weak binding) can be highly specific. The thermodynamic relationship

   $\Delta G^0=-RT\ln K_{eq}$

   is introduced to link binding affinity to free energy changes, reinforcing the concept that specificity in biological systems is determined not solely by binding strength but also by the context of the interaction.

Overall, this chapter provides both a quantitative framework and an intuitive understanding of how ligands interact with macromolecules, setting the stage for further exploration of binding mechanisms, allosteric effects, and the regulation of biological activity.

This page titled 7.5: Binding - The First Step Towards Protein Function is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **5.1: Binding - The First Step Towards Protein Function** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

8: Enzyme Regulation

8.1: Regulation of Enzyme Activity
8.2: Mechanisms of Catalysis
8.3: Control of Enzymatic Activity
8.4: Oxygen-Binding Proteins and Allosterism

8: Enzyme Regulation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8.1: Regulation of Enzyme Activity

Figure **8.1.7** (and 9) also illustrates the effects of two different types of inhibition on the different components of enzyme kinetics. Enzymes can be slowed down or even prevented from catalyzing reactions in many ways including preventing the substrate from entering the active site or preventing the enzyme from altering conformation to catalyze the reaction. The inhibitors that do this can do so either reversibly or irreversibly. The irreversible inhibitors are also called **inactivators**, and either bind to the enzyme with such high afinity as to be virtually irreversible, or they actually form covalent bonds with the enzyme. Reversible inhibitors are generally grouped into two basic types: competitive and non-competitive.

Finasteride (trade names include Propecia and Proscar) is an irreversible inhibitor that binds very tightly to the enzyme 5-a-reductase, used in converting testosterone to dihydrotestosterone. It is used in the treatment of male pattern baldness, benign prostatic hyperplasia, and prostate cancer.

Aspirin is an example of an irreversible inhibitor that actually forms a covalent bond with the enzyme. The aspirin (acetylsalicylic acid) transfers its acetyl group onto a serine residue on cyclooxygenase-2 (COX-2). This stops the production of in ammation-producing prostaglandins and thromboxanes by COX-2.

Methotrexate is a competitive inhibitor of dihydrofolate reductase (DHFR), an enzyme that synthesizes tetrahydrofolate, which is a precursor for purine synthesis, and therefore for DNA and RNA. It has a very similar molecular structure for folic acid, the natural substrate of DHFR. Methotrexate is used as an anti-cancer drug because it affects rapidly reproducing cells (which need to make DNA sooner than other cells) more than non-cancerous cells.

Competitive inhibition is perhaps the simplest to understand. The inhibitor molecule competes directly with the substrate for the active site of an unbound enzyme. If an inhibitor binds to the active site, the substrate is unable to do so until the inhibitor has vacated the site. Thus, one could potentially overwhelm competitive inhibition with sufficiently larger concentrations of substrate so that the probability that the enzyme bumps into a substrate to bind becomes exceeding large compared to the probability of bumping into an inhibitor. Normal, uninhibited V_{max} is then achieved despite the presence of the competitive inhibitor, which has only affected the K_m, that is, the concentration of substrate needed to reach $V_{max}/2$. This is the kinetic signature of competitive inhibitors: with increasing inhibitor concentrations, K_M is increased but V_{max} is unaffected.

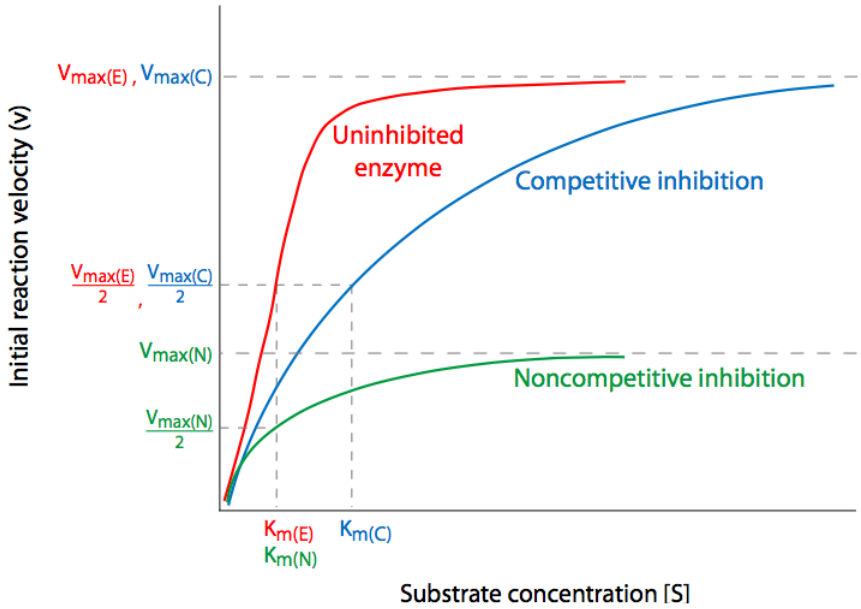


Figure **8.1.9**. Saturation curves for enzyme-catalyzed reactions without inhibitor (red) with competitive inhibitor (blue) at a constant concentration, and with noncompetitive inhibi- tor (green) at a constant concentration.

Non-competitive inhibition involves inhibiting the enzyme by altering its ability to complete the catalyzed reaction through binding of the enzyme at a position that is not the active site. When the inhibitor binds to the enzyme, it causes a change, usually conformational, that may either prevent the enzyme from binding the substrate, or prevent the enzyme from acting upon a bound substrate. In either case, increasing the availability of substrate will not ultimately overcome the effect of the inhibitor. Thus, V_{max} is reduced not because some proportion of the enzymes are no longer usable, but because the enzymes that are available have the same access to substrate as it would without inhibitor (that is, it is not in competition with an inhibitor), the K_m is not affected.

Allostery

Non-competitive regulation is one example of *allosteric regulation* of enzymes. Allosteric interactions occur when the binding of a ligand (not necessarily a substrate) to a protein influences the binding of another ligand to the protein at a separate binding site. These kinds of interactions can be either positive (activating) or negative (inhibitory), and either homotropic (both ligands are identical) or heterotropic (ligands are different). Interestingly, sometimes the regulator ligand may actually be a product of the catalyzed reaction. In this kind of feedback mechanism, the progress of a reaction is self-regulating.

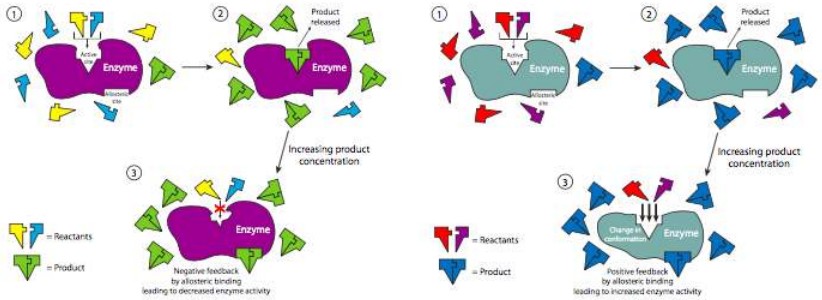


Figure 8.1.10. Feedback control loops. (Left) Negative feedback is a common biological mechanism for self-regulating processes. The product of the reaction, either by itself or inconjunction with other molecules, acts as an allosteric inhibitor of the enzyme. (Right) Positive feedback is less common because it can lead to rapid expansion of the scope of a reaction, and requires an external (relative to the enzymatic reaction) mechanism for slowing or stopping the reaction.

Many of the enzymes in metabolic pathways (chapters 5 and 6) are regulated by a naturally occurring non-competitive inhibitor. One example is phosphofructokinase (PFK), which is involved in glycolysis, which produces ATP for the cell. However, if there are high enough levels of ATP in the cell that other cellular processes aren't using, then it can bind to phosphofructokinase outside of its active site (it binds fructose-6-phosphate) and turn it off. This blocks glycolysis and production of excess ATP when the cell does not need it. As ATP is used up, there is less available to inhibit PFK and glycolysis starts back up.

Modeling Allosteric Behavior

There are two models for such allosteric interactions.

- The **symmetry model**, also known as the concerted model, or MWC model (Monod, Wyman, and Changeux, 1965), proposes that the allosteric enzyme is an oligomer of several subunits, each of which are symmetrically related, and can be in either a "tensed" or "relaxed" state, but all of the subunits are in the same state and at equilibrium. When a ligand binds, it changes the state of the subunit(s) to which it binds, and to maintain equilibrium, that in turn causes the state of the other subunits to match, thus altering binding properties for a subsequent ligand.
- The **sequential model**, or the KNF model (Koshland, Nemethy, and Filmer, 1966), proposes something quite different: although it also supposes subunits in either tensed or relaxed states, it does not require all subunits to be connected in such a way as to mandate that all subunits be in the same state, and therefore conformational changes in one subunit need not be propagated to all. Instead, using an induced-fit model of ligand binding rather than the more rigid basic lock and key mechanism, it suggests that when a ligand binds to the enzyme, it induces a slight conformational change in the active site that increases its affinity for the ligand. The conformational change may slightly alter the conformation of other subunits of the enzyme, but may not constitute a state change between relaxed and tensed depending on how tightly the subunits are interacting. However, the change is enough to increase substrate affinity in adjacent subunits.

Advanced: Uncompetitive and Mixed inhibition

Most cell/molecular courses stop the discussion of enzyme inhibitors at competitive vs non-competitive based on their kinetic profiles. However, it should be noted that if you take a biochemistry course, you may encounter the terms *uncompetitive* inhibitor and *mixed* inhibitor. These terms are defined not just by the enzyme kinetics, but the mechanism of interaction:

- *Uncompetitive inhibitors* bind only to the enzyme-substrate (ES) complex, and not to the enzyme before it has encountered substrate. This leads to decreased V_{max} and decreased K_m.
- *Mixed inhibition* means that the inhibitor can bind to either enzyme alone or the enzyme-substrate complex. Because the affinities of the inhibitor for the two forms of the enzyme are different, and because part of it depends on substrate concentration while the other kind of binding does not, generally, V_{max} decreases and K_m increases. Non-competitive inhibition is a special case of mixed inhibition in which the catalytic activity of the enzyme is diminished or abolished, but the ability to bind substrate is unaltered.

Other Mechanisms of Inhibition

While important, especially pharmaceutically, the use of enzyme inhibitors is not the only way to regulate enzymes. There are numerous examples of one type of enzyme activating or inhibiting another. The most common general example are the protein kinases. These enzymes phosphorylate (transfer phosphate group to) other enzymes and thereby activate them. Kinases are generally fast and very specific, and this is an efficient method for activating large numbers of particular enzymes quickly. Conversely, protein phosphatases are enzymes (also quite fast, but much less specific than kinases) that remove the phosphate groups from phosphorylated proteins, thereby turning off those enzymes. Keep in mind that this is a generalization, and that not all phosphorylations are activating. In addition to enzymatic inhibition of enzymes, there is also inhibition by binding and sequestration of the substrates. In fact, the antibiotic vancomycin works just this way, binding to the substrate peptide for transpeptidase and preventing the enzyme from recognizing it. Transpeptidase normally helps stabilize the cell wall of certain bacteria by altering some of the proteins, and without its activity, the protection of the cell wall is compromised and the bacteria may be more easily killed.

Well over half of the enzymes discovered so far do not act in the simplistic Michaelis-Menten one-substrate-one-product mechanism, but rather operate with two substrates and two products, usually with the transfer of an active group. These types of reactions are sometimes known as Bi Bi reactions. There are two major classes of these reactions: the *sequential reactions*, in which all substrates bind with the enzyme before the reaction proceeds, and the *ping pong reactions*, in which one or more products are created and released before all of the substrates have been bound. In fact, unlike sequential reactions, the two substrates do not interact with one another while bound to the enzyme.

Optimal Conditions

The activity of enzymes is greatly influenced by both pH and temperature, as expected from the discussion of protein structure in the previous chapter. Activity profiles of most enzymes show a peak of activity that tails off on either side, whether it is pH or temperature. This is an innate characteristic of the enzyme. For example, pepsin, a digestive enzyme secreted into the stomach (pH 2) does not function when the pH > 5. On the other hand, another digestive enzyme, trypsin, which is secreted into the duodenum (proximal small intestine) where the pH is ~8, does not work in acidic environments. Changes in pH can change ionization of amino acid side chains that can thereby alter interaction with the substrate, or lead to changes in tertiary structure.

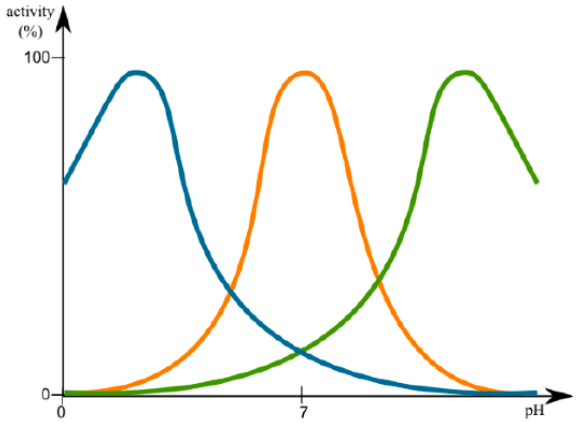


Figure **8.1.11**. pH dependence of enzymatic activity. This graph depicts three hypothetical enzymes with acidic, neutral, and basic pH optima.

Similarly, at suboptimal temperatures, the likelihood of protein-substrate interaction is low but above the optimal temperature, the increased energy can lead to breaking of hydrogen bonds within the structure of the enzyme, resulting in changes that in- activate the catalytic ability of enzyme or prevent it from binding substrate with sufficient affinity. The temperature optimum of most enzymes is very close to its typical environment. Thus, a human enzyme would operate optimally around 37°C, while an enzyme from bacteria that live in deep-sea volcanic vents (e.g. *Thermophilus aquaticus*) might have temperature optima over 90°C. This is one of the reasons that refrigeration can slow down growth of microorganisms (which obviously have no ability to regulate their temperature), and why most microorganisms are killed (enzymes permanently denatured) when put into sustained high temperature environments. Interestingly, the DNA polymerase from the *T. aquaticus* bacteria, also commonly called Taq polymerase, is used in a rapid DNA-amplifying lab technique known as PCR (polymerase chain reaction, see Methods chapter) in which samples are repeatedly heated to high temperatures to separate DNA strands in preparation for making copies of them. DNA polymerases from most prokaryotic or eukaryotic species would be denatured and in- activated by the high heat, but Taq has evolved (with respect to its tertiary structure) for extraordinary structural stability even in heat extremes.

Finally, many enzymes require a molecular partner that has no catalytic activity of its own, but like a catalyst, is not permanently altered by the chemical reaction. These molecules are *cofactors*. Some are simple: elemental, in fact, including metal ions such as Zn^{2+} or Ca^{2+}. Others are slightly more complex: small organic cofactors are called *coenzymes*, and accomplish the same thing, acting as a required partner to the enzyme in catalyzing a reaction. The interaction with the enzyme itself varies and may be only transient, as in NAD^+/NADH which are coenzymes used in redox reactions, or permanently bound to the enzyme by covalent bond like the heme group of hemoglobin. Often the function of the coenzyme is to provide an active group to facilitate the catalyzed reaction. Coenzyme A, in various metabolic pathways such as glycolysis or the tricarboxylic acid cycle, can be bound to a substrate to form a stable product that then acts as an intermediate. The Co-A is released from the molecule as it undergoes the next step in a series of reactions in the metabolic pathway (see Chapter 5).

From a human health standpoint, it is interesting to note that many coenzymes are vitamins, or derived from vitamins. These are the B vitamins biotin (B_7), cobalamin (B_{12}), folic acid (B_3), niacin/nicotinamide (B_9), pantothenic acid (B_5), pyridoxine (B_6), riboflavin (B_2), and thiamine (B_1). Vitamins are small organic compounds that are not synthesized by an organism and must therefore be ingested. They are generally needed only in small quantities, but necessary nonetheless. Naturally, the vitamins we are familiar with are those required by humans. The specific roles of these vitamins and the con- sequences of not having enough of them are discussed later in this textbook, as the enzymes that they work with are introduced in detail.

This page titled 8.1: Regulation of Enzyme Activity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by E. V. Wong via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Regulation of Enzyme Activity** by E. V. Wong is licensed CC BY-NC-SA 3.0. Original source: https://www.axopub.com/wp01/2012/02/26/download-cells-molecules-and-mechanisms/.

8.2: Mechanisms of Catalysis

A printable version of this section is here: BiochemFFA_4_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

The magic of enzymes, as noted, is in their ability to create electronic environments conducive to initiation of a reaction. There are more mechanisms of reaction than we could ever hope to cover in a book like this, and comprehensive discussion of these is not our aim. Instead, we will cite some examples and go into detail on one of them - the mechanism of action of serine proteases.

Chymotrypsin

We will begin with mechanism of action of one enzyme - chymotrypsin. Found in our digestive system, chymotrypsin's catalytic activity is cleaving peptide bonds in proteins and it uses the side chain of a serine in its mechanism of catalysis. Many other protein-cutting enzymes employ a very similar mechanism and they are known collectively as serine proteases (Figure 4.52).

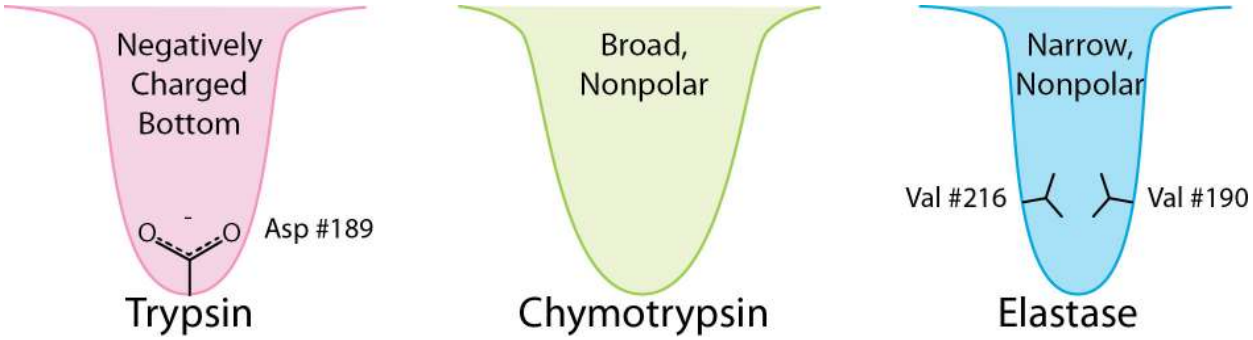


Figure 4.52 - Substrate binding sites (S1 pockets) of three serine proteases. Image by Aleia Kim

These enzymes are found in prokaryotic and eukaryotic cells and all use a common set of three amino acids in the active site called a catalytic triad (Figure 4.53). It consists of aspartic acid, histidine, and serine. The serine is activated in the reaction mechanism to form a nucleophile in these enzymes and gives the class their name. With the exception of the recognition that occurs at the substrate binding site, the mechanism shown here for chymotrypsin would be applicable to any of the serine proteases.

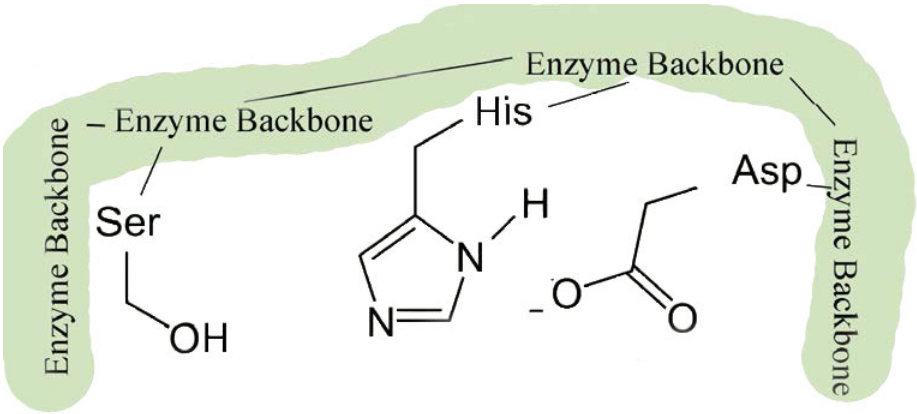


Figure 4.53 - 1. Active site of chymotrypsin showing the catalytic triad of serine - histidine-aspartic acid

Specificity

As a protease, chymotrypsin acts fairly specifically, cutting not all peptide bonds, but only those that are adjacent to relatively non-polar amino acids in the protein. One of the amino acids it cuts adjacent to is phenylalanine. The enzyme's action occurs in two phases – a fast phase that occurs first and a slower phase that follows. The enzyme has a substrate binding site that includes a region of the enzyme known as the S1 pocket. Let us step through the mechanism by which chymotrypsin cuts adjacent to phenylalanine.

Substrate binding

The process starts with the binding of the substrate in the S1 pocket (Figure 4.54). The S1 pocket in chymotrypsin has a hydrophobic hole in which the substrate is bound. Preferred substrates will include amino acid side chains that are bulky and hydrophobic, like phenylalanine. If an ionized side chain, like that of glutamic acid binds in the S1 pocket, it will quickly exit, much like water would avoid an oily interior.

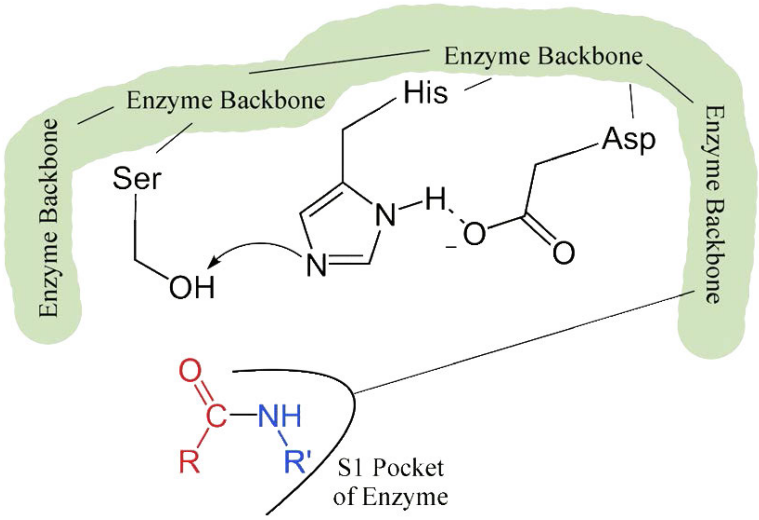


Figure 4.54 - 2. Binding of substrate to S1 pocket in the active site

Shape change on binding

When the proper substrate binds in the S1 pocket, its presence induces an ever so slight change in the shape of the enzyme. This subtle shape change on the binding of the proper substrate starts the steps of the catalysis. Since the catalytic process only starts when the proper substrate binds, this is the reason that the enzyme shows specificity for cutting at specific amino acids in the target protein. Only amino acids with the side chains that interact well with the S1 pocket start the catalytic wheels turning.

The slight changes in shape involve changes in the positioning of three amino acids (aspartic acid, histidine, and serine) in the active site known as the *catalytic triad*.

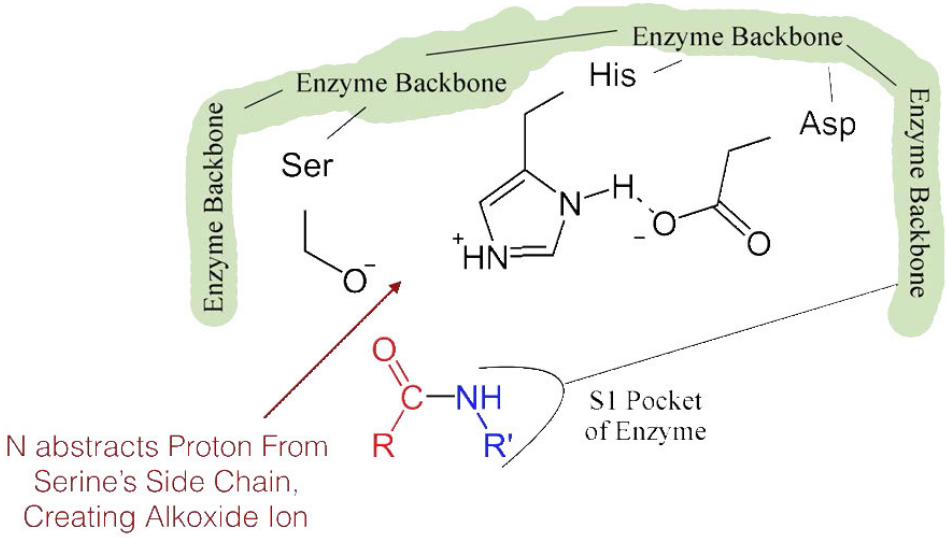


Figure 4.55 - 3. Formation of alkoxide ion

The shift of the negatively charged aspartic acid towards the electron rich histidine ring favors the abstraction of a proton by the histidine from the hydroxyl group on the side chain of serine, resulting in production of a very reactive alkoxide ion in the active

site (Figure 4.55).

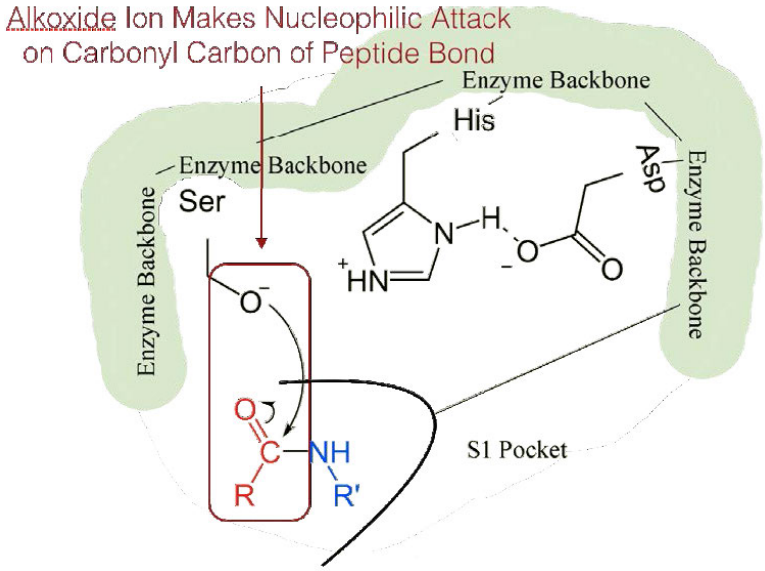


Figure 4.56 - 4. Nucleophilic attack

Since the active site at this point also contains the polypeptide chain positioned with the phenylalanine side chain embedded in the S1 pocket, the alkoxide ion performs a nucleophilic attack on the peptide bond on the carboxyl side of phenylalanine sitting in the S1 pocket (Figure 4.56). This reaction breaks the peptide bond (Figure 4.57) and causes two things to happen.

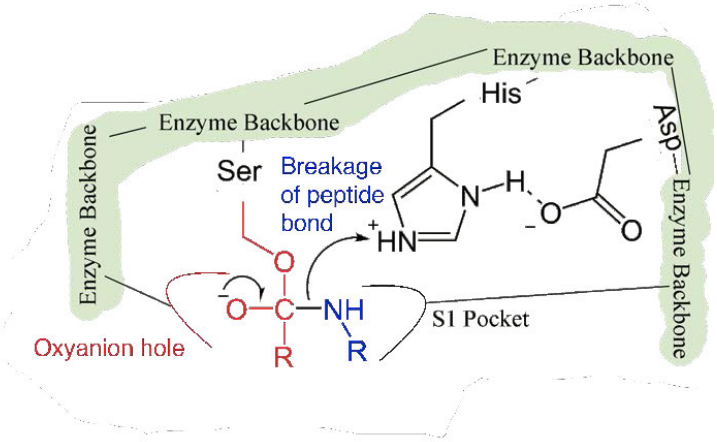


Figure 4.57 - 5. Stabilization by oxyanion hole. Breakage of peptide bond.

First, one end of the original polypeptide is freed and exits the active site (Figure 4.58). The second is that the end containing the phenylalanine is covalently linked to the oxygen of the serine side chain. At this point we have completed the first (fast) phase of the catalysis.

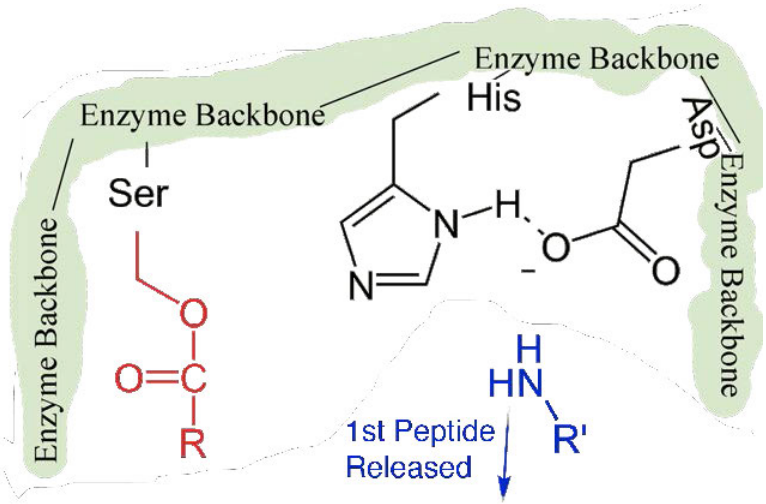


Figure 4.58 - 6. First peptide released. Other half bonded to serine.

Slower second phase

The second phase of the catalysis by chymotrypsin is slower. It requires that the covalent bond between phenylalanine and serine's oxygen be broken so the peptide can be released and the enzyme can return to its original state. The process starts with entry of water into the active site. Water is attacked in a fashion similar to that of the serine side chain in the first phase, creating a reactive hydroxyl group (Figure 4.59) that performs a nucleophilic attack on the phenylalanine-serine bond (Figure 4.60), releasing it and replacing the proton on serine. The second peptide is released in the process and the reaction is complete with the enzyme back in its original state (Figure 4.61).

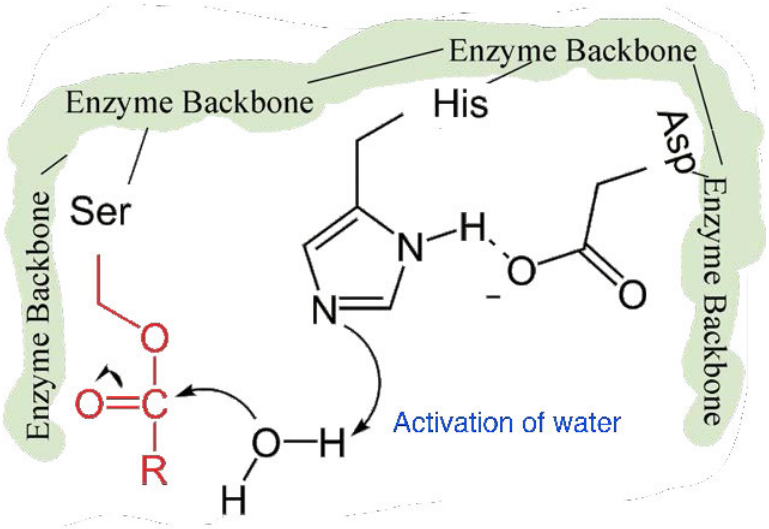


Figure 4.59 - 7. Activation of water by histidine

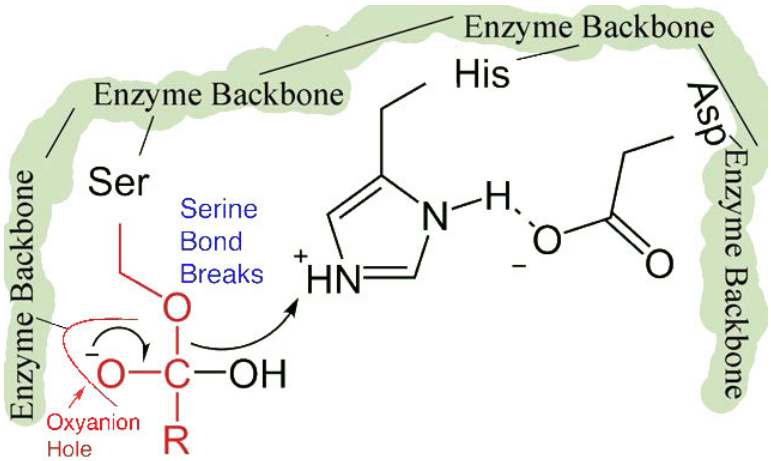


Figure 4.60 - 8. Nucleophile attack by hydroxyl creates tetrahydryl intermediate stabilized by oxyanion hole. Bond to serine breaks.

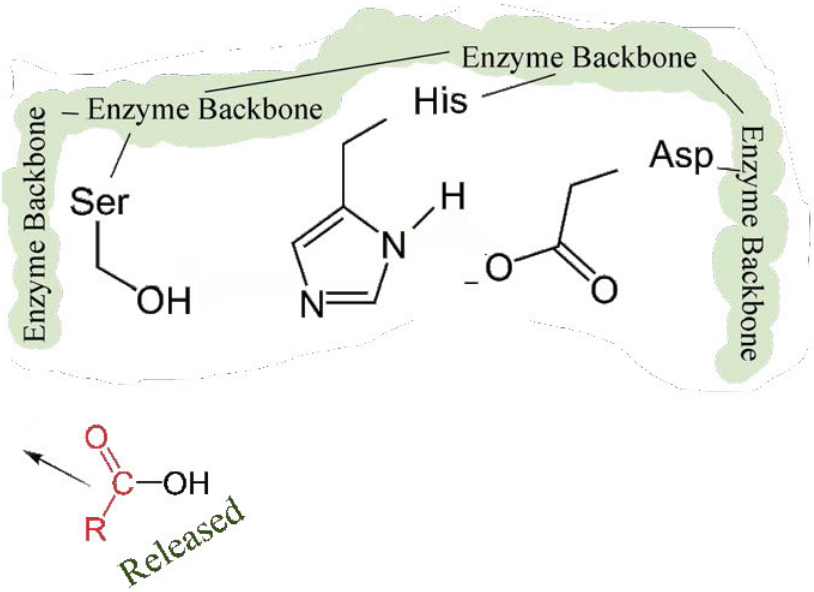


Figure 4.61 - 9. Second half of peptide released. Enzyme active site restored.

Serine proteases

The list of serine proteases is quite long. They are grouped in two broad categories - 1) those that are chymotrypsin-like and 2) those that are subtilisin-like. Though subtilisin-type and chymotrypsin-like enzymes use the same mechanism of action, including the catalytic triad, the enzymes are otherwise not related to each other by sequence and appear to have evolved independently. They are, thus, an example of convergent evolution - a process where evolution of different forms converge on a structure to provide a common function.

The serine protease enzymes cut adjacent to specific amino acids and the specificity is determined by the size/shape/charge of amino acid side chain that fits into the enzyme's S1 binding pocket (Figure 4.62).

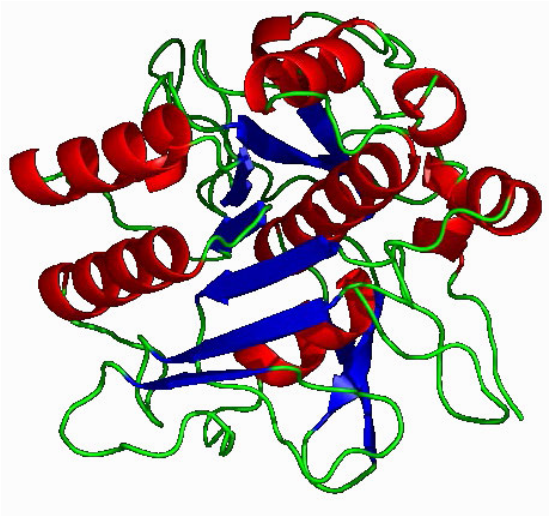


Figure 4.62 - Subtilisin - A serine protease

Examples of serine proteases include trypsin, chymotrypsin, elastase, subtilisin, signal peptidase I, and nucleoporin. Serine proteases participate in many physiological processes, including blood coagulation, digestion, reproduction, and the immune response.

Cysteine proteases

Cysteine proteases (also known as thiol proteases) catalyze the breakdown of proteins by cleaving peptide bonds using a nucleophilic thiol from a cysteine (Figure 4.63). The cysteine is typically found in a catalytic dyad or triad also involving histidine and (sometimes) aspartic acid (very much like serine proteases). The sulfhydryl group of cysteine proteases is more acidic than the hydroxyl of serine proteases, so the aspartic acid of the triad is not always needed.

The mechanism of action is very similar to that of serine proteases. Binding of proper substrate results in activation of the thiol (removal of the proton by the histidine group). The activated thiol acts as a nucleophile, attacking the peptide bond and causing it break. One peptide is released and the other peptide becomes covalently linked to the sulfur. Hydrolysis by water releases the second peptide and completes the cycle. Examples of cysteine proteases include papain, caspases, hedgehog protein, calpain, and cathepsin K.

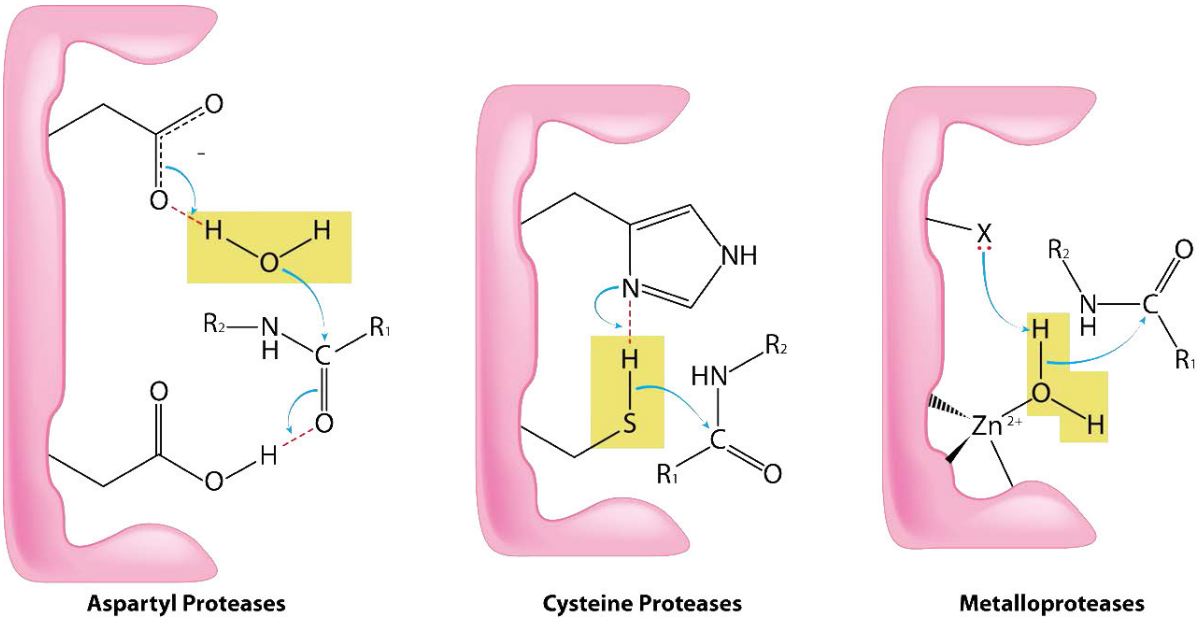


Figure 4.63 - Mechanism of action of proteases. In each case, a nucleophile is created - hydroxyl (aspartyl proteases), thiol (cysteine proteases), and hydroxyl (metalloproteases). Image by Aleia Kim

Caspases

Caspases (Cysteine-ASPartic ProteASEs) are a family of cysteine proteases that play important roles in the body. At the cellular level they function in apoptosis and necrosis and in the body, they are involved in inflammation and the immune system. Maturation of lymphocytes is one such role. They are best known, however, for their role in apoptosis, which has given rise to descriptions of them as “executioner” proteins or “suicide proteases” that dismantle cells in programmed cell death.

There are 12 known human caspases. The enzymes are synthesized as pro-caspase zymogens with a prodomain and two other subunits. The prodomain contains regions that allow it to interact with other molecules that regulate the enzyme’s activity. The caspases come in two forms. The initiator caspases, when activated, activate the effector caspases. The effector caspases cleave other proteins in the cell. Targets for effector caspase cleavage action include the nuclear lamins (fibrous proteins providing structural integrity to the nucleus), ICAD/DFF45 (an inhibitor of DNAse), PARP (flags areas where DNA repair needed), and PAK2 (apoptotic regulation).

The caspase activation cascade can itself be activated by granzyme B (a serine protease secreted by natural killer cells and cytotoxic T-cells), cellular death receptors, and the apoptosome (large protein structure in apoptotic cells stimulated by release of cytochrome C from the mitochondria). Each of these activators is responsible for activating different groups of caspases.

Metalloproteases

Metalloproteases (Figure 4.64) are enzymes whose catalytic mechanism for breaking peptide bonds involves a metal. Most metalloproteases use zinc as their metal, but a few use cobalt, coordinated to the protein by three amino acid residues with a labile water at the fourth position. A variety of side chains are used - histidine, aspartate, glutamate, arginine, and lysine. The water is the target of action of the metal which, upon binding of the proper substrate, abstracts a proton to create a nucleophilic hydroxyl group that attacks the peptide bond, cleaving it (Figure 4.64). Since the nucleophile here is not attached covalently to the enzyme, neither of the cleaved peptides ends up attached to the enzyme during the catalytic process. Examples of metalloproteases include carboxypeptidases, aminopeptidases, insulinases and thermolysin.

Figure 4.64 - Carboxypeptidase - A metalloprotease

Aspartyl proteases

As the name suggests, aspartyl proteases use aspartic acid in their catalytic mechanism (Figures 4.63 & 4.65). Like the metalloproteases, aspartyl proteases activate a water to create a nucleophile for catalysis (Figure 4.65). The activated water attacks the peptide bond of the bound substrate and releases the two pieces without the need to release a bound intermediate, since water is not covalently attached to the enzyme. Common aspartyl proteases include pepsin, signal peptidase II, and HIV-1 protease.

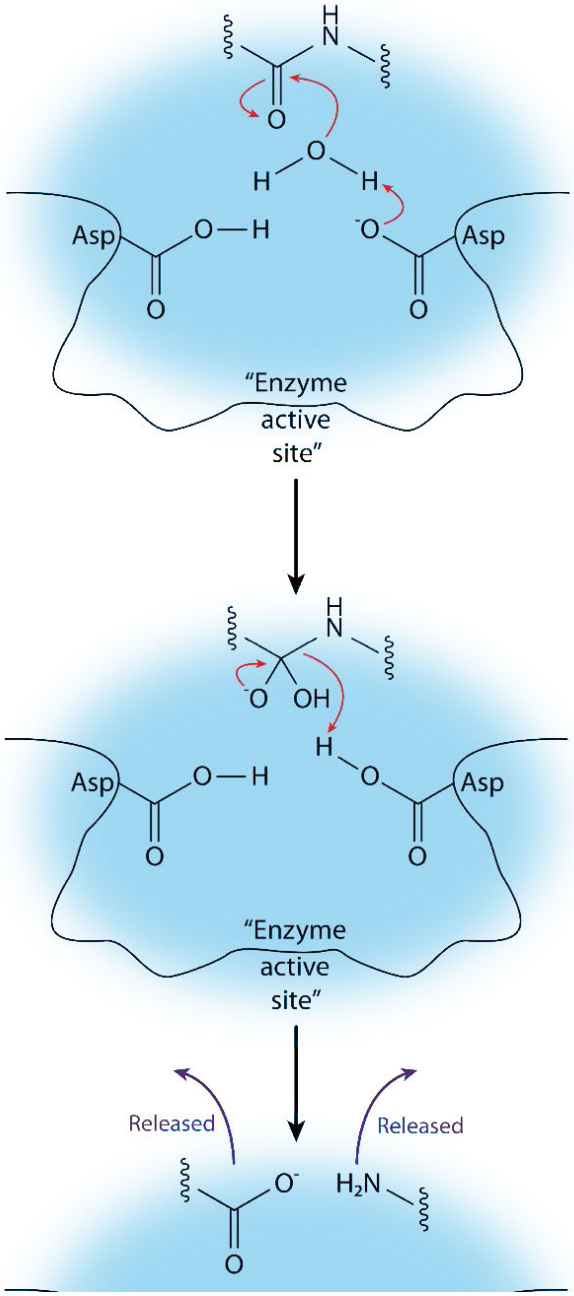


Figure 4.65 - Activation of an aspartyl protease - aspartate side chain abstracts proton (top). Hydroxyl attacks peptide bond (middle). Broken peptide pieces released (bottom). Image by Pehr Jacobson

Threonine proteases

Though threonine has an R-group with a hydroxyl like serine, the mechanism of action of this class of proteases differs somewhat from the serine proteases. There are some similarities. First, the threonine's hydroxyl plays a role in catalysis and that is to act as a nucleophile. The nucleophile is created, however, not by a catalytic triad, but rather as a result of threonine's own α-amine group abstracting a proton.

Because of this, the nucleophilic threonine in a threonine protease must be at the n-terminus of the enzyme. Nucleophilic attack of the peptide bond in the target protease results in breakage of the bond to release one peptide and the other is covalently attached to serine, like the serine proteases. Also, as with the serine proteases, water must come in to release the covalently linked second peptide to conclude the catalytic mechanism.

Examples

Examples of threonine proteases include the catalytic subunits of the proteasome. Some acyl transferases (such as ornithine acyltransferase) have evolved the same catalytic mechanism by convergent evolution. The latter enzymes use ornithine instead of water to break the enzyme-substrate covalent bond, with the result that the acyl-group becomes attached to ornithine, instead of water.

Protease inhibitors

Molecules which inhibit the catalytic action of proteases are known as protease inhibitors. These come in a variety of forms and have biological and medicinal uses. Many biological inhibitors are proteins themselves. Protease inhibitors can act in several ways, including as a suicide inhibitor, a transition state inhibitor, a denaturant, and as a chelating agent. Some work only on specific classes of enzymes. For example, most known aspartyl proteases are inhibited by pepstatin. Metalloproteases are sensitive to anything that removes the metal they require for catalysis. Zinc-containing metalloproteases, for example, are very sensitive to EDTA, which chelates the zinc ion.

One category of proteinaceous protease inhibitors is known as the serpins. Serpins inhibit serine proteases that act like chymotrypsin. 36 of them are known in humans.

Serpins are unusual in acting by binding to a target protease irreversibly and undergoing a conformational change to alter the active site of its target. Other protease inhibitors act as competitive inhibitors that block the active site.

Serpins can be broad in their specificity. Some, for example, can block the activity of cysteine proteases. One of the best known biological serpins is α-1-anti-trypsin (A1AT - Figure 4.66) because of its role in lungs, where it functions to inhibit the elastase protease. Deficiency of A1AT leads to emphysema. This can arise as a result of genetic deficiency or by cigarette smoking. Reactive oxygen species produced by cigarette smoking can oxidize a critical methionine residue (#358 of the processed form) in A1AT, rendering it unable to inhibit elastase. Uninhibited, elastase can attack lung tissue and cause emphysema. Most serpins work extracellularly. In blood, for example, serpins like antithrombin can help to regulate the clotting process.

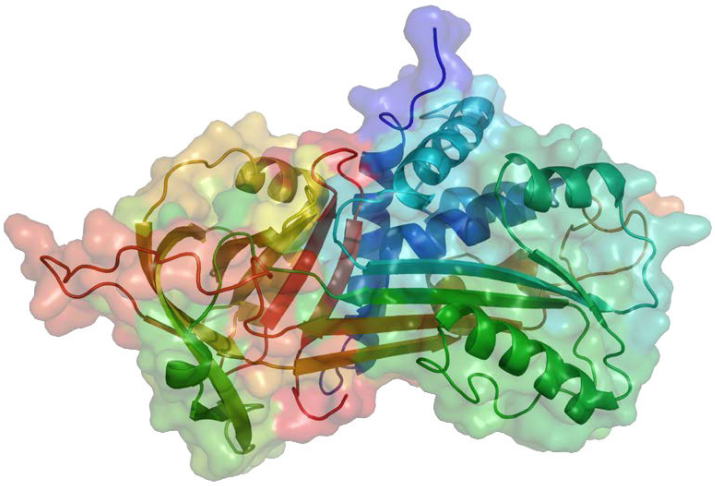


Figure 4.66 - α-1-antitrypsin

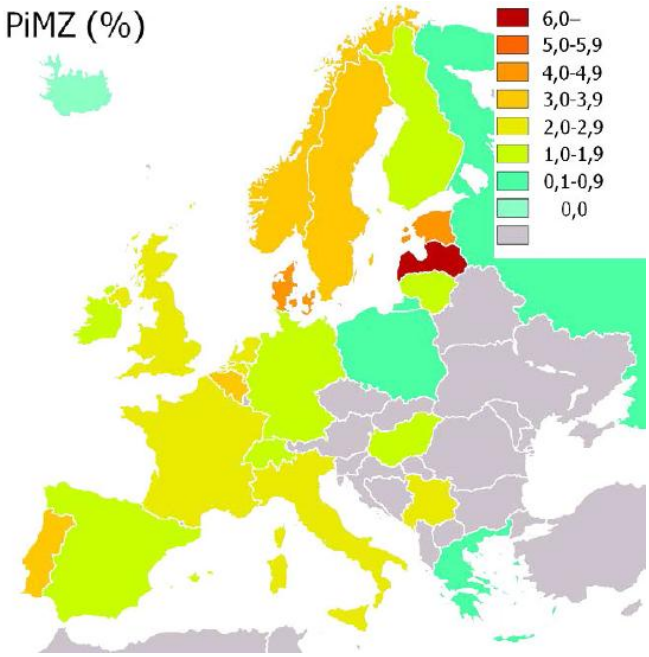


Figure 4.67 - Incidence of α-1-antitrypsin (PiMZ) deficiency in Europe by percent. Wikipedia

Anti-viral Agents

Protease inhibitors are used as anti-viral agents to prohibit maturation of viral proteins - commonly viral coat proteins.

They are part of drug “cocktails” used to inhibit the spread of HIV in the body and are also used to treat other viral infections, including hepatitis C. They have also been investigated for use in treatment of malaria and may have some application in anti-cancer therapies as well.

This page titled 8.2: Mechanisms of Catalysis is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **4.3: Mechanisms of Catalysis** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **7.3: Mechanisms of Catalysis** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

8.3: Control of Enzymatic Activity

A printable version of this section is here: BiochemFFA_4_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Regulation of enzyme activity

Apart from their ability to greatly speed the rates of chemical reactions in cells, enzymes have another property that makes them valuable. This property is that their activity can be regulated, allowing them to be activated and inactivated, as necessary. This is tremendously important in maintaining homeostasis, permitting cells to respond in controlled ways to changes in both internal and external conditions.

Inhibition of specific enzymes by drugs can also be medically useful. Understanding the mechanisms that control enzyme activity is, therefore, of considerable importance.

Inhibition

We will first discuss four types of enzyme inhibition – competitive, non-competitive, uncompetitive, and suicide inhibition. Of these, the first three types are reversible. The last one, suicide inhibition, is not.

Competitive inhibition

Probably the easiest type of enzyme inhibition to understand is competitive inhibition and it is the one most commonly exploited pharmaceutically. Molecules that are competitive inhibitors of enzymes resemble one of the normal substrates of an enzyme. An example is methotrexate, which resembles the folate substrate of the enzyme dihydrofolate reductase (DHFR). This enzyme normally catalyzes the reduction of folate, an important reaction in the metabolism of nucleotides.

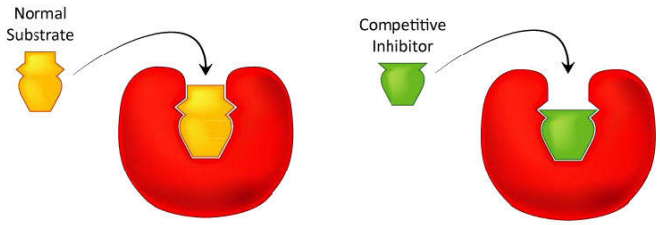


Figure 4.33 - Competitive inhibitors resemble the normal substrate and compete for binding at the active site. Image by Aleia Kim

Inhibitor binding

When the drug methotrexate is present, some of the DHFR enzyme binds to it, instead of to folate, and during the time methotrexate is bound, the enzyme is inactive and unable to bind folate. Thus, the enzyme is inhibited. Notably, the binding site on DHFR for methotrexate is the active site, the same place that folate would normally bind. As a result, methotrexate 'competes' with folate for binding to the enzyme. The more methotrexate there is, the more effectively it competes with folate for the enzyme's active site. Conversely, the more folate there is, the less of an effect methotrexate has on the enzyme because folate outcompetes it.

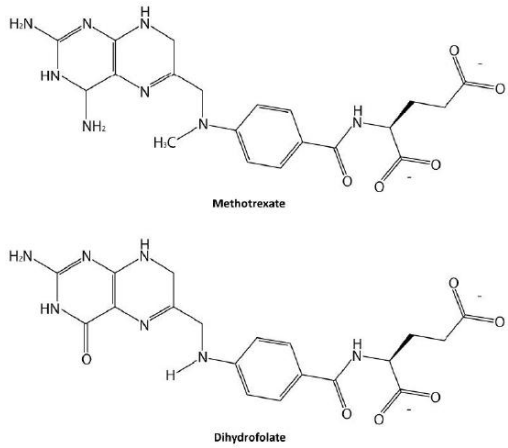


Figure 4.34 - Methotrexate and dihydrofolate. Image by Ben Carson

No effect on Vmax

How do we study competitive inhibition? It is typically done as follows. First, one performs a set of V0 vs. [S] reactions without inhibitor (20 or so tubes, with buffer and constant amounts of enzyme, varying amounts of substrate, equal reaction times). V0 vs. [S] is plotted (Figure 4.35 red line), as well as 1/V0 vs. 1/[S] (Figure 4.36 green line). Next, a second set of reactions is performed in the same manner as before, except that a fixed amount of the methotrexate inhibitor is added to each tube. At low concentrations of substrate, the methotrexate competes for the enzyme effectively, but at high concentrations of substrate, the inhibitor will have a much reduced effect, since the substrate outcompetes it, due to its higher concentration (remember that the inhibitor is at fixed concentration).

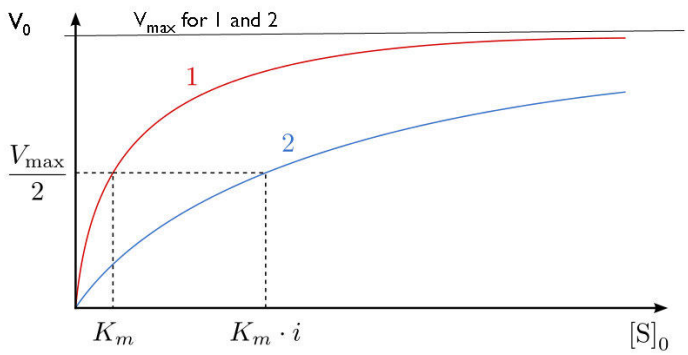


Figure 4.35 V0 vs [S] Plots for uninhibited reactions (red) and competitively inhibited reactions (blue). Both ultimately have same Vmax

Graphically, the results of these inhibitor experiments are shown in Figure 4.35 (blue line) and Figure 4.36 (orange line). Notice that at high substrate concentrations, the competitive inhibitor has essentially no effect, causing the V_{max} for the enzyme to remain unchanged. To reiterate, this is due to the fact that at high substrate concentrations, the inhibitor doesn't compete well. However, at lower substrate concentrations, it does.

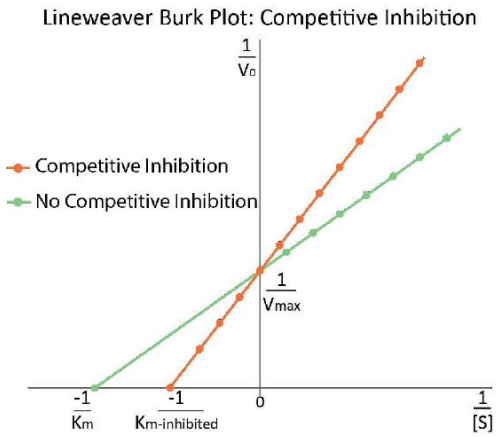


Figure 4.36 - Lineweaver-Burk plots - uninhibited reactions (green). Competitively inhibited reactions (orange). Lines cross on Y-axis at 1/Vmax Since Vmax is the same for both reactions. Image by Aleia Kim

Increased K_m

In competitively inhibited reactions, the apparent Km of the enzyme for the substrate increases ($-1/K_m$ gets closer to zero - red line in Figure 4.36) when the inhibitor is present compared to when the inhibitor is absent, thus illustrating the better competition of the inhibitor at lower substrate concentrations. It may not be obvious why we call the changed Km the apparent Km of the enzyme. The reason is that the inhibitor doesn't actually change the enzyme's affinity for the folate substrate. It only appears to do so. This is because of the way that competitive inhibition works. When the competitive inhibitor binds the enzyme, it is effectively 'taken out of action.' Inactive enzymes have NO affinity for substrate and no activity either. We can't measure Km for an inactive enzyme.

The enzyme molecules that are not bound by methotrexate can, in fact, bind folate and are active. Methotrexate has no effect on them and their Km values are unchanged. Why then, does Km appear higher in the presence of a competitive inhibitor? The reason is that the competitive inhibitor is having a greater effect of reducing the amount of active enzyme at lower concentrations of substrate than it does at higher concentrations of substrate. When the amount of enzyme is reduced, one must have more substrate to supply the reduced amount of enzyme sufficiently to get to Vmax/2.

It is worth noting that in competitive inhibition, the percentage of inactive enzyme changes drastically over the range of [S] values used. To start, at low [S] values, the greatest percentage of the enzyme is inhibited. At high [S], no significant percentage of enzyme is inhibited. This is not always the case, as we shall see in non-competitive inhibition.

Non-competitive inhibition

A second type of inhibition employs inhibitors that do not resemble the substrate and bind not to the active site, but rather to a separate site on the enzyme (Figure 4.37). The effect of binding a non-competitive inhibitor is significantly different from binding a competitive inhibitor because there is no competition. In the case of competitive inhibition, the effect of the inhibitor could be reduced and eventually overwhelmed with increasing amounts of substrate. This was because increasing substrate made increasing percentages of the enzyme active. With non-competitive inhibition, increasing the amount of substrate has no effect on the percentage of enzyme that is active. Indeed, in non-competitive inhibition, the percentage of enzyme inhibited remains the same through all ranges of [S].

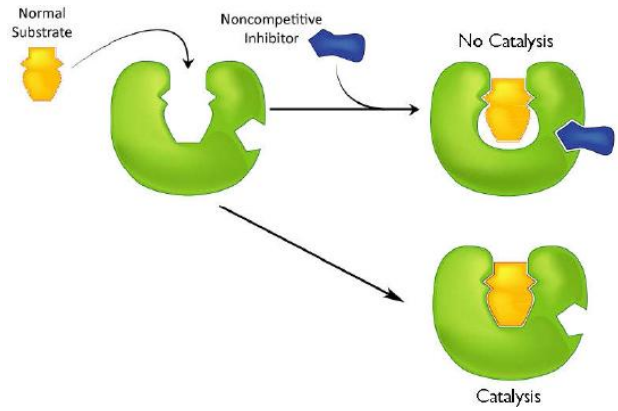


Figure 4.37 - Non-competitive inhibition - inhibitor does not resemble the substrate and binds to a site other than the active site. Image by Aleia Kim

This means, then, that non-competitive inhibition effectively reduces the amount of enzyme by the same fixed amount in a typical experiment at every substrate concentration used The effect of this inhibition is shown in Figure 4.38 & 4.39. As you can see, V_{max} is reduced in non-competitive inhibition compared to uninhibited reactions.

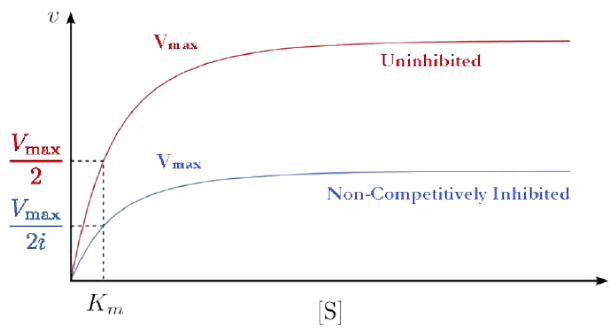


Figure 4.38 - V0 vs [S] plots of uninhibited reactions (red) and non-competitively inhibited reactions (blue). Vmax is reduced, but Km values are unchanged in non-competitively inhibited reactions

This makes sense if we remember that Vmax is dependent on the amount of enzyme present. Reducing the amount of enzyme present reduces V_{max}. In competitive inhibition, this doesn't occur detectably, because at high substrate concentrations, there is essentially 100% of the enzyme active and the V_{max} appears not to change. Additionally, Km for non-competitively inhibited reactions does not change from that of uninhibited reactions. This is because, as noted previously, one can only measure the K_m of active enzymes and K_m is a constant for a given enzyme.

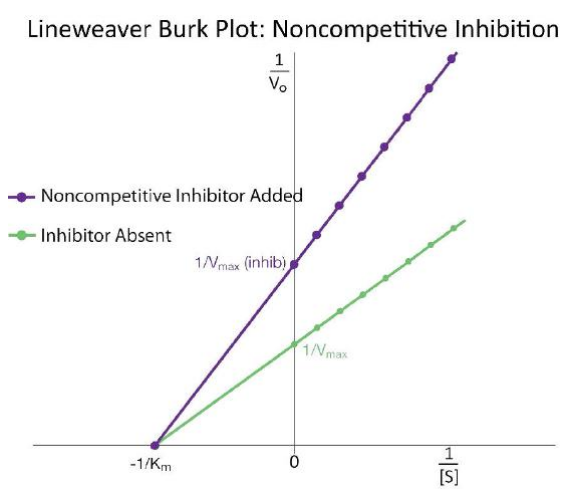


Figure 4.39 - Lineweaver Burk plots of uninhibited reactions (green) and non-competitively inhibited reactions (purple). Image by Aleia Kim

Uncompetitive inhibition

A third type of enzymatic inhibition is that of uncompetitive inhibition, which has the odd property of a reduced Vmax as well as a reduced Km. The explanation for these seemingly odd results is rooted in the fact that the uncompetitive inhibitor binds only to the enzyme-substrate (ES) complex (Figure 4.40). The inhibitor-bound complex forms mostly under concentrations of high substrate and the ES-I complex cannot release product while the inhibitor is bound, thus explaining the reduced V_{max}.

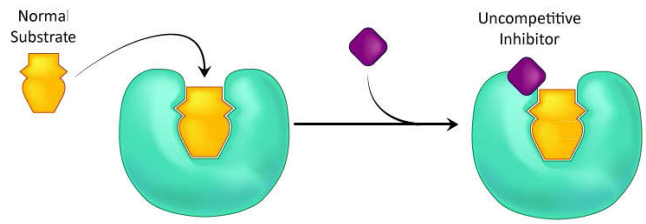


Figure 4.40 - Uncompetitive inhibition. Image by Aleia Kim

The reduced Km is a bit harder to conceptualize. The reason is that the inhibitor-bound complex effectively reduces the concentration of the ES complex. By Le Chatelier's Principle, a shift occurs to form additional ES complex, resulting in less free enzyme and more enzyme in the forms ES and ESI (ES with inhibitor). Decreases in free enzyme correspond to an enzyme with greater affinity for its substrate. Thus, paradoxically, uncompetitive inhibition both decreases V_{max} and increases an enzyme's affinity for its substrate (K_m - Figures 4.41 & 4.42).

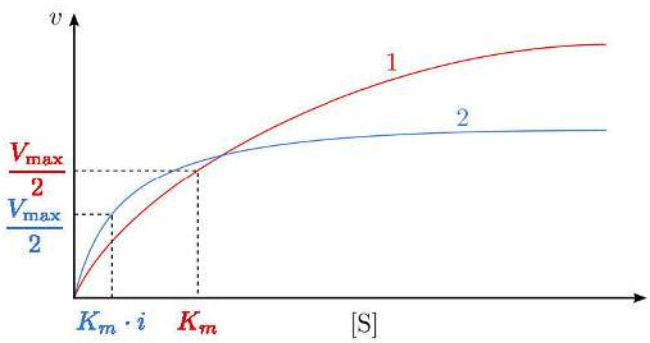


Figure 4.41 - V0 vs [S] plot for uncompetitive inhibition (blue) and uninhibited reactions (red)

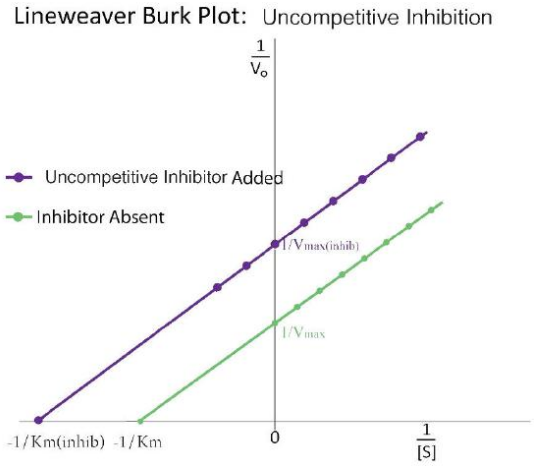


Figure 4.42 - Uncompetitive inhibition (purple) and uninhibited reactions (green). Image by Aleia Kim

Suicide inhibition

In contrast to the first three types of inhibition, which involve reversible binding of the inhibitor to the enzyme, suicide inhibition is irreversible, because the inhibitor becomes covalently bound to the enzyme during the inhibition. Suicide inhibition rather closely resembles competitive inhibition because the inhibitor generally resembles the substrate and binds to the active site of the enzyme.

The primary difference is that the suicide inhibitor is chemically reactive in the active site and makes a bond with it that precludes its removal. Such a mechanism is that employed by penicillin (Figure 4.43), which covalently links to the bacterial enzyme, DD transpeptidase and stops it from functioning. Since the normal function of the enzyme is to make a bond necessary for the peptidoglycan complex of the bacterial cell wall, the cell wall cannot properly form and bacteria cannot reproduce.

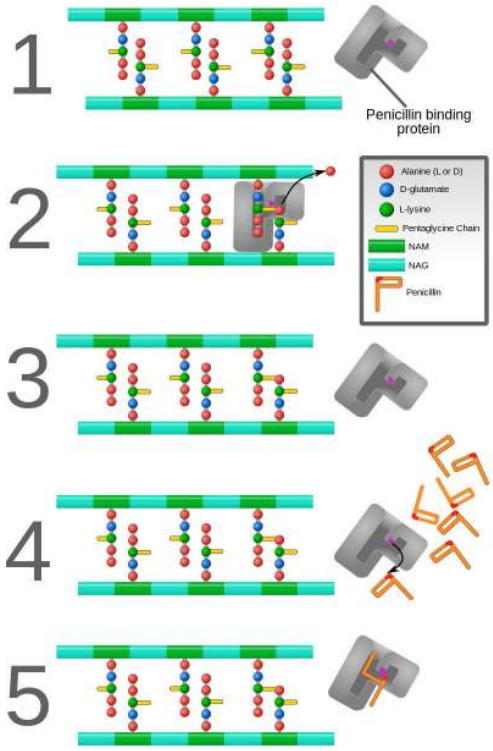


Figure 4.43 - Action of penicillin. DD-transpeptidase builds peptidoglycan layer of bacterial cell wall (1-3). Binding of penicillin by DD-transpeptidase stops peptidoglycan synthesis (4-5). Wikipedia

Control of enzymes

It is appropriate to talk at this point about mechanisms cells use to control enzymes. There are four general methods that are employed:

1. allosterism,
2. covalent modification,
3. access to substrate, and
4. control of enzyme synthesis/breakdown.

Some enzymes are controlled by more than one of these methods.

Allosterism

The term allosterism refers to the fact that the activity of certain enzymes can be affected by the binding of small molecules. Molecules causing allosteric effects come in two classifications. Ones that are substrates for the enzymes they affect are called **homotropic effectors** and those that are not substrates are called heterotropic effectors.

The homotropic effectors usually are activators of the enzymes they bind to and the results of their action can be seen in the conversion of the hyperbolic curve typical of a V0 vs. [S] plot for an enzyme (Figure 4.18), being converted to a sigmoidal plot (Figure 4.44). This is due to the conversion of the enzyme from the T-state to the R-state on binding the substrate/homotropic effector.

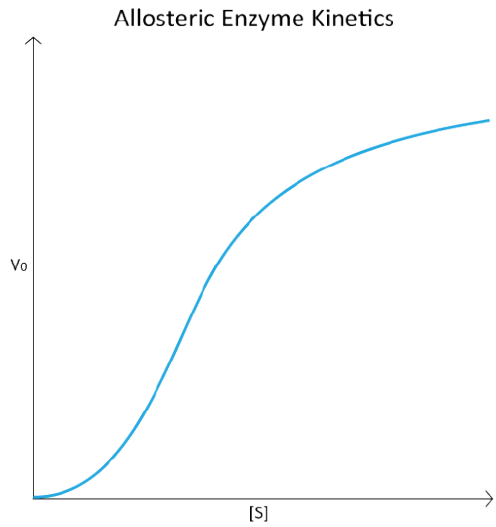


Figure 4.44 - Kinetic profile of an allosteric enzyme whose activity is controlled by a homotropic effector. Image by Aleia Kim

The V0 vs. [S] plot of allosteric enzyme reactions resembles the oxygen binding curve of hemoglobin (see Figure 2.83). Even though hemoglobin is not an enzyme and is thus not catalyzing a reaction, the similarity of the plots is not coincidental. In both cases, the binding of an external molecule is being measured – directly, in the hemoglobin plot, and indirectly by the V0 vs. [S] plot, since substrate binding is a factor in enzyme reaction velocity.

Allosteric inhibition

Allosterically, regulation of these enzymes works by inducing different physical states (shapes, as it were) that affect their ability to bind to substrate. When an enzyme is inhibited by binding an effector, it is converted to the T-state (T=tight), it has a reduced affinity for substrate and it is through this means that the reaction is slowed.

Allosteric activation

On the other hand, when an enzyme is activated by effector binding, it converts to the R-state (R=relaxed) and binds substrate much more readily. When no effector is present, the enzyme may be in a mixture of T- and R-states.

Feedback inhibition

An interesting kind of allosteric control is exhibited by HMG-CoA reductase, which catalyzes an important reaction in the pathway leading to the synthesis of cholesterol. Binding of cholesterol to the enzyme reduces the enzyme's activity significantly. Cholesterol is not a substrate for the enzyme, so it is therefore a heterotropic effector.

Notably, though, cholesterol is the end-product of the pathway that HMG-CoA reductase catalyzes a reaction in. When enzymes are inhibited by an end-product of the pathway in which they participate, they are said to exhibit feedback inhibition.

Feedback inhibition always operates by allosterism and further, provides important and efficient control of an entire pathway. By inhibiting an early enzyme in a pathway, the flow of materials (and ATP hydrolysis required for their processing) for the entire pathway is stopped or reduced, assuming there are not alternate supply methods.

Pathway control

In the cholesterol biosynthesis pathway, stopping this one enzyme has the effect of shutting off (or at least slowing down) the entire pathway. This is significant because after catalysis by HMG-CoA reductase, there are over 20 further reactions necessary to make cholesterol, many of them requiring ATP energy. Shutting down one reactions stops all of them. Another excellent example of allosteric control and feedback inhibition is the enzyme ATCase, discussed below.

ATCase

Another interesting example of allosteric control and feedback inhibition is associated with the enzyme Aspartate Transcarbamoylase (ATCase). This enzyme, which catalyzes a step in the synthesis of pyrimidine nucleotides, has 12 subunits. These include six identical catalytic subunits and six identical regulatory subunits. The catalytic subunits bind to substrate and

catalyze a reaction. The regulatory subunits bind to either ATP or CTP. If they bind to ATP, the enzyme subunits arrange themselves in the R-state.

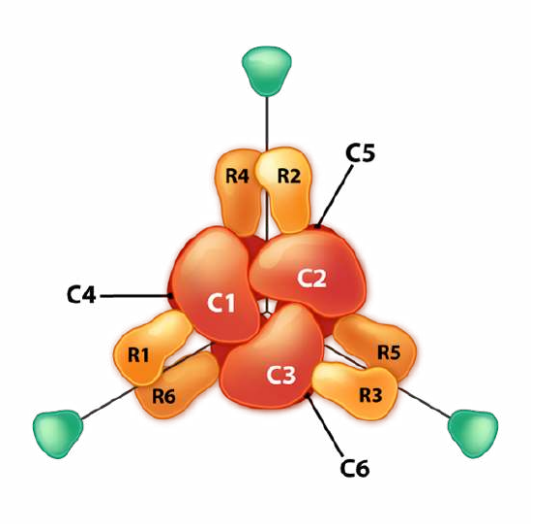


Figure 4.45 - Schematic structure of ATCase. Regulatory Units = R, Catalytic Units = C. Image by Aleia Kim

R-state

The R-state of ATCase allows the substrate to have easier access to the six active sites and the reaction occurs more rapidly. For the same amount of substrate, an enzyme in the R-state will have a higher velocity than the same enzyme that is not in the R-state. By contrast, if the enzyme binds to CTP on one of its regulatory subunits, the subunits will arrange in the T-state and in this form, the substrate will not have easy access to the active sites, resulting in a slower velocity for the same concentration of substrate compared to the R-state. ATCase is interesting in that it can also flip into the R-state when one of the substrates (aspartate) binds to an active site within one of the catalytic subunits.

Aspartate has the effect of activating the catalytic action of the enzyme by favoring the R-state. Thus, aspartate, which is a substrate of the enzyme is a homotropic effector and ATP and CTP, which are not substrates of the enzyme are heterotropic effectors of ATCase.

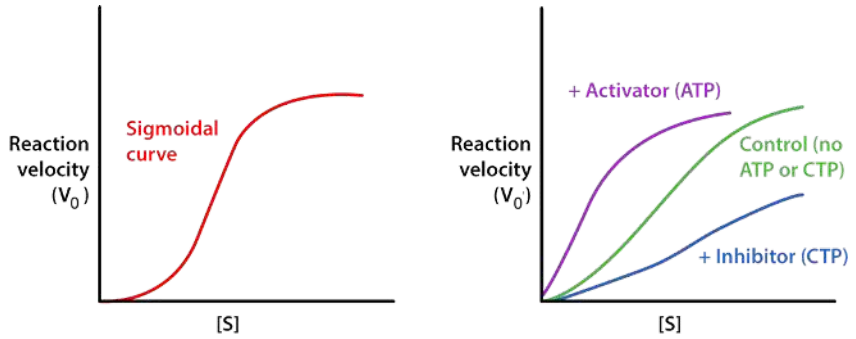


Figure 4.46 - Plots of V0 vs. [S] for ATCase. Left - Allosteric effect of aspartate. Right - Allosteric effects of ATP (activator) and CTP (inhibitor). Image by Pehr Jacobson

Allosteric models

There are three models commonly used to explain how allosterism regulates multi-subunit enzyme activity. They are known as

- the Monod-Wyman-Changeux (MWC) model (also known as the concerted model),
- the sequential model (also known as KNF),
- and the morpheein model.

All models describe a Tense (T) state that is less catalytically active and a Relaxed (R) state that is more catalytically active. The models differ in how the states change.

Sequential model

In the sequential model, binding of an allosteric effector by one subunit causes it to change from T to R state (or vice versa) and that change makes it easier for adjacent subunits to similarly change state. With this model , there is a cause/effect relationship between binding of an effector by one subunit and change of state by an adjacent subunit.

In hemoglobin, for example, binding of one oxygen by one unit of the complex may induce that unit to flip to the R-state and, through interactions with other subunits, cause them to favor adopting the R configuration before they bind to oxygen. In this way, binding of one subunit favors binding of others and cooperativity can be explained by the change in binding affinity as oxygen concentration changes.

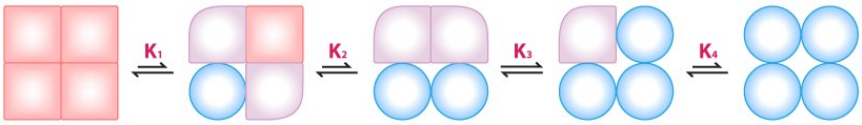


Figure 4.47 - Sequential model of allosteric regulation. Round = R-State. Square = T-state.. Image by Aleia Kim

MWC model

The MWC model is less intuitive. In it, the entire complex changes state from T to R (or vice versa) independently of the binding of effectors. Flipping between T-states and R-states is postulated to be in an equilibrium of states in the absence of effector (for example, a 50 to 1 ratio of T/R. This ratio is referred to as L, so L = T/R). Binding of effector by the enzyme complex has the tendency of “locking” the complex in a state. Binding of inhibitors will increase the ratio of T/R whereas binding of activators will increase R and thus decrease the ration of T/R.

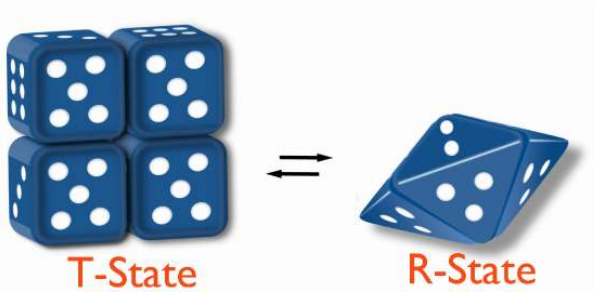


Figure 4.48 - The MWC Model - The multisubunit complex flips as a whole and is in equilibrium. Wikipedia

Morpheein model

The morpheein model is similar to the MWC model, but with an added step of dissociation of the subunits. The MWC model proposes that flipping between R and T states occurs by the complex as a whole and occurs on all units simultaneously. The morpheein model instead proposes that the multi-subunit enzyme breaks down to individual units which can then flip in structure and re-form the complex. In the morpheein model, only identically shaped units (all R, for example) can come together in the complex, thus explaining the “all-R-” or “all-T-” state found in the MWC model.

A large number of enzymes, including prominent ones like citrate synthase, acetyl-CoA carboxylase, glutamate dehydrogenase, ribonucleotide reductase and lactate dehydrogenase have behavior consistent with the morpheein model.

Figure 4.49 - Morpheein model of allosterism. States flip as monomers and aggregate only with same monomer. Wikipedia

Covalent control of enzymes

Some enzymes are synthesized in a completely inactive form and their activation requires covalent bonds in them to be cleaved. Such inactive forms of enzymes are called zymogens. Examples include the proteins involved in blood clotting and proteolytic

enzymes of the digestive system, such as trypsin, chymotrypsin, pepsin, and others.

Synthesizing some enzymes in an inactive form makes very good sense when an enzyme's activity might be harmful to the tissue where it is being made. For example, the painful condition known as pancreatitis arises when digestive enzymes made in the pancreas are activated too soon and end up attacking the pancreas.

Cascades

For both the blood clotting enzymes and the digestive enzymes, the zymogens are activated in a protease cascade. This occurs when activation of one enzyme activates others in a sort of chain reaction. In such a scheme the first enzyme activated proteolytically cleaves the second zymogen, causing it to be activated, which in turn activates a third and this may proceed through several levels of enzymatic action (Figure 4.50).

The advantage of cascades is that they allow a large amount of zymogens to become activated fairly quickly, since there is an amplification of the signal at each level of catalysis.

Zymogens are also abundant in blood. Blood clotting involves polymerization of a protein known as fibrin. Since random formation of fibrin is extremely hazardous because it can block the flow of blood, potentially causing heart attack/stroke, the body synthesizes fibrin as a zymogen (fibrinogen) and its activation results from a "cascade" of activations of proteases that arise when a signal is received from a wound. Similarly, the enzyme catalyzing removal of fibrin clots (plasmin) is also synthesized as a zymogen (plasminogen), since random clot removal would also be hazardous (see below also).

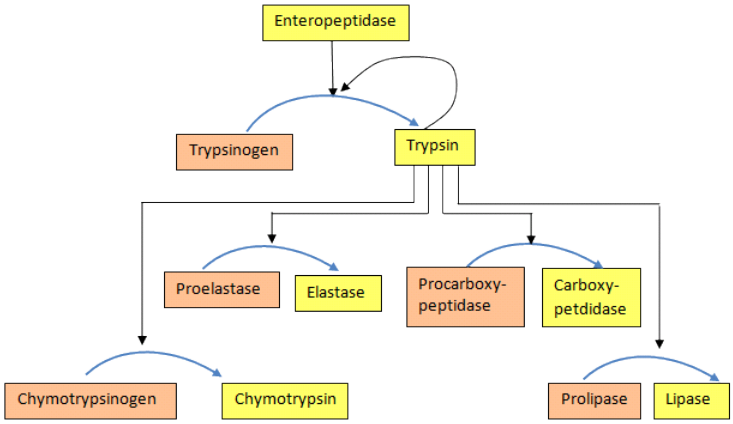


Figure 4.50 - Protease activation scheme. Wikipedia

Phosphorylation/dephosphorylation

Another common mechanism for control of enzyme activity by covalent modification is phosphorylation. The phosphorylation of enzymes (on the side chains of serine, threonine or tyrosine residues) is carried out by protein kinases. Enzymes activated by phosphorylation can be regulated by the addition of phosphate groups by kinases or their removal by phosphatases. Thus, this type of covalent modification is readily reversible, in contrast to proteolytic cleavage.

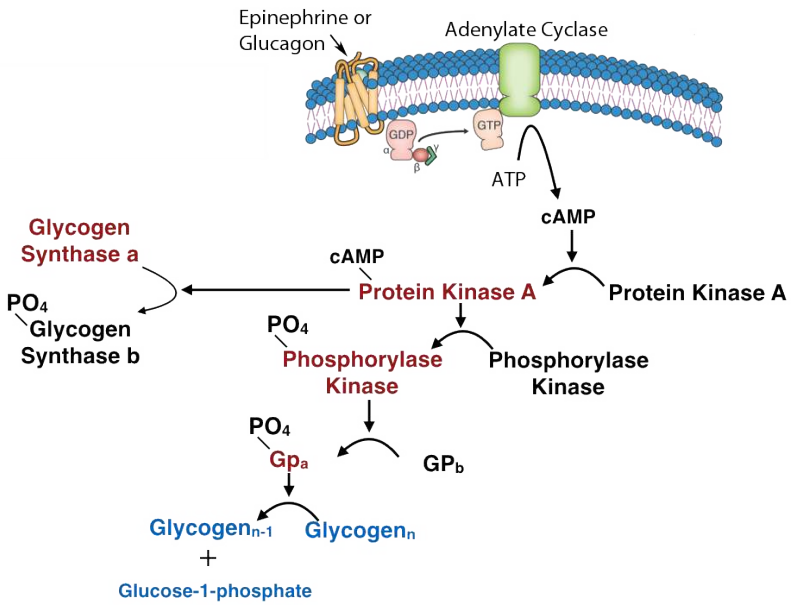


Figure 4.51 - Regulation by covalent modification of glycogen catabolism enzymes

Reduction/oxidation

An interesting covalent control of enzymes using reduction/oxidation is exhibited in photosynthetic plants. In the light phase of photosynthesis, electrons are excited by light and flow through carriers to NADP+, forming NADPH. Thus, in the light, the NADPH concentration is high. When NADPH concentration is high, the concentration of reduced ferredoxin (a molecule donating electrons to NADP+) is also high.

Reduced ferredoxin can transfer electrons to thioredoxin, reducing it. Reduced thioredoxin can, in turn, transfer electrons to proteins to reduce their disulfide bonds. Four enzymes related to the Calvin cycle can receive electrons from thioredoxin and become activated, as a result.

These include sedoheptulose 1,7-bisphosphatase, ribulose-5-phosphate kinase, fructose 1,6-bisphosphatase, and glyceraldehyde 3-phosphate dehydrogenase. Thus, in the light, electrons flow, causing NADPH to accumulate and ferredoxin to push electrons in the direction of these enzymes above, activating them and favoring the Calvin cycle. In the dark, the concentration of reduced NADPH, reduced ferredoxin, and reduced thioredoxin fall, resulting in loss of electrons by the Calvin cycle enzymes (oxidations that re-form disulfide bonds) and the Calvin cycle inactivates.

Other enzyme control mechanisms

Other means of controlling enzymes relate to access to substrate (substrate-level control) and control of enzyme synthesis. Hexokinase is an enzyme that is largely regulated by availability of its substrate, glucose. When glucose concentration is low, the product of the enzyme's catalysis, glucose-6-phosphate, inhibits the enzyme's function.

Regulation of enzymes by controlling their synthesis is covered later in the book in the discussion relating to control of gene expression.

This page titled 8.3: Control of Enzymatic Activity is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **4.2: Control of Enzymatic Activity** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **7.4: Control of Enzymatic Activity** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

8.4: Oxygen-Binding Proteins and Allosterism

5.03: Oxygen-Binding Proteins and Allosterism

This page titled 8.4: Oxygen-Binding Proteins and Allosterism is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

CHAPTER OVERVIEW

9: Carbohydrates

9.1: Carbohydrates overview
9.2: Monosaccharides
9.3: Disaccharides
9.4: Polysaccharides
9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls
9.6: The Sugar Code and Lectin Decoding
9.7: Chapter 7 Problems
9.8: Chapter 7 Problems - Answer Key

9: Carbohydrates is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.1: Carbohydrates overview

Carbohydrates

Carbohydrates are one of the four main classes of macromolecules that make up all cells and are an essential part of our diet; grains, fruits, and vegetables are all natural sources. While we may be most familiar with the role carbohydrates play in nutrition, they also have a variety of other essential functions in humans, animals, plants, and bacteria. In this section, we will discuss and review basic concepts of carbohydrate structure and nomenclature, and a variety of functions they play in cells.

Molecular structures

In their simplest form, **carbohydrates** can be represented by the stoichiometric formula $(CH_2O)_n$, where $_n$ is the number of carbons in the molecule. For simple carbohydrates, the ratio of carbon-to-hydrogen-to-oxygen in the molecule is 1:2:1. This formula also explains the origin of the term “carbohydrate”: the components are carbon (“carbo”) and the components of water (“hydrate”). Simple carbohydrates are classified into three subtypes: monosaccharides, disaccharides, and polysaccharides, which will be discussed below. While simple carbohydrates fall nicely into this 1:2:1 ratio, carbohydrates can also be structurally more complex. For example, many carbohydrates contain functional groups (remember them from our basic discussion about chemistry) besides the obvious hydroxyl. For example, carbohydrates can have phosphates or amino groups substituted at a variety of sites within the molecule. These functional groups can provide additional properties to the molecule and will alter its overall function. However, even with these types of substitutions, the basic overall structure of the carbohydrate is retained and easily identified.

Nomenclature

One issue with carbohydrate chemistry is the nomenclature. Here are a few quick and simple rules:

1. Simple carbohydrates, such as glucose, lactose, or dextrose, end with an "-ose."
2. Simple carbohydrates can be classified based on the number of carbon atoms in the molecule, as with triose (three carbons), pentose (five carbons), or hexose (six carbons).
3. Simple carbohydrates can be classified based on the functional group found in the molecule, i.e ketose (contains a ketone) or aldose (contains an aldehyde).
4. Polysaccharides are often organized by the number of sugar molecules in the chain, such as in a monosaccharide, disaccharide, or trisaccharide.

For a short video on carbohydrate classification, see the 10-minute Khan Academy video by clicking here.

Monosaccharides

Monosaccharides ("mono-" = one; "sacchar-" = sweet) are simple sugars; the most common is glucose. In monosaccharides, the number of carbons usually ranges from three to seven. If the sugar has an aldehyde group (the functional group with the structure R-CHO), it is known as an aldose; if it has a ketone group (the functional group with the structure RC(=O)R'), it is known as a ketose.

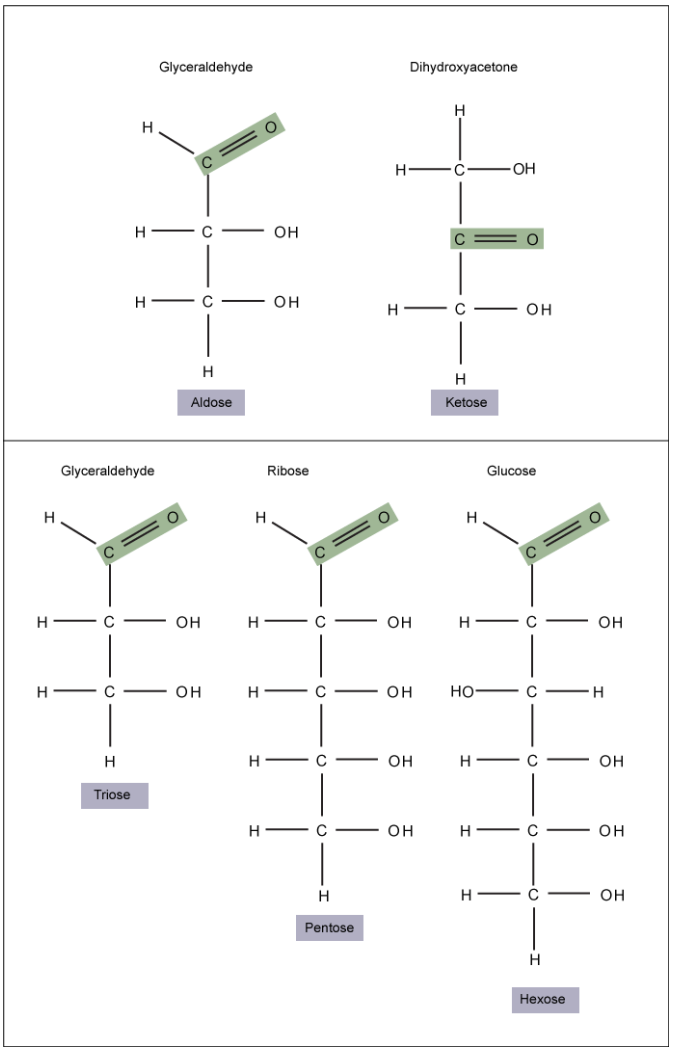


Figure 1. *Monosaccharides are classified based on the position of their carbonyl group and the number of carbons in the backbone. Aldoses have a carbonyl group (indicated in green) at the end of the carbon chain and ketoses have a carbonyl group in the middle of the carbon chain. Trioses, pentoses, and hexoses have three, five, and six carbons in their backbones, respectively. Attribution: Marc T. Facciotti (own work).*

Glucose versus galactose

Galactose (part of lactose, or milk sugar) and **glucose** (found in sucrose, glucose disaccharride) are other common monosaccharides. The chemical formula for glucose and galactose is $C_6H_{12}O_6$; both are hexoses, but the arrangements of the hydrogens and hydroxyl groups are different at position C_4. Because of this small difference, they differ structurally and chemically and are known as chemical isomers because of the different arrangement of functional groups around the asymmetric carbon; both of these monosaccharides have more than one asymmetric carbon (compare the structures in the figure below).

Fructose versus both glucose and galactose

A second comparison can be made when looking at glucose, galactose, and **fructose** (the second carbohydrate that with glucose makes up the disaccharide sucrose and is a common sugar found in fruit). All three are hexoses; however, there is a major structural difference between glucose and galactose versus fructose: the carbon that contains the **carbonyl** (C=O).

In glucose and galactose, the carbonyl group is on the C_1 carbon, forming an **aldehyde** group. In fructose, the carbonyl group is on the C_2 carbon, forming a **ketone** group. The former sugars are called **aldoses** based on the aldehyde group that is formed; the latter is designated as a **ketose** based on the ketone group. Again, this difference gives fructose different chemical and structural

properties from those of the aldoses, glucose, and galactose, even though fructose, glucose, and galactose all have the same chemical composition: $C_6H_{12}O_6$.

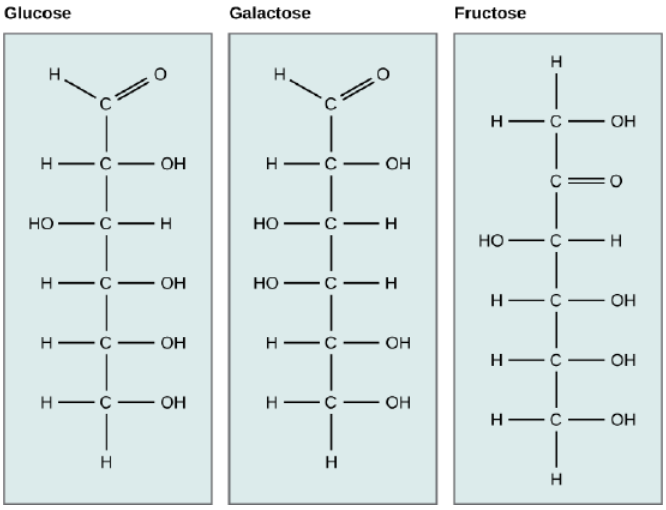


Figure 2. *Glucose, galactose, and fructose are all hexoses. They are structural isomers, meaning they have the same chemical formula ($C_6H_{12}O_6$) but a different arrangement of atoms.*

Linear versus ring form of the monosaccharides

Monosaccharides can exist as a linear chain or as ring-shaped molecules. In aqueous solutions, monosaccharides are usually found in ring form (Figure 3). Glucose in a ring form can have two different arrangements of the hydroxyl group (OH) around the anomeric carbon (C_1 that becomes asymmetric in the process of ring formation). If the hydroxyl group is below C_1 in the sugar, it is said to be in the alpha (α) position, and if it is above C_1 in the sugar, it is said to be in the beta (β) position.

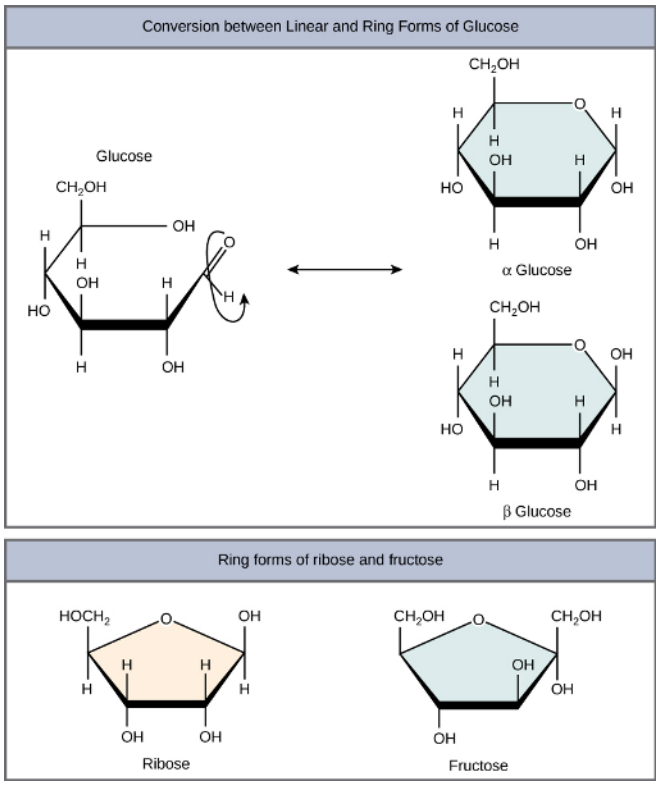


Figure 3. *Five- and six-carbon monosaccharides exist in equilibrium between linear and ring form. When the ring forms, the side chain it closes on is locked into an α or β position. Fructose and ribose also form rings, although they form five-membered rings as*

opposed to the six-membered ring of glucose.

Disaccharides

Disaccharides ("di-" = two) form when two monosaccharides undergo a dehydration reaction (also known as a condensation reaction or dehydration synthesis). During this process, the hydroxyl group of one monosaccharide combines with the hydrogen of another monosaccharide, releasing a molecule of water and forming a covalent bond. A covalent bond formed between a carbohydrate molecule and another molecule (in this case, between two monosaccharides) is known as a **glycosidic** bond. Glycosidic bonds (also called glycosidic linkages) can be of the alpha or the beta type.

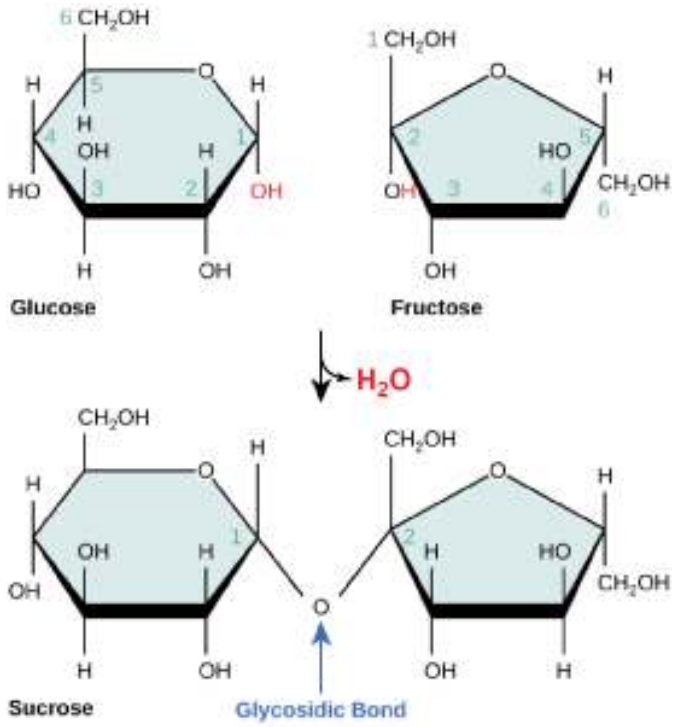
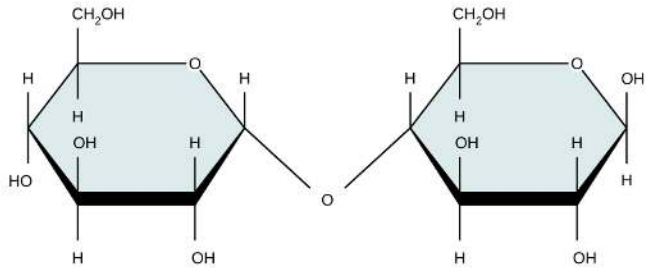
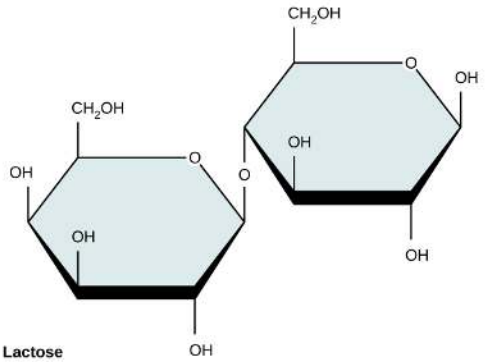


Figure 4. *Sucrose is formed when a monomer of glucose and a monomer of fructose are joined in a dehydration reaction to form a glycosidic bond. In the process, a water molecule is lost. By convention, the carbon atoms in a monosaccharide are numbered from the terminal carbon closest to the carbonyl group. In sucrose, a glycosidic linkage is formed between the C_1 carbon in glucose and the C_2 carbon in fructose.*

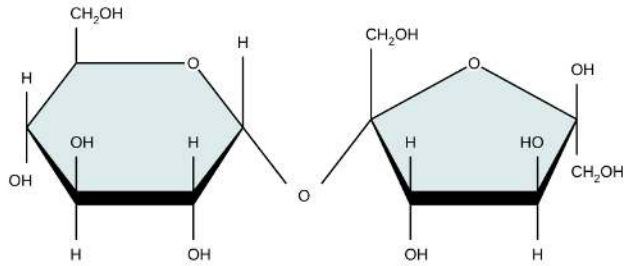
Common disaccharides include lactose, maltose, and sucrose (Figure 5). Lactose is a disaccharide consisting of the monomers glucose and galactose. It is found naturally in milk. Maltose, or malt/grain sugar, is a disaccharide formed by a dehydration reaction between two glucose molecules. The most common disaccharide is sucrose, or table sugar, which is composed of the monomers glucose and fructose.



Maltose

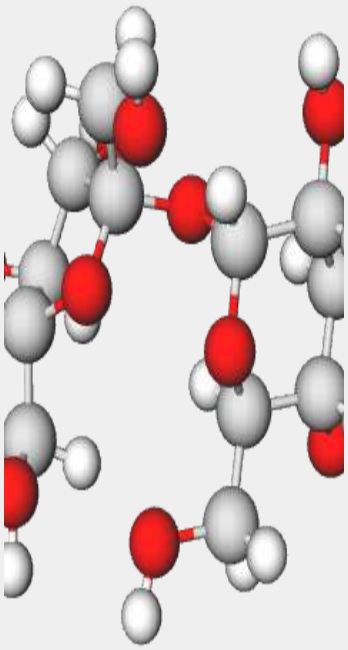
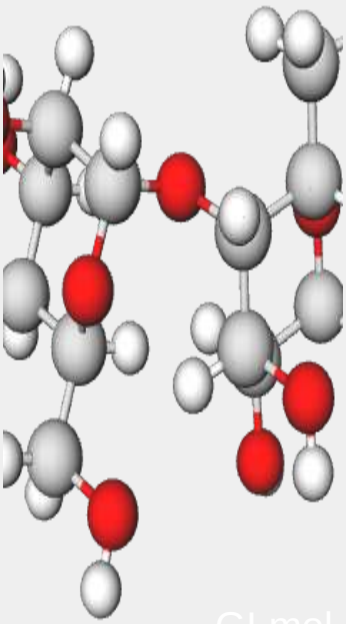
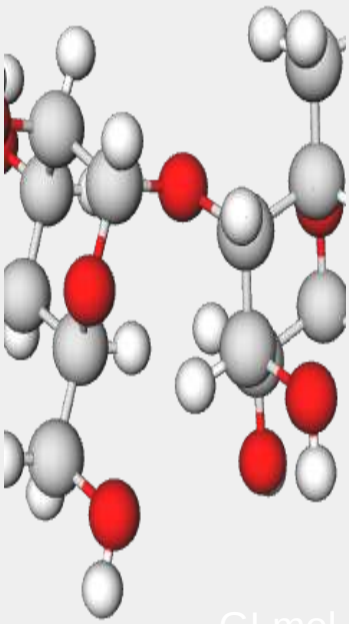


Lactose



Sucrose

Figure 5. *Common disaccharides include maltose (grain sugar), lactose (milk sugar), and sucrose (table sugar).*

Sucrose	Lactose	Maltose
		

Polysaccharides

A long chain of monosaccharides linked by glycosidic bonds is known as a **polysaccharide** ("poly-" = many). The chain may be branched or unbranched, and it may contain different types of monosaccharides. The molecular weight may be 100,000 Daltons or more, depending on the number of monomers joined. Starch, glycogen, cellulose, and chitin are primary examples of polysaccharides.

Starch is the stored form of sugars in plants and is made up of a mixture of amylose and amylopectin; both are polymers of glucose. Plants are able to synthesize glucose. Excess glucose, the amount synthesized that is beyond the plant's immediate energy needs, is stored as starch in different plant parts, including roots and seeds. The starch in the seeds provides food for the embryo as it germinates and can also act as a source of food for humans and animals who may eat the seed. Starch that is consumed by humans is broken down by enzymes, such as salivary amylases, into smaller molecules, such as maltose and glucose.

Starch is made up of glucose monomers that are joined by 1-4 or 1-6 glycosidic bonds; the numbers 1-4 and 1-6 refer to the carbon number of the two residues that have joined to form the bond. As illustrated in Figure 6, amylose is starch formed by unbranched chains of glucose monomers (only 1-4 linkages), whereas amylopectin is a branched polysaccharide (1-6 linkages at the branch points).

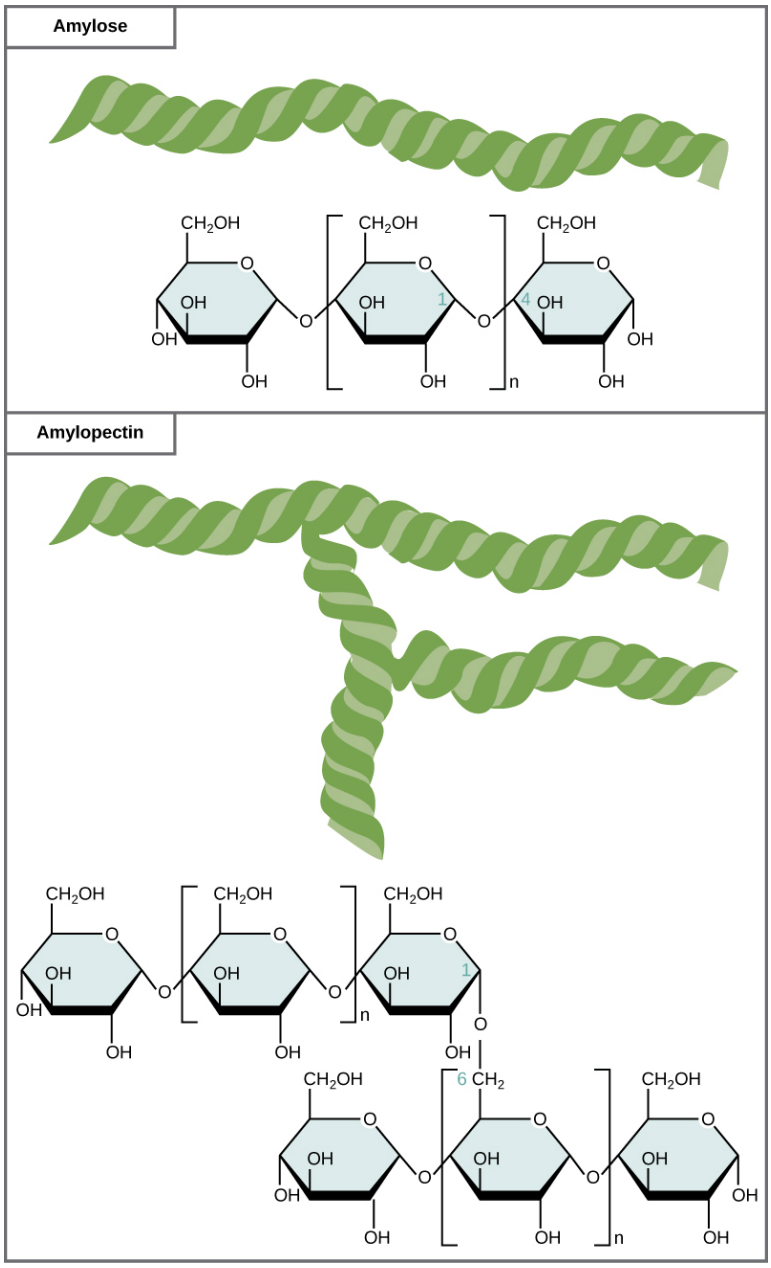
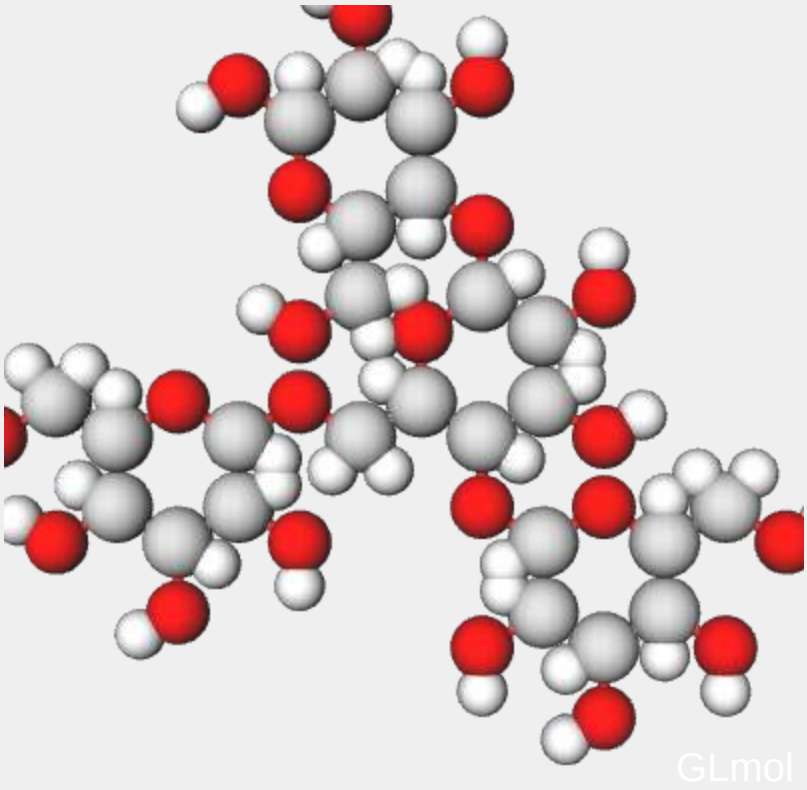


Figure 6. *Amylose and amylopectin are two different forms of starch. Amylose is composed of unbranched chains of glucose monomers connected by 1-4 glycosidic linkages. Amylopectin is composed of branched chains of glucose monomers connected by 1-4 and 1-6 glycosidic linkages. Because of the way the subunits are joined, the glucose chains have a helical structure. Glycogen (not shown) is similar in structure to amylopectin but more highly branched.*

Glycogen

Glycogen is a common stored form of glucose in humans and other vertebrates. Glycogen is the animal equivalent of starch and is a highly branched molecule usually stored in liver and muscle cells. Whenever blood glucose levels decrease, glycogen is broken down to release glucose in a process known as glycogenolysis.

Glycogen



Cellulose

Cellulose is the most abundant natural biopolymer. The cell wall of plants is mostly made of cellulose, which provides structural support to the cell. Wood and paper are mostly cellulosic in nature. Cellulose is made up of glucose monomers that are linked by β 1-4 glycosidic bonds.

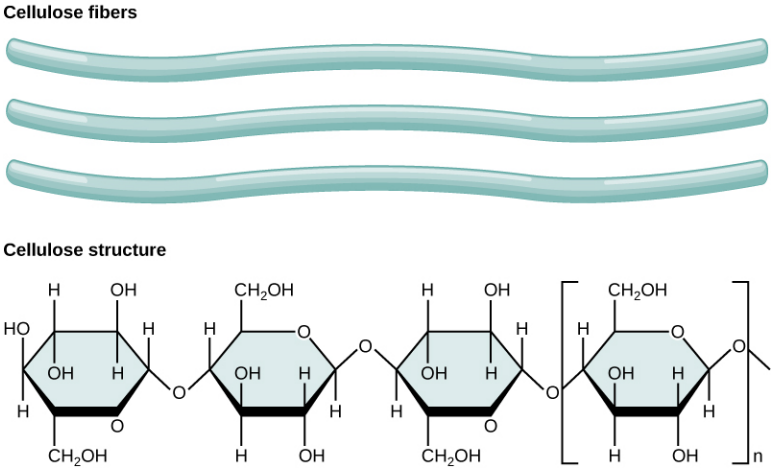


Figure 7. *In cellulose, glucose monomers are linked in unbranched chains by β 1-4 glycosidic linkages. Because of the way the glucose subunits are joined, every glucose monomer is flipped relative to the next one, resulting in a linear, fibrous structure.*

As shown in the figure above, every other glucose monomer in cellulose is flipped over, and the monomers are packed tightly as extended, long chains. This gives cellulose its rigidity and high tensile strength—which is so important to plant cells. While the β 1-4 linkage cannot be broken down by human digestive enzymes, herbivores such as cows, koalas, buffalos, and horses are able, with the help of the specialized flora in their stomach, to digest plant material that is rich in cellulose and use it as a food source. In these animals, certain species of bacteria and protists reside in the rumen (part of the digestive system of herbivores) and secrete the enzyme cellulase. The appendix of grazing animals also contains bacteria that digest cellulose, giving it an important role in the digestive systems of ruminants. Cellulases can break down cellulose into glucose monomers that can be used as an energy source by the animal. Termites are also able to break down cellulose because of the presence of other organisms in their bodies that secrete cellulases.

Interactions with carbohydrates

We have just discussed the various types and structures of carbohydrates found in biology. The next thing to address is how these compounds interact with other compounds. The answer to that is that it depends on the final structure of the carbohydrate. Because carbohydrates have many hydroxyl groups associated with the molecule, they are therefore excellent **H-bond donors** and **acceptors**. Monosaccharides can quickly and easily form H-bonds with water and are readily soluble. All of those H-bonds also make them quite "sticky". This is also true for many disaccharides and many short-chain polymers. Longer polymers may not be readily soluble.

Finally, the ability to form a variety of H-bonds allows polymers of carbohydrates or **polysaccharides** to form strong intramolecular and intermolocular bonds. In a polymer, because there are so many H-bonds, this can provide a lot of strength to the molecule or molecular complex, especially if the polymers interact. Just think of cellulose, a polymer of glucose, if you have any doubts.

Possible NB Discussion Point

Lipids and carbohydrates are not just classes of macromolecules that we discuss in BIS 2A but are also two of the essential macronutrients that we can obtain from eating various foods. Some popularized diet programs (e.g., Atkins, ketogenic) suggest limiting carbohydrates and/or fats. As you learn more about biomolecules and their roles in living systems, are you refining your perspective on foods and diets? What have you learned so far? Do you think there is anything missing in your understanding? Are you able to better understand and evaluate certain diets, such as the aforementioned?

This page titled 9.1: Carbohydrates overview is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Marc Facciotti.

- **8.1: Carbohydrates overview** by Marc Facciotti has no license indicated.

9.2: Monosaccharides

Learning Goals (ChatGPT 5, 11/11/25)

- **Core terminology:** Define sugar, carbohydrate, and glycan; distinguish monosaccharides from larger glycans.
- **Structures and representations:** Identify stereocenters and convert among Fischer, Haworth, chair, and wedge/dash for D-glucose, D-ribose, and D-fructose.
- **Cyclization and anomers:** Explain hemiacetal ring formation to furanose or pyranose, locate the anomeric carbon, distinguish α and β, and describe mutarotation and ring populations.
- **Reducing behavior and derivatives:** Explain reducing behavior via ring opening to an aldehyde or α-hydroxyketone, and recognize common derivatives (oxidation, phosphorylation, amination, acetylation, lactonization) and their roles.

These goals prepare students to analyze and predict properties of monosaccharides.

9.2.1: Introduction

Carbohydrate or glycan biochemistry is very complex and challenging owing to the stereochemical complexity of simple sugars, the large number of positions on the sugars used to form linkages between other sugars to create polymers, the large number of chemical modifications to base sugars, and the lack of a genetic template to instruct glycan polymer formation. No wonder our understanding of complex glycans has developed after that of the chemically simpler polymers like nucleic acids and proteins.

In addition, the terminology used to describe them varies as well. We use these general descriptions of them:

Sugar: usually refers to low molecular weight carbohydrates like glucose, lactose, and sucrose, but it can also refer broadly to any carbohydrate.

Carbohydrate: a general term that applies to simple sugars to complex sugar polymers like glycogen, starch, and cellulose. The name derives from the formula for simple sugars like glucose ($C_6H_{12}O_6$), which can be written as $C_6(H_2O)_6$ - a carbo (C) - hydrate (H_2O).

Glycan: a general term for molecules containing simple sugars and sugar derivatives linked in a polymer, either standalone molecules or attached to other molecules like proteins.

9.2.2: Classification of Monosaccharides

Monosaccharides are classified along a few orthogonal axes. Use these labels together when naming and comparing structures.

- **Carbonyl type:** *aldose* (aldehyde at C1) or *ketose* (ketone). For ketoses, specify the carbonyl position when not at C2 (for example, *3-ketose* or the stem *hex-3-ulose*).
- **Carbon count:** triose, tetrose, pentose, hexose, heptose, etc. Combine with carbonyl type (for example, *aldohexose*, *ketopentose*).
- **D/L configuration:** assigned by the stereocenter farthest from the carbonyl, by reference to D-glyceraldehyde. D/L does not predict optical rotation.
- **Ring size in solution:** furanose (5-member) or pyranose (6-member). Report the common ring if relevant (for example, β-D-ribofuranose vs β-D-glucopyranose).
- **Common epimer pairs:** D-glucose/D-mannose (C2 epimers), D-glucose/D-galactose (C4 epimers); D-ribose/D-arabinose/D-xylose/D-lyxose vary at C2–C4.
- **Frequent derivatives (name stems):** deoxy- (for example, 2-deoxyribose), *-onic* acids (C1 oxidation), *-uronic* acids (C6 oxidation), amino sugars (GlcN, GalN), N-acetylated forms (GlcNAc, GalNAc), phosphates (for example, Glc-6-P).

Naming tip: A precise name can stack these labels, for example “β-D-arabino-hex-3-ulose” or “α-D-glucopyranose-6-phosphate.”

9.2.2.1: Carbonyl type

Aldose has an aldehyde in the open chain at C1. *Ketose* has a ketone internal to the chain. When the ketone is not at C2, state its position (for example, 3-ketose; systematic stem “hex-3-ulose”). In a Fischer projection, an aldose shows the carbonyl at the top carbon; a ketose shows a C=O lower in the chain. In a Haworth projection, find the carbon bonded to two oxygens (ring O and the anomeric substituent); C1 implies an aldose, C2 a 2-ketose, C3 a 3-ketose.

9.2.2.2: Carbon count

Triose, tetrose, pentose, hexose, heptose, and so on. Combine with carbonyl type (for example, aldohexose, ketopentose). In a Fischer projection, number from the carbonyl carbon and count the backbone. In a Haworth projection, count ring atoms and include the exocyclic CH_2OH to reach the open-chain carbon total.

9.2.2.3: D/L configuration

Assign by the stereocenter farthest from the carbonyl in the open chain. If the OH at that center points right in the Fischer projection it is D, left is L. D/L does not predict optical rotation. In the standard Haworth orientation used here, D-pyranoses show the terminal CH_2OH drawn up, L-pyranoses down.

9.2.2.4: Ring size in solution

Five-member rings are furanoses and six-member rings are pyranoses. For common D-aldoses, intramolecular attack of O5 on C1 gives a pyranose; attack of O4 on C1 gives a furanose. For 2-ketoses, attack of O5 on C2 gives a furanose, while attack of O6 on C2 gives a pyranose. In Haworth, count the ring atoms to name furanose or pyranose and report the prevalent form when relevant.

9.2.2.5: Anomeric carbon and α/β

The anomeric carbon is the former carbonyl after ring closure (C1 in aldoses, C2 in 2-ketoses, C3 in 3-ketoses). Assign α or β only at anomeric carbons. In Haworth, compare the anomeric OH to the CH_2OH-bearing carbon: for the D series, α is trans to CH_2OH and β is cis. In pyranoses drawn in the standard orientation, this appears as α down and β up.

9.2.3: Monosaccharides Structures

The above definition of sugar needs some further nuance. From a chemical perspective, sugars can be defined as polyhydroxy-aldehydes or ketones. The simplest sugars contain at least three carbon atoms, and the most common are the aldo- and keto-trioses, tetroses, pentoses, and hexoses. The 3C sugars are glyceraldehyde and dihydroxyacetone, as shown in Figure **9.2.1**.

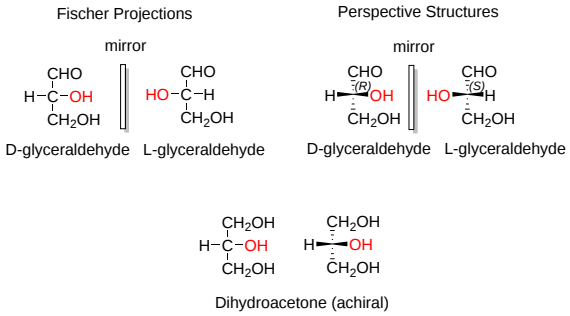


Figure **9.2.1**: Three-carbon sugars

Glucose, an aldohexose, is a central sugar in metabolism. It and other 5C and 6C sugars can cyclize through intramolecular nucleophilic attack of one of the free hydroxyl groups on the carbonyl carbon of the aldehyde or ketone. Such intramolecular reactions occur if stable 5- or 6-member rings can form. The resulting rings are labeled furanose (5-member) or pyranose (6-member) based on their similarity to furan and pyran. On nucleophilic attack to form the ring, the carbonyl O becomes an OH that points either below (α anomer) or above (β anomer) the ring.

Figure **9.2.2** shows different representations of the linear and cyclic forms of the sugars D-glucose, D-ribose, and D-fructose

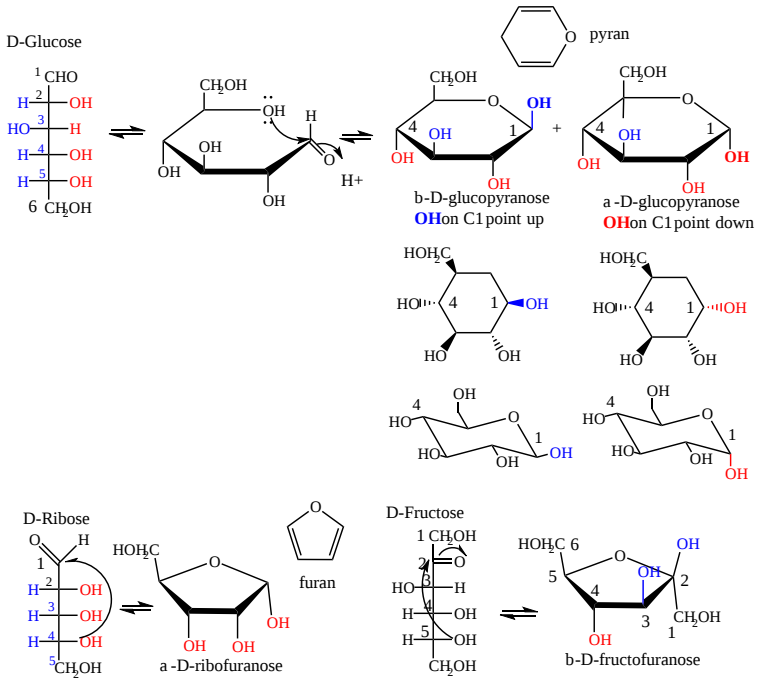


Figure 9.2.2: Linear and cyclic forms of D-glucose, D-ribose, and D-fructose

Monosaccharides in solution exist as equilibrium mixtures of the straight and cyclic forms. In solution, glucose (Glc) is mainly in the pyranose form, fructose is 67% pyranose and 33% furanose, and ribose is 75% furanose and 25% pyranose. However, in polysaccharides, Glc is exclusively pyranose, and fructose and ribose are furanoses.

Sugars can be drawn in the straight chain form as **Fischer projections** or perspective structural formulas.

In the Fischer projection, the vertical bonds point down into the plane of the paper. That's easy to visualize for 3C sugars, but more complicated for larger ones. For those, draw a wedge and dash line drawing of the molecule. When determining the orientation of the OHs on each C, orient the wedge and dash drawing in your mind so that the C atoms adjacent to the one of interest are pointing down. Sighting towards the carbonyl C, if the OH is pointing to the right in the Fischer project, it should be pointing to the right in the wedge and dash drawing, as shown below for D-threose and D-glucose. Figure 9.2.3 shows how to convert Fischer projections to wedge dash representations.

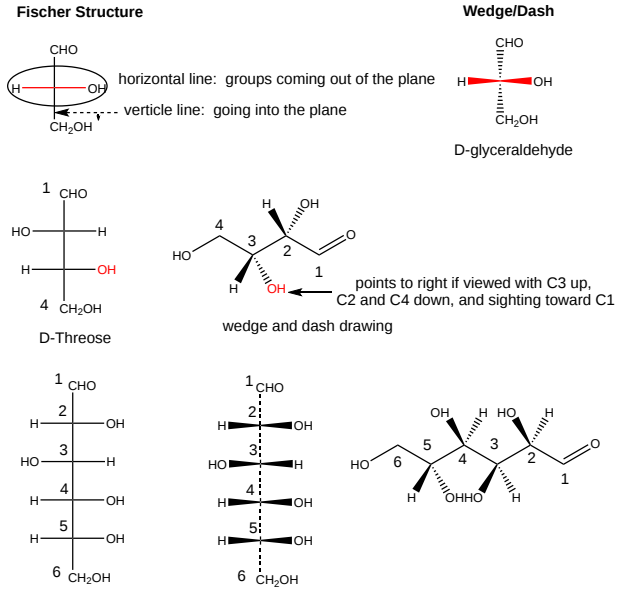


Figure 9.2.3: Converting Fischer projections to wedge dash representations.

Figure 9.2.4 shows an interactive iCn3D model of D-glucose in a linear form.

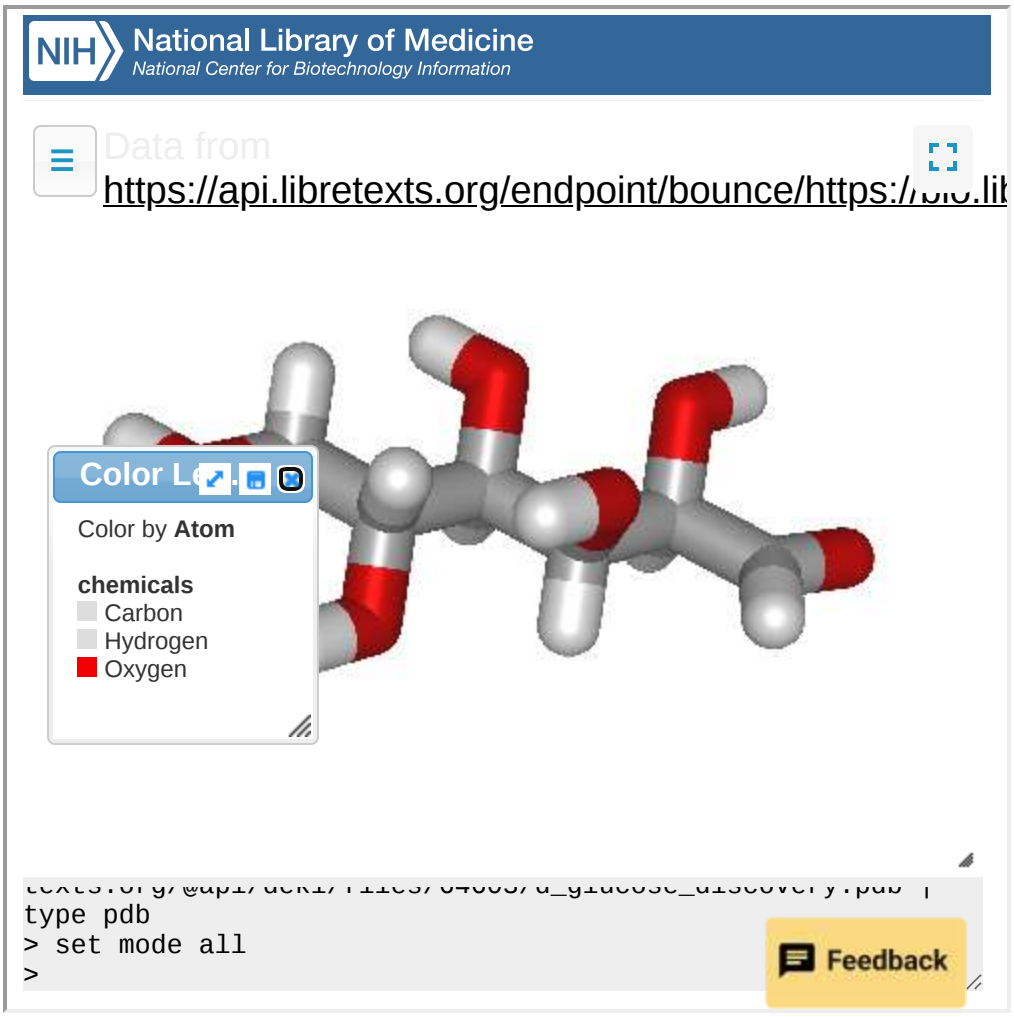


Figure 9.2.4: D-glucose (Copyright; author via source).

Orient the molecule as shown in Figure 9.2.5 below, with the carbonyl oxygen pointed to the far right, and compare it to the orientation shown in Figure 9.2.5 to reinforce your understanding of Fischer and wedge/dash projections.

Figure 9.2.5:

Cyclic forms can be drawn either as the **Haworth projections**, which show the molecule as cyclic and planar with substituents above or below the ring, or the more plausible bent forms (showing glucose in the chair or boat conformations). β-D-glucopyranose is the only aldohexose that can be drawn with all its bulky substituents (OH and CH_2OH) in equatorial positions, which probably accounts for its widespread prevalence in nature. Figure 9.2.6 shows four different representations of glucose.

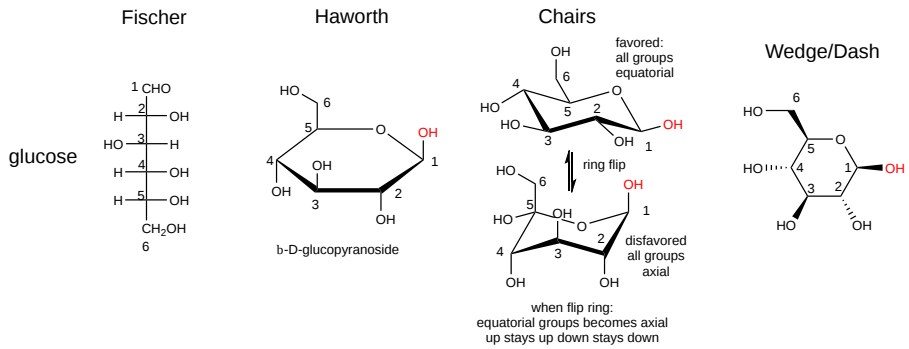


Figure 9.2.6: Fischer and cyclic Haworth, chair and wedge/dash representations of glucose

Haworth projections are more realistic than the Fischer projections, but you should be able to draw both structures. Generally, if a substituent points to the right in the Fischer projection, it points down in the Haworth projection. If it points left, it points up. Generally, the OH on the **α**-anomer points down (**α**nts down) while on the **β**-anomer, it points up (**β**utterflies up) as illustrated in Figure 9.2.7

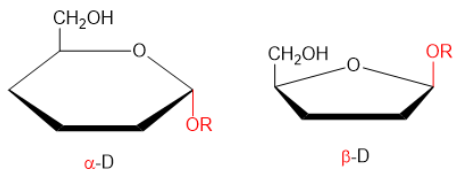


Figure 9.2.7: Alpha and beta Haworth representations of sugars

In the Haworth projections, the bulky R group of the next carbon after the carbon whose OH group was the nucleophile for ring formation is pointed up if the OH engaged in the attack was on the right-hand side in the straight chain Fischer diagram (as in α-D-glucopyranose above, when the CH_2OH group is up). Conversely, it is pointed down if the OH engaged in the attack was on the left-hand side in the straight chain Fischer diagram (as in α-D-galactofuranose above when the $(CHOH)CH_2OH$ group is down). The rest of the OH groups still follow the simple rule that if they point to the right in the Fischer projection, they point down in the Haworth projection.

The Fischer structures of the most common monosaccharides (other than glyceraldehyde and dihydroxyacetone), which you will encounter most frequently, are shown in Figure 9.2.8.

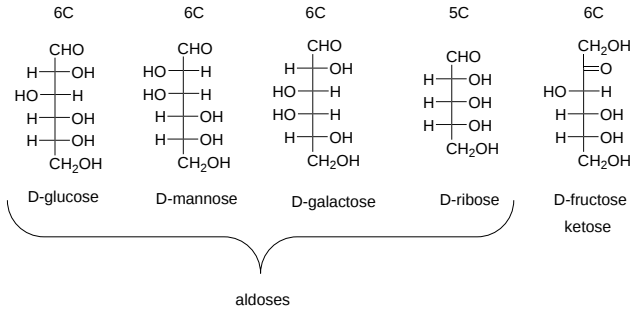


Figure 9.2.8: Most common monosaccharides discussed in this book

The mirror image of D-Glc is L-Glc. The D- and L- designations refer to the center of asymmetry most remote from the aldehyde or ketone. By convention, all chiral centers are related to D-glyceraldehyde, so sugar isomers related to D-glyceraldehyde at their last asymmetric center are D sugars.

Figure 9.2.9 shows multiple renderings of common hexoses.

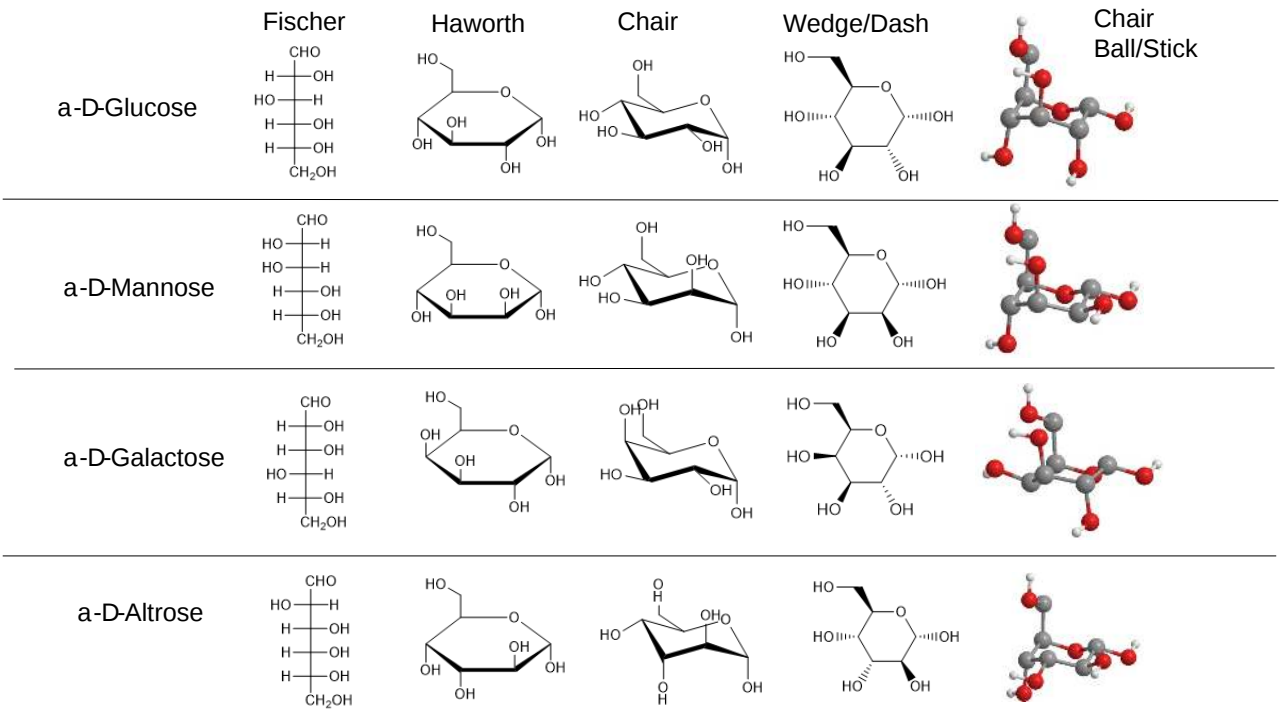


Figure 9.2.9: Multiple renderings of common hexoses

9.2.4: Isomers

Sugars can be configurational (interconverted only by breaking covalent bonds) or conformational isomers. Figure 9.2.10 reviews different configurational isomers.

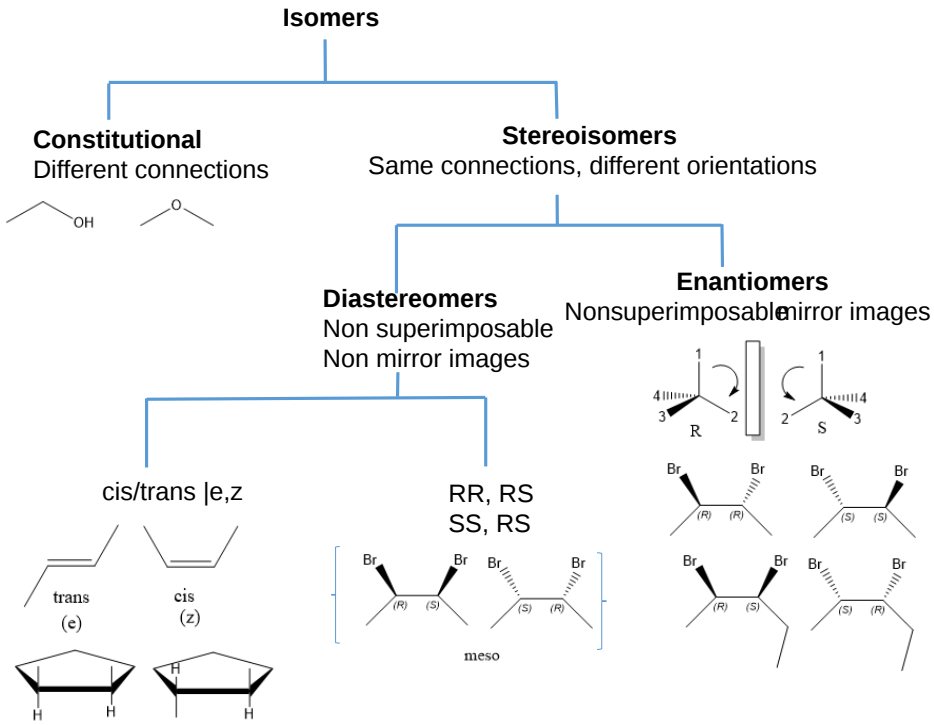


Figure 9.2.10: Types of isomers

The configurational isomers include enantiomers (stereoisomers that are mirror images of each other), diastereomers (stereoisomers that are not mirror images), **epimers** (diastereomers that differ at one stereocenter), and **anomers** (a particular form of stereoisomer, diastereomer, and epimer). Figure 9.2.11 below shows enantiomers, diastereomers, epimers, and anomers of 6-carbon sugars.

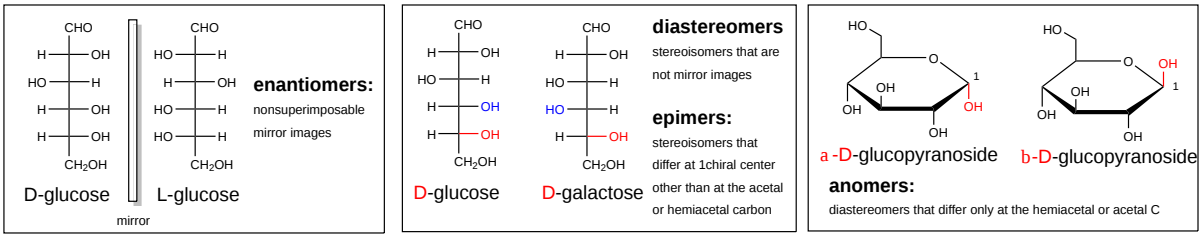


Figure 9.2.11: Enantiomers, diastereomers, epimers, and anomers of 6-carbon sugars.

Sugars can also exist as conformational isomers, which interchange without breaking covalent bonds. These include chair and boat conformations of the cyclic sugars as shown in Figure 9.2.12.

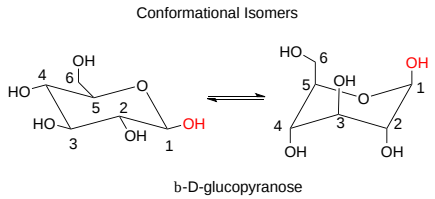


Figure 9.2.12: Conformational isomers of b-D-glucopyranose

9.2.5: Monosaccharide Derivatives

Many derivatives of monosaccharides are found in nature. These include

- oxidized forms in which the aldehyde and/or alcohol functional groups are oxidized to carboxylic acids
- phosphorylated forms in which phosphates are transferred from ATP to form phosphoester derivatives
- amine derivatives such as glucosamine or galactosamine
- acetylated amine derivatives such as N-Acetyl-GlcNAc (GlcNAc) or GalNAc
- lactone forms (intramolecular esters) in which an OH group attacks a carbonyl C that was previously oxidized to a carboxylic acid
- condensation products of sugar derivatives with lactate ($CH_3CHOHCO_2^-$) and pyruvate ($CH_3COCO_2^-$), both from the glycolytic pathway, to form muramic acid and neuraminic acids (also called sialic acids), respectively.

Figure 9.2.13 some simple monosaccharide derivatives.

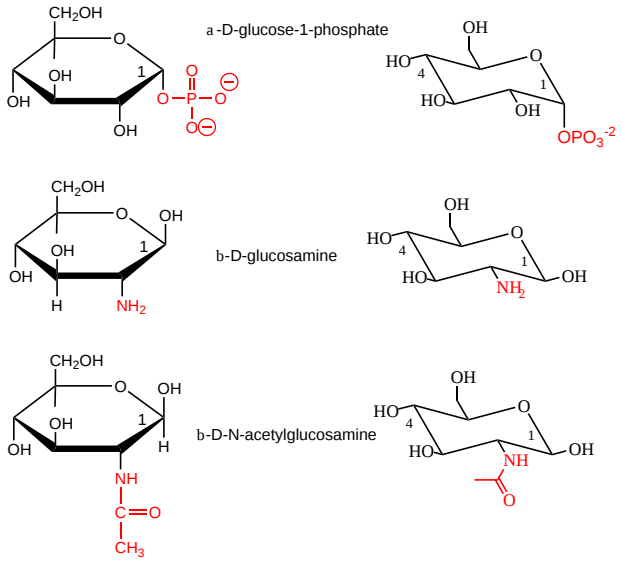


Figure 9.2.13: Monosaccharide derivatives

Figure 9.2.14 shows some additional oxidative derivatives of glucose shown in Fischer projections.

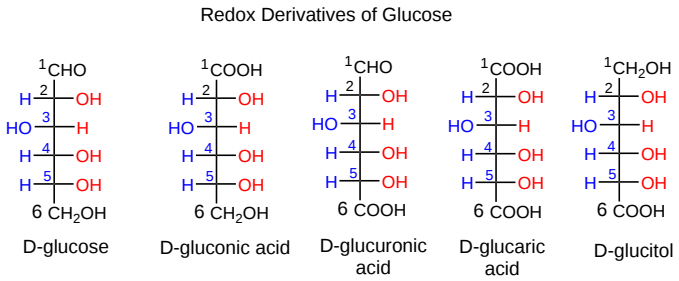


Figure 9.2.14: Redox derivatives of glucose

Other important derivatives of monosaccharides are **sialic acids**. **N-acetylmuramic acid**, found in bacterial cell walls, consists of GlcNAc in an ether link at C3 with lactate, while **N-acetylneuraminic acid** results from an intramolecular cyclization of a condensation product of ManNAc and pyruvate. These sialic acids are shown in Figure 9.2.15.

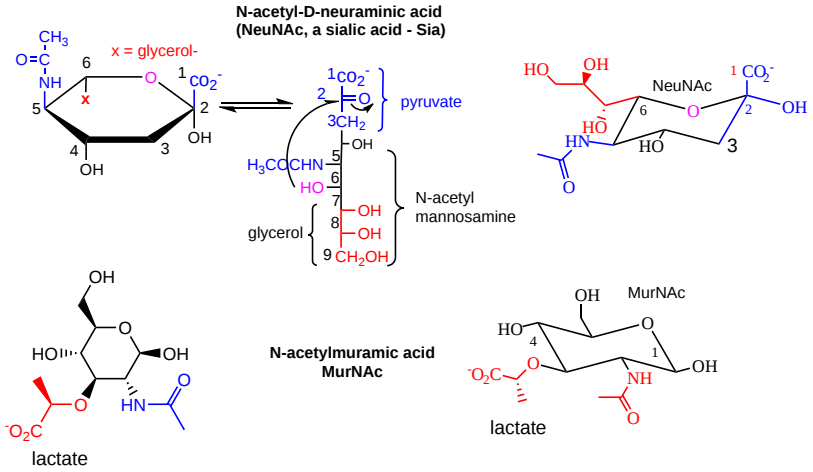


Figure 9.2.15: Sialic Acids

Sugars are very complicated, as the linkages and substituents are so diverse. Figure 9.2.16 shows differences in sialic acids between humans and chimps.

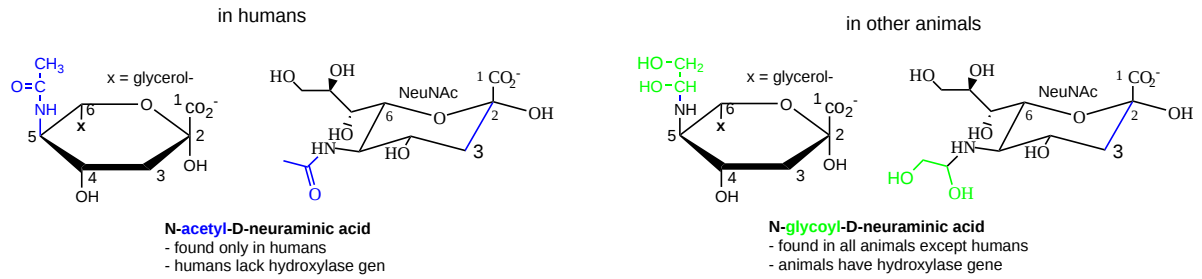


Figure 9.2.16: Sialic acids between humans and chimps.

What happens when non-vegan humans eat animal products (meat, milk) with N-glycoyl neuraminic acids (Neu5Gc)? Some get incorporated into human membrane glycans. Sialic acids on surface proteins can serve as "receptors" that allow the binding of self-cells as well as foreign cells or proteins that have evolved to bind them. A toxin, SubAB, secreted by E. Coli 0157, can bind Neu5Gc. Hence, eating meat products can make us more susceptible to bacteria that recognize Neu5Gc.

9.2.6: Formation of Hemiacetals and Acetals

Monosaccharides that contain aldehydes can cyclize through an intramolecular nucleophilic attack of an OH at the carbonyl carbon in an addition reaction to form a **hemiacetal.** In the past, the group was called a **hemiketal** if the attack was on a ketone, but now they are also called hemiacetals. On the addition of acid (which protonates the anomeric OH, forming water as a potential leaving group), another alcohol can add, forming an **acetal** with water leaving. These reactions are shown in Figure 9.2.17.

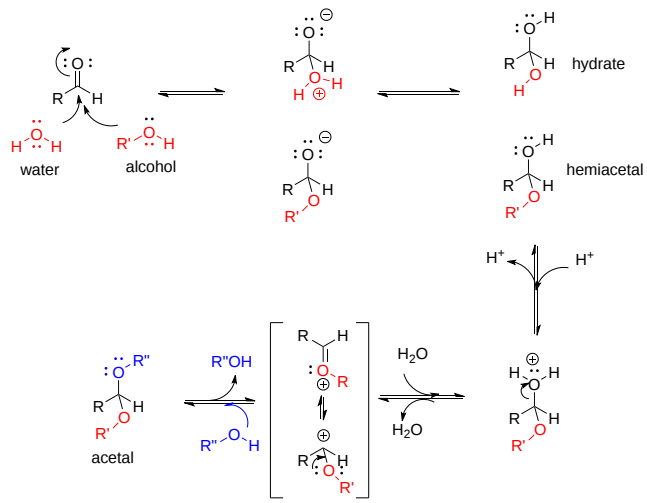


Figure 9.2.17: Hemiacetal and acetal formation

9.2.7: Summary

This chapter provides an in-depth exploration of the complexity of carbohydrate chemistry, emphasizing the structural, stereochemical, and functional diversity of sugars and glycans. It highlights why glycan biochemistry is particularly challenging due to the vast array of possible structures that arise from simple monosaccharide building blocks.

Key topics include:

1. **Terminology and Definitions:**
   The chapter begins by clarifying the terms "sugar," "carbohydrate," and "glycan." While "sugar" typically refers to small, low molecular weight carbohydrates like glucose, the term "carbohydrate" encompasses everything from simple sugars to complex polymers such as glycogen and cellulose. "Glycan" is used more generally to describe any polymer of sugars, whether free or attached to proteins and lipids.

2. **Monosaccharide Structures and Representations:**
   An overview of monosaccharide structures is provided, including the chemical basis of sugars as polyhydroxy-aldehydes or ketones. The chapter explains how simple sugars (e.g., trioses, tetroses, pentoses, hexoses) can cyclize to form stable 5-

membered (furanose) or 6-membered (pyranose) rings. Various methods of structural representation are compared, such as Fischer projections, Haworth projections, and chair conformations, highlighting the stereochemical intricacies and the significance of α- and β-anomers.

3. **Isomerism and Structural Diversity:**
   The text reviews different types of isomers found in sugars, including configurational isomers (enantiomers, diastereomers, epimers, and anomers) and conformational isomers (chair and boat forms). This section underscores how subtle differences in stereochemistry can have profound effects on glycan function and recognition.

4. **Monosaccharide Derivatives and Modifications:**
   Beyond the basic sugar units, the chapter discusses various chemical modifications of monosaccharides, such as oxidation, phosphorylation, amination, acetylation, lactone formation, and complex condensation reactions that lead to the formation of important derivatives like sialic acids. These modifications are critical in determining the biological roles and properties of glycans.

Overall, this chapter lays a foundation for understanding the intricate world of carbohydrate biochemistry by examining the molecular structures, stereochemical variations, and modifications that contribute to the functional complexity of glycans in biological systems.

This page titled 9.2: Monosaccharides is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **7.1: Monosaccharides and Disaccharides** by Henry Jakubowski and Patricia Flatt has no license indicated.

9.3: Disaccharides

Learning Goals (ChatGPT 5, 11/11/25)

- **Read linkage notation:** Interpret α or β at the donor anomeric carbon and positions in 1→4, 1→6, and 1→2 linkages.
- **Predict reducing behavior:** Determine whether a disaccharide is reducing or nonreducing and explain Benedict and Fehling outcomes.
- **Understand glycosidic bond chemistry:** Describe formation of acetals from anomeric hemiacetals, acid sensitivity, and enzymatic hydrolysis by specific glycosidases; assumes prior knowledge of ring-closing hemiacetals from Monosaccharides.
- **Recognize common examples:** Assign compositions and linkages for maltose, lactose, cellobiose, sucrose, and trehalose.

These goals support structural analysis, naming, and property prediction for disaccharides.

9.3.1: Introduction

Scope: This page builds on the Monosaccharides page and focuses on how two monosaccharides join to form disaccharides. The emphasis is on linkage notation, anomeric configuration, and reducing versus nonreducing behavior.

What you should already know: Numbering of carbons in common monosaccharides, identification of anomeric carbons, and conversion among Fischer, Haworth, and chair forms.

Carbohydrate or glycan biochemistry is very complex and challenging owing to the stereochemical complexity of simple sugars, the large number of positions on the sugars used to form linkages between other sugars to create polymers, the large number of chemical modifications to base sugars, and the lack of a genetic template to instruct glycan polymer formation. No wonder our understanding of complex glycans has developed after that of the chemically simpler polymers like nucleic acids and proteins.

In addition, the terminology used to describe them varies as well. We use these general descriptions of them:

Sugar: usually refers to low molecular weight carbohydrates like glucose, lactose, and sucrose, but it can also refer broadly to any carbohydrate.

Carbohydrate: a general term that applies to simple sugars to complex sugar polymers like glycogen, starch, and cellulose. The name derives from the formula for simple sugars like glucose ($C_6H_{12}O_6$), which can be written as $C_6(H_2O)_6$ - a carbo (C) - hydrate (H_2O).

Glycan: a general term for molecules containing simple sugars and sugar derivatives linked in a polymer, either standalone molecules or attached to other molecules like proteins.

9.3.2: Formation of Disaccharides

This section focuses on acetal formation that creates glycosidic bonds between two sugars. Ring-closing hemiacetal formation in single sugars is treated on the Monosaccharides page.

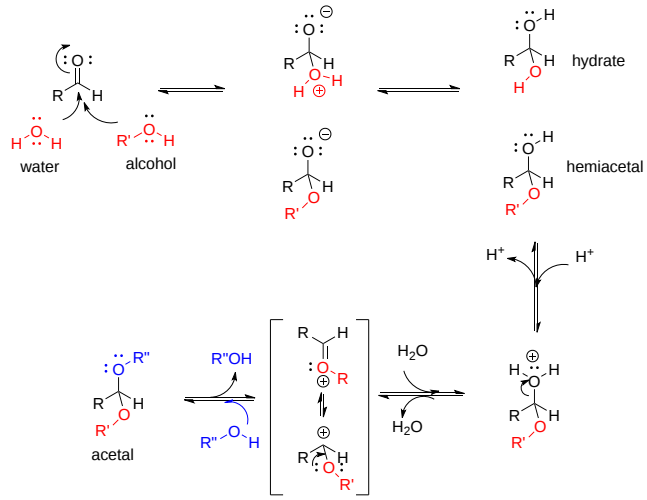


Figure 9.3.17: Hemiacetal and acetal formation

Assign α or β only to a residue whose *anomeric carbon* participates in the glycosidic bond. If one anomeric carbon participates, give a single α or β for that residue. If both anomeric carbons participate, give α/β for both residues.

If the other alcohol is a second monosaccharide, a disaccharide results. The acetal link bonding the two monosaccharides is called a **glycosidic bond**. Because any hydroxyl can serve as the acceptor, linkages are diverse, including 1→2, 1→3, 1→4, and 1→6. Here are examples of disaccharides:

- maltose: Glc(α1→4)Glc, a disaccharide hydrolysis product of glycogen or starch
- cellobiose: Glc(β1→4)Glc, a disaccharide hydrolysis product of cellulose
- lactose: Gal(β1→4)Glc, also known as milk sugar
- sucrose: Glc(α1→2β)Fru. Because fructose participates through its anomeric carbon, sucrose does not equilibrate to an open-chain keto form and is nonreducing.

Figure 9.3.18 illustrates the differences between lactose and sucrose. Note that the β-D-fructofuranose ring is flipped (left to right as in turning one of your hands over) compared to Figure 9.3.16.

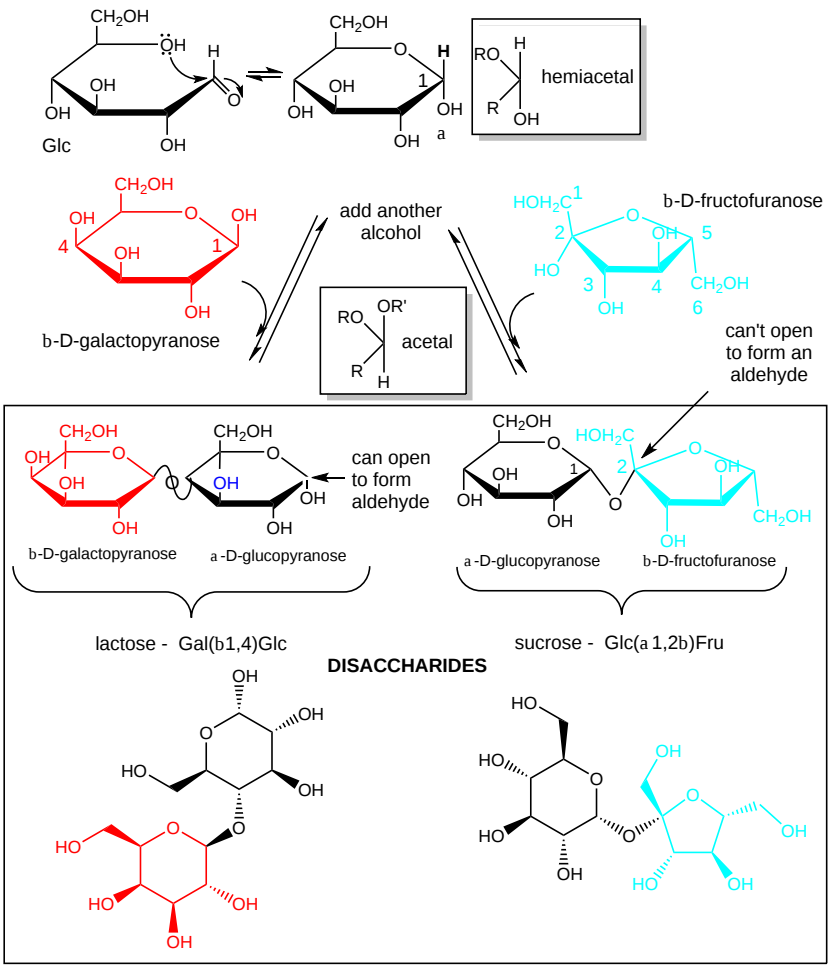


Figure 9.3.18: Structures of lactose and sucrose

Acetal links between sugars in glycans can be hydrolyzed by water under acid catalysis, and by specific glycosidases.

The disaccharides above that retain a free anomeric carbon (for example maltose and lactose) are **reducing sugars**. In lactose, galactose is attached to glucose through the OH on C4, so the anomeric carbon of glucose (C1) can revert to the open-chain aldehyde. Benedict and Fehling reagents oxidize that aldehyde, reducing Cu^{2+} to Cu_2O. Sugars that possess a potentially open-chain aldehyde or an *α-hydroxy ketone* that tautomerizes to an aldehyde under basic conditions (such as fructose) give positive tests.

Any mono-, di-, or polysaccharide with at least one hemiacetal remains reducing. If all anomeric carbons are locked as acetals, as in sucrose where both anomeric carbons form the 1→2 link, the sugar is nonreducing.

Alpha-gal syndrome

Alpha-gal syndrome (AGS) arises after tick exposure to the disaccharide galactose-α-1,3-galactose (alpha-gal), which elicits IgE antibodies. Subsequent exposures, including ingestion of red meat or cow's milk, can trigger allergic reactions up to anaphylaxis.

The structure of Gal(α1,3)Gal is shown in Figure 9.3.19.

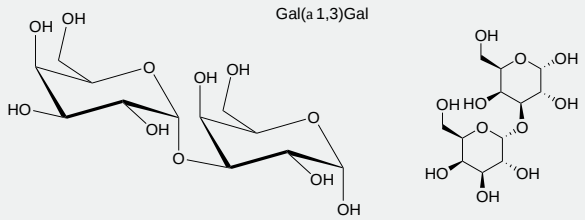


Figure 9.3.19: The disaccharide Gal(α1,3)Gal

Figure 9.3.20 shows an interactive iCn3D model of α-D-galactosyl-(1,3)-α-D-galactose (PubChem 9840966).

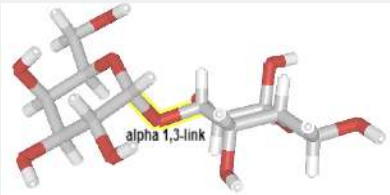
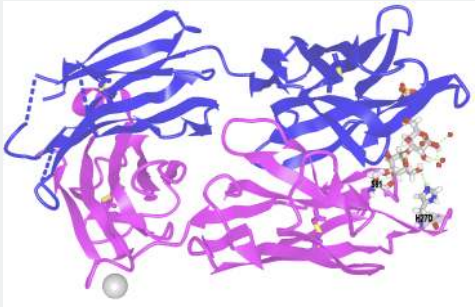


Figure 9.3.21 shows an interactive iCn3D model of the human anti-alpha-galactosyl antibody complex (7UEN). The Fab light chain is magenta and the heavy chain is blue. Side chains H27 and S91 contact the disaccharide.



9.3.3: Summary

Disaccharides form when the anomeric carbon of one sugar reacts to give an acetal (the glycosidic bond) with a hydroxyl on the same or a second sugar. Configuration at the anomeric carbon (α or β) and the carbons joined (for example 1→4, 1→2, 1→6) define structure and properties. If at least one anomeric carbon remains a hemiacetal, the disaccharide is reducing and can mutarotate. If both anomeric carbons are locked as acetals, the disaccharide is nonreducing.

Key topics:

1. **Glycosidic bond chemistry:** Formation of acetals from anomeric hemiacetals, acid-catalyzed hydrolysis, and specificity of glycosidases.
2. **Nomenclature and notation:** Reading α or β at the donor anomeric carbon and mapping positions in 1→4, 1→6, and 1→2 linkages.
3. **Reducing vs nonreducing disaccharides:** Structural criteria and outcomes with Benedict or Fehling reagents.
4. **Representative structures:** Maltose [Glc(α1→4)Glc], lactose [Gal(β1→4)Glc], cellobiose [Glc(β1→4)Glc], sucrose [Glc(α1→2β)Fru], and trehalose [Glc(α1→1α)Glc].
5. **Biological relevance:** Dietary sources, digestion and intolerance, and immune recognition exemplified by the α-Gal epitope.

Overall, this chapter lays a foundation for understanding the intricate world of carbohydrate biochemistry by examining the molecular structures, stereochemical variations, and modifications that contribute to the functional complexity of glycans in biological systems.

This page titled 9.3: Disaccharides is shared under a not declared license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **7.1: Monosaccharides and Disaccharides** by Henry Jakubowski and Patricia Flatt has no license indicated.

9.4: Polysaccharides

Learning Goals (ChatGPT o1, 1/30/25)

- **Define and Differentiate Carbohydrates and Glycans:**
  - Distinguish among the terms “sugar,” “carbohydrate,” and “glycan,” and explain how these concepts encompass everything from simple monosaccharides to highly complex glycan polymers.
- **Understand Monosaccharide Structure and Representation:**
  - Describe the chemical structure of monosaccharides, including key features such as stereochemistry, the formation of cyclic (pyranose and furanose) structures, and the distinction between α- and β-anomers.
  - Convert between different structural representations (Fischer projections, Haworth projections, chair and wedge/dash forms) and understand how these representations inform our perception of 3D carbohydrate conformation.
- **Glycosidic Linkages and Polymer Structure:**
  - Explain the mechanism of glycosidic bond formation (hemiacetal and acetal formation) and describe how different linkage types (e.g., α 1,4, α 1,6, β 1,4) give rise to the structural diversity of polysaccharides.
  - Compare and contrast homopolymers such as starch, glycogen, dextran, cellulose, and chitin in terms of their linkage types, branching patterns, and overall 3D conformation.
- **Structure–Function Relationships in Storage vs. Structural Polysaccharides:**
  - Discuss how the type of glycosidic linkage (e.g., α in starch/glycogen versus β in cellulose/chitin) influences the polymer’s physical properties, biological roles (energy storage vs. structural support), and digestibility.
  - Analyze the role of branching (e.g., α 1,6 linkages in glycogen and amylopectin) in enhancing solubility and rapid mobilization of stored glucose.
- **Utilizing Symbolic Nomenclature for Glycans (SNFG):**
  - Learn and apply the SNFG system to accurately represent glycan structures using specific colors and shapes, facilitating clear communication of complex carbohydrate architectures.
- **3D Conformational Analysis and Modeling:**
  - Interpret interactive 3D models and computer-generated representations (e.g., iCn3D models) of polysaccharide fragments to understand their tertiary and quaternary arrangements, such as the helical structures of V-amylose or the fiber assembly in cellulose.
- **Exploration of Glycosaminoglycans (GAGs):**
  - Identify the structure and function of key GAGs (e.g., hyaluronic acid, keratan sulfate, chondroitin sulfate, dermatan sulfate, and heparin) and describe how variations in disaccharide repeat units and sulfation patterns contribute to their biological roles, such as in extracellular matrix composition and cell signaling.
- **Comparative Structural Analysis in Biological Contexts:**
  - Contrast the molecular structures of polysaccharides from different biological sources (e.g., cellulose in plants versus chitin in invertebrates) and discuss the impact of subtle structural differences (e.g., the presence of N-acetyl groups) on function.
- **Practical and Clinical Implications:**
  - Understand the physiological significance of polysaccharide structure in energy storage (glycogen and starch) and structural support (cellulose, chitin, and GAGs) and discuss how these structures influence their biochemical properties, such as solubility and reactivity.
  - Explore real-world applications, such as the use of iodine staining in starch detection, the role of glycogenin in glycogen synthesis, and the impact of dietary glycans on human health (e.g., Neu5Gc incorporation).

Achieving these goals will equip students with the fundamental and advanced concepts needed to analyze and interpret the structural and functional diversity of carbohydrates and glycans in biological systems.

Polysaccharides contain many monosaccharides in glycosidic links and may have many branches. They serve as either structural components or energy storage molecules. Polysaccharides consisting of single monosaccharides are homopolymers. Starch, glycogen, dextran, cellulose, and chitin are the most common. We'll discuss based on whether the acetal link is alpha or beta.

9.4.1: α 1,4 main chain links

Starch and **Glycogen**: These polysaccharides are polymers of glucose linked in α 1,4 links with α 1,6 branches. **Starch**, found in plants, is subdivided into **amylose**, which has no branches, and **amylopectin,** which does. Starch granules consist of about 20% amylose and 80% amylopectin. **Glycogen,** the main CHO storage in animals, is found in muscle and liver and consists of glucose residues in α 1,4 links with lots of α 1,6 branches (many more branches than in starch).

Here are various ways to render in 2D the chemical structure of a branched glycogen and starch fragment, as shown in Figure 9.4.1.

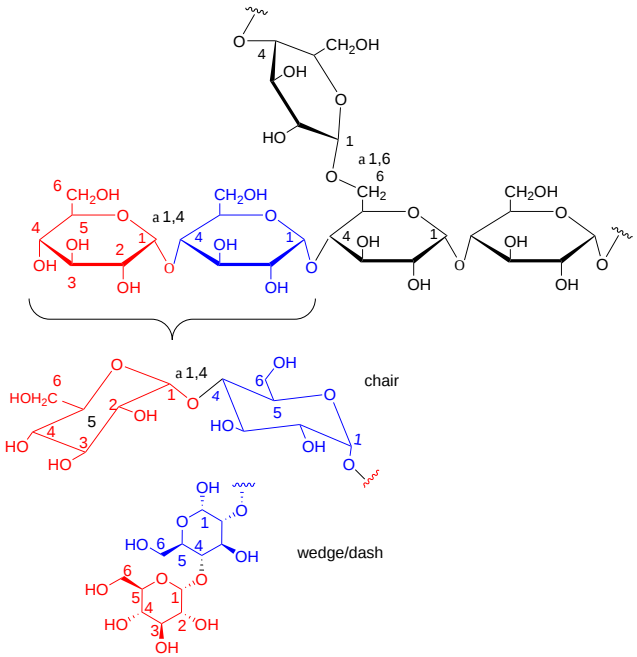


Figure 9.4.1: Structure of a branched glycogen and starch fragment

The top part of the figure shows the Haworth structure. The bottom part shows two glucose units in red and blue in the more structurally clear chair and wedge/dash representations.

Figure 9.4.2 shows an interactive iCn3D model of 10 glucose monosaccharides in an α-(1,4) linkage with five glucose units with α-(1,4) linkages attached to the main chain through an α-(1,6) branch at glucose 6 of the main chain. The type of substructure would be found in starch (amylopectin) and glycogen.

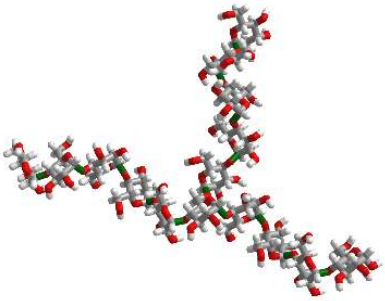


Figure 9.4.2: α-(1,4) linked glucose with an α-(1,6) branch (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TXXy589xPNgmG9

Figure 9.4.3 shows the structure in the iCn3D model in a diagrammatic fashion in which glucose is represented as a blue circle with the acetal/glycosidic/glucosidic linkages between the monosaccharides written between the circles. The 14A label shows that the acetal linkage is an α-(1,4) link with a single α-(1,6) branch.

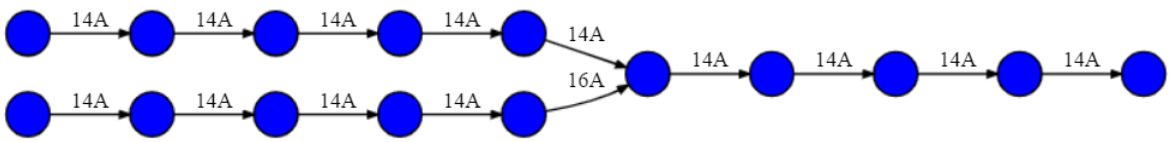


Figure 9.4.3: Symbolic Nomenclature For Glycans (SNFG) for an α-(1,4) linked glucose polymer with an α-(1,6) branch.

The linkages are written in a variety of conventions. These include 14A, 14α, 4A and 4α. The link between many sugars is often a 1,x link where x is 2,3, 4, 5, or 6. In those cases, the number 1 can be omitted. The program-generated images in this text use numbers and A or B.

What makes carbohydrates so complex is their 3D structures. Like proteins and nucleic acids, they can adopt a myriad of conformations. As the monomeric units are so homogeneous, especially in homopolymers, it can be difficult to obtain crystal structures for them, so computer models are often used.

Studies have shown that the simple starch fraction amylose, an α 1,4 polymer of glucose, often envisioned as a straight chain, can adopt three main conformations. They are double-helical A- (found chiefly in cereals), double helical B-(found primarily on tubers) amyloses, and single-helical V-amylose (or simply A, B, and V structures). The A and B do **NOT** represent alpha or beta in this classification system. The A and B forms consist of double helices aligned in a parallel fashion with about six glucose residues per turn. The helices appear to be left or right-handed, and this ambiguity might arise from a lack of crystal structures.

In contrast, a well-defined structure of the V helix is known. It folds into a left-handed helix with six glucose residues per turn and a pitch of about 8Å. Unlike alpha helices of proteins and the double-stranded helix of DNA, the center of the helix is **NOT** packed tightly and can accommodate small molecules. One is iodine (triiodide, I_3^-), which exhibits a dark blue color when bound in amylose in starch. This is the basis for starch indicators that you may have used in titration reactions in biology and chemistry courses.

Alpha helices might self-associate during folding in proteins to form a 4-helix bundle. Likewise, the helices in V-amylose can associate into bundles. Figure 9.4.4 shows an interactive iCn3D model of the actual structure of a V-amylose, cycloamylose 26 (1C58). It consists of a linear cycloamylose strand of 26 glucose monomers, which has collapsed to form secondary structure with 6-residue helices packed together into a tertiary structure of 4-helix bundle. The blue sphere "cartoon" color coding of each glucose residue corresponds to the blue circles in the diagrammatic representation above.

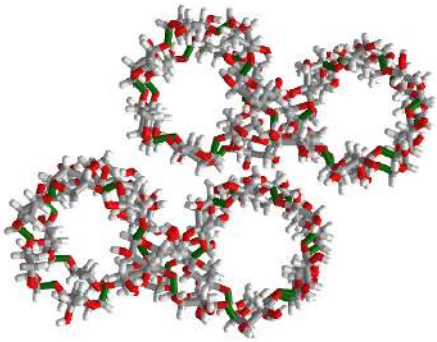


Figure 9.4.4: V-amylose, cycloamylose 26 (1C58). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...6tghM5HMTypm88

Rotate the model to explore it. Trace the chains by following the blue sphere symbolic representation of glucose as you trace the main chain. Rotate it to view down the helix axes to see the four holes that can each accommodate an I_3^-. In the menu button (☰), choose Style, Chemicals, Sphere to see a spacefilling model that shows the holes within each helix. Remember, NO unoccupied holes exist within either a protein alpha helix or a double-stranded DNA molecule.

The well-known macrocyclic compound cyclodextrins (for example, α-cyclodextrin) are structures equivalent to one turn of V-amylose. The V-amylose helix is partially stabilized by hydrogen bonds from donors and acceptors within the helix from the OH3 on the ith glucose and the OH2 on the $i+1^{th}$ glucose, as well as from the OH6 on the ith glucose and the OH2 on the $i+6^{th}$ glucose.

In vivo, glycogen is synthesized by attaching glucose monomers to a core protein called glycogenin. Figure 9.4.5 shows a model of a glycogen particle with glycogenin at its core.

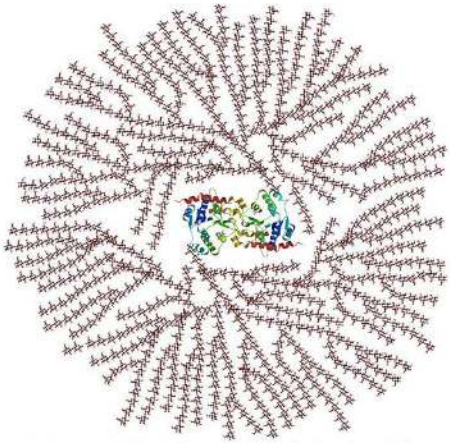


Figure 9.4.5: Glycogen particle with glycogenin at its core

The dimeric protein glycogenin is an enzyme that autoglucosylates itself stepwise. The first glucose is added at Tyr 195. At some point, the active site must get buried, and the protein can no longer add more monomers.

Storing glucose residues as either glycogen or starch, one large molecule, makes chemical sense. A review of colligative properties would inform you that if glucose were stored as the monosaccharide, a great osmotic pressure difference would be found between the outside and inside of the cell. Glycogen, with its many branches, is a single molecule. When glucose is needed, it is cleaved one residue at a time from all the branches (at the nonreducing ends) of glycogen, producing a large amount of free glucose quickly.

Phi/Psi angles can also be used in Ramachandran plots to show the conformations around the acetal link for the starch/glycogen main chain in a fashion comparable to that for proteins (around the alpha carbon). The phi torsion angle describes rotation around the C1-O bond of the acetal link, and the psi angle describes rotation around the O-C4 bond of the same acetal link, with the glucopyranose ring considered as a rigid rotator (just as the six atoms in the planar peptide bond unit). The most extended form of a glucose polymer occurs when the glycosidic link is β 1,4 (as in cellulose), which forms linear chains. This would be analogous to the more extended parallel beta strand (phi/psi angles of -119^0, -113^0) and antiparallel beta strands (phi/psi angles of -139^0, $+135^0$) of proteins. The α 1,4-linked main chain of glycogen and starch causes the chain to turn and form a large helix. Iodine (or I_3^-) can

fit into the helix, which turns a solution/suspension of starch blue, which turns starch purple. The less extended structure is analogous to the less extended protein alpha helix, which has phi/psi angles of -57^0,-47^0.

Figure **9.4.6** shows phi/psi angles for acetal/glycosidic linkage in maltose, a disaccharide of glucose, as shown below.

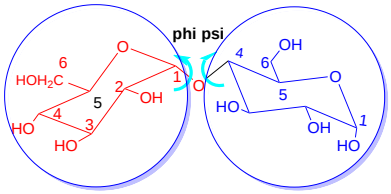


Figure **9.4.6**: Phi/psi angles for acetal/glycosidic linkage in the glucoside disaccharide maltose

9.4.2: α 1,6 main chain links

Dextran is a branched polymer of glucose in α 1,6 links with α 1,2, α 1,3, or α 1,4 linked side chains. This polymer is used in some chromatography resins. Figure **9.4.7** shows chair structures (A) and wedge/dash structures (B) for dextran, showing the main chain α 1,6 link with one α 1,3 branch.

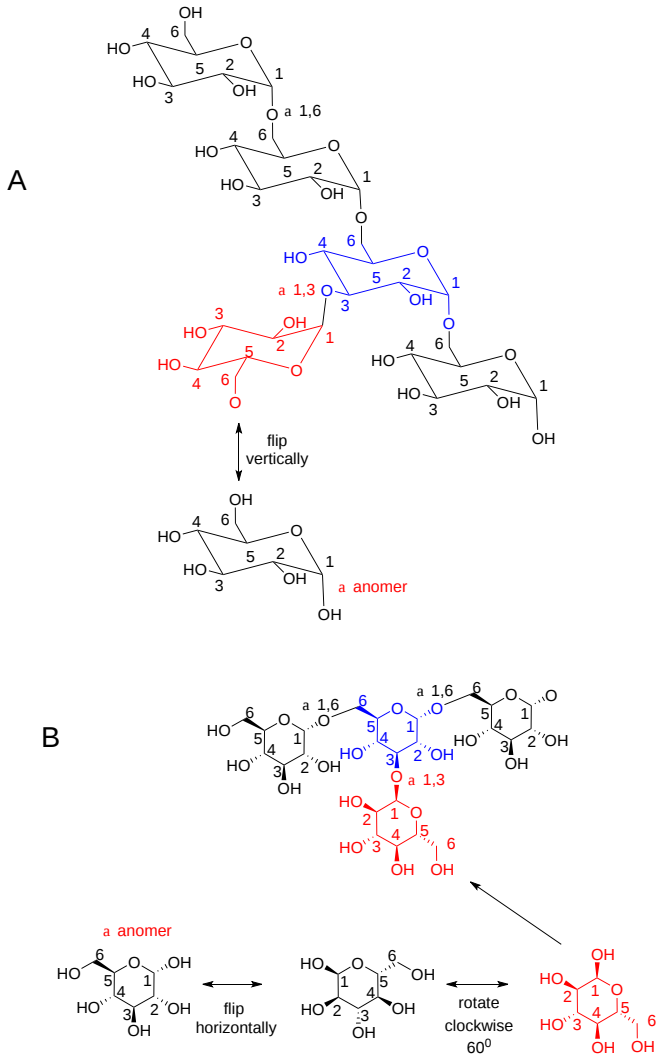


Figure **9.4.7**: Chair structures (A) and wedge/dash structures (B) for dextran showing the main chain α 1,6 link with one α 1,3 branch

Depending on its molecular weight, it is soluble in water (forming viscous solutions) and organic solvents. It is also used as a food thickener and stabilizer. It is synthesized by lactic acid-forming bacteria using sucrose as an energy source. Most uses are

commercial.

9.4.3: β 1,4 links

Cellulose, a structural homopolymer of glucose in plants, has β 1,4 main chain links without branching. Multiple chains are held together by intra- and inter-chain H-bonds. It is the most abundant biological molecule in nature. Various renderings of 4 glucose residues in cellulose are shown in Figure **9.4.8**. Haworth structures are not shown. Instead, more chemically informative chairs and wedge/dash structures are used. It's important to see the structures displayed in many ways since different representations of carbohydrate structures can be found in different sources.

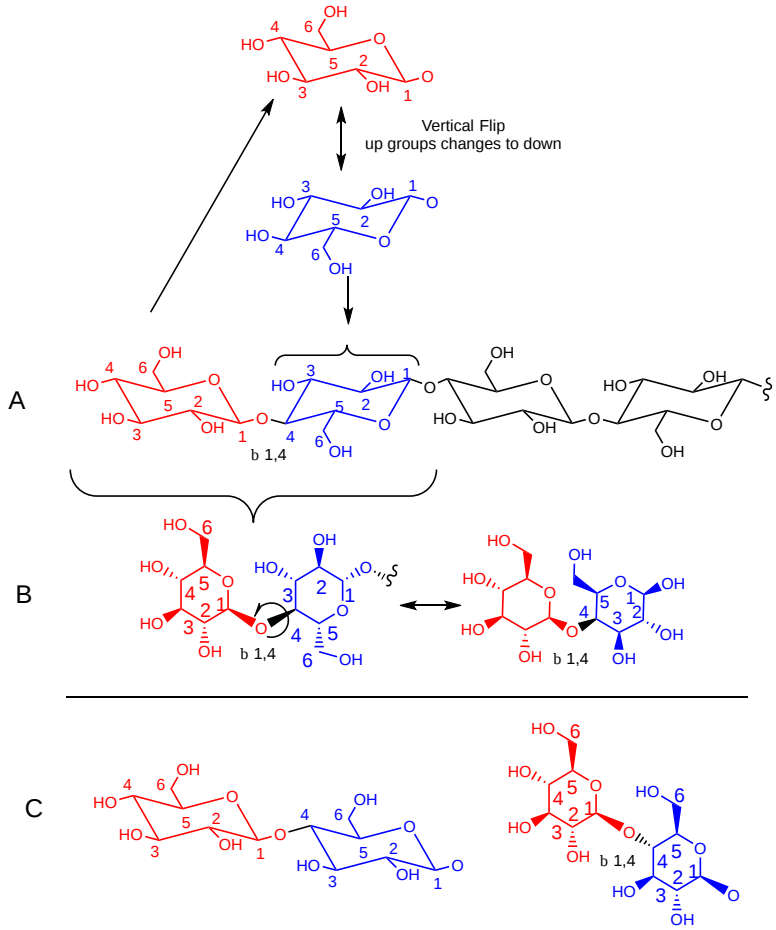


Figure **9.4.8**: Rendering of 4 glucose residues in cellulose

In A, the most common chair representation, the 2nd and 4th residues from the right-hand end are flipped versions of residues 1 and 3. Residues 1 and 2 are colored red and blue for clarity. This unit is repeated to generate the full chain. The top part of A shows a simplified version of the flip of the red ring to produce the blue ring to help you see that they are indeed identical structures.

The same structure as in A is shown in the left part of B in wedge/dash from (looking down on the ring). The right-hand side of B shows a variant of B's left-hand side generated by a simple 180^0 rotation around the bond indicated on the left of B.

The simple repeat is shown in C without the chain flips in A and B. The acetal/glycosidic/glucosidic bond seems to be shown in a straight line in the chair structures (a bit confusing and structurally deceptive), but more clearly in the adjacent wedge/dash structure.

All structures are correct, but the one shown in A is most often used.

One long chain can interact with other chains in a structure stabilized by intrachain and interchain hydrogen bonds. Different sources display different hydrogen bonds. Some common ones are shown below. These chains align in parallel and twist to form larger cellulose fibers. Figure **9.4.9** shows an interactive iCn3D model of cellulose chains.

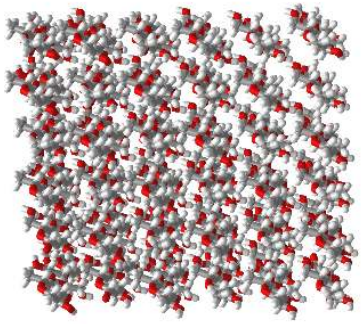


Figure 9.4.9: Cellulose chains (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TUHKki2VN4jeTA

In addition, hydrophobic interactions occur between adjacent planes of cellulose strands. How can that be, considering the strongly polar nature of glucose and a single cellulose strand? Figure 9.4.9*B*: below, with two representations of cellulose, shows how axial hydrogens project vertically from the plane of the glucose monomer rings to form a weak but biologically significant hydrophobic surface.

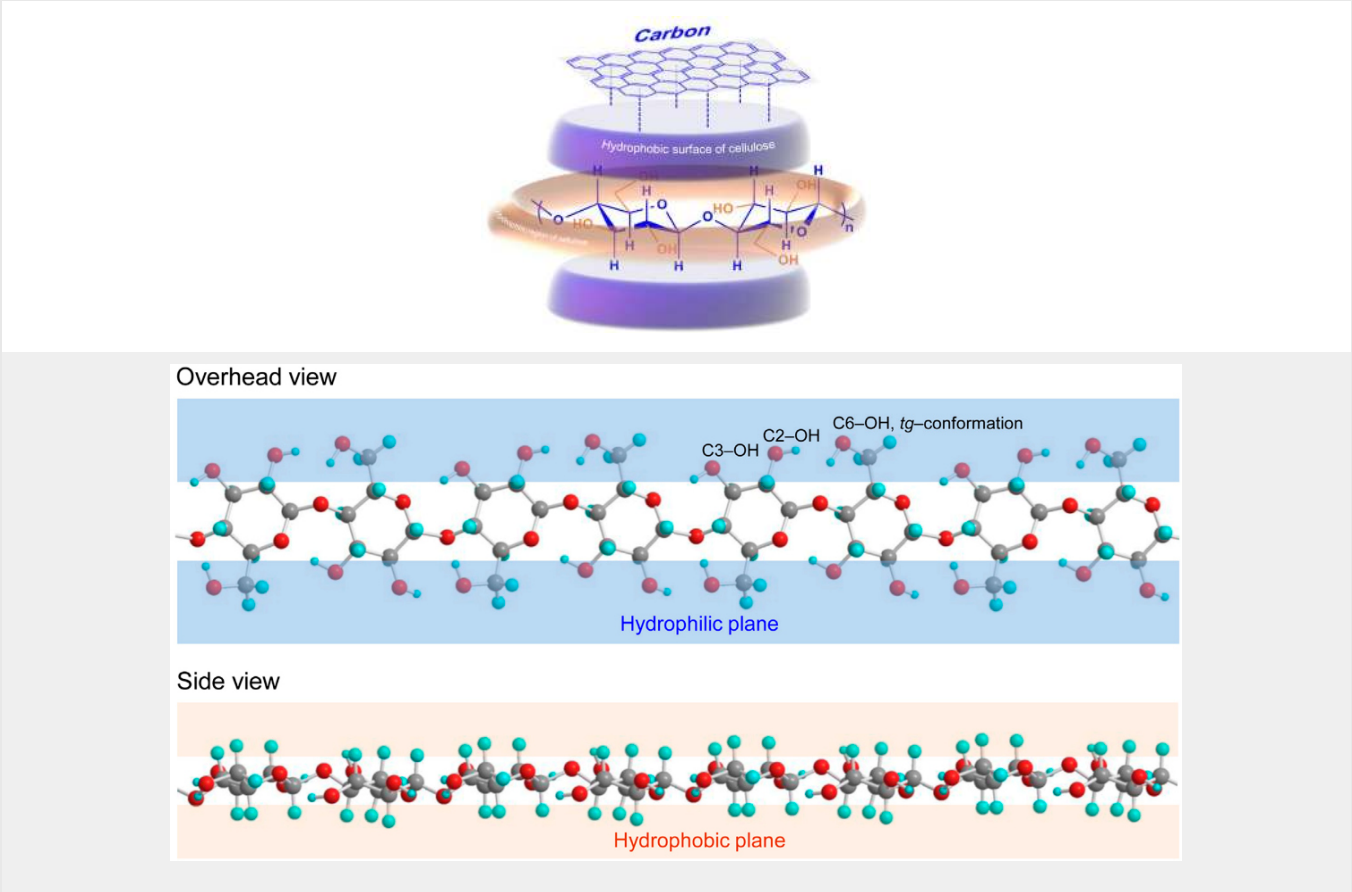


Figure 9.4.9: Two representations showing the nonpolar axial surface of cellulose strands that contribute to hydrophobic interactions stabilizing the cellulose assemblies. The top panel shows how cellulose could interact with a nonpolar planar substance, such as a layer of graphite. Axial H atoms are shown in green/cyan in the bottom panel.

Top Panel: Yang G, Luo X and Shuai L (2021) Bioinspired Cellulase-Mimetic Solid Acid Catalysts for Cellulose Hydrolysis. *Front. Bioeng. Biotechnol.* 9:770027. doi: 10.3389/fbioe.2021.770027. **Creative Commons Attribution License (CC BY).**

Bottom Panel: Uusi-Tarkka, E.-K.; Skrifvars, M.; Haapala, A. Fabricating Sustainable All-Cellulose Composites. *Appl. Sci.* **2021**, *11*, 10069. https://doi.org/10.3390/app112110069. Creative Commons Attribution (CC BY) license

(https://creativecommons.org/licenses/by/
4.0/).

Chitin

The glycan is the major component in the exoskeletons of anthropoids and mollusks. It is a β 1,4-linked polymer of N-acetylglucosamine (GlcNAc). Compare this with cellulose, which is a β 1,4-linked polymer of glucose. What a difference an N-acetyl substituent makes!

The basic chemical structure of chitin is shown in chair form in Figure **9.4.10** along with the symbolic nomenclature for glycans (SNFG).

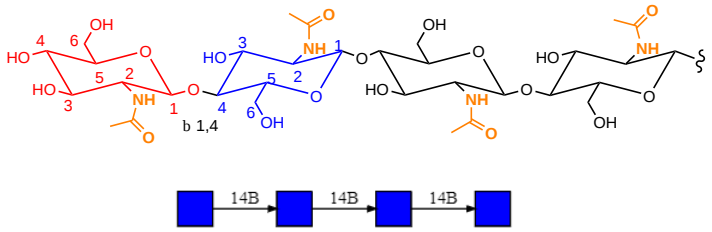


Figure **9.4.10**: Chitin - Chain and symbolic nomenclature for glycans SNFG)

9.4.4: Symbolic nomenclature for glycans (SNFG) -

Before we go further into the complexities of glycan structure, let's explore the symbolic nomenclature for glycan structures. The Consortium for Functional Glycomics (2005) proposed a scheme based on specific colored geometric shapes for each, as shown for the example glycan shown in Figure **9.4.11** for a complex glycan.

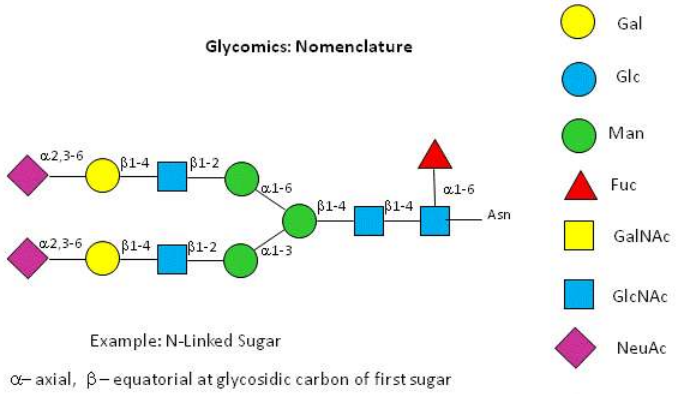


Figure **9.4.11**: SNFG representation of a complex glycan

This nomenclature has recently been updated in Appendix 1B of Essentials of Glycobiology, 3rd Edition (Glycobiology 25(12): 1323-1324, 2015. doi: 10.1093/glycob/cwv091 (PMID 26543186) and is summarized in the Figure **9.4.12**.

SHAPE	White (Generic)	Blue	Green	Yellow	Orange	Pink	Purple	Light Blue	Brown	Red
Filled Circle	Hexose	Glc	Man	Gal	Gul	Alt	All	Tal	Ido	
Filled Square	HexNAc	GlcNAc	ManNAc	GalNAc	GulNAc	AltNAc	AllNAc	TalNAc	IdoNAc	
Crossed Square	Hexosamine	GlcN	ManN	GalN	GulN	AltN	AllN	TalN	IdoN	
Divided Diamond	Hexuronate	GlcA	ManA	GalA	GulA	AltA	AllA	TalA	IdoA	
Filled Triangle	Deoxyhexose	Qui	Rha			6dAlt		6dTal		Fuc
Divided Triangle	DeoxyhexNAc	QuiNAc	RhaNAc							FucNAc
Flat Rectangle	Di-deoxyhexose	Oli	Tyv		Abe	Par	Dig	Col		
Filled Star	Pentose		Ara	Lyx	Xyl	Rib				
Filled Diamond	Nonulosonate		Kdn				Neu5Ac	Neu5Gc	Neu	
Flat Hexagon	Unknown	Bac	LDManHep	Kdo	Dha	DDManHep	MurNAc	MurNGc	Mur	
Pentagon	Assigned	Api	Fru	Tag	Sor	Psi				

Figure 9.4.12: Symbolic Nomenclature for Glycans (SNFG) for the most common monosaccharides

9.4.5: Glycosaminoglycans - Heteropolysaccharides with Disaccharide Repeating Units

Many polysaccharides consist of repeating disaccharide units. A major class of polysaccharides with disaccharide repeats includes the **glycosaminoglycans** (**GAGs**), all of which contain one amino sugar in the repeat, and in which one or both of the sugars contain negatively charged sulfate and/or carboxyl groups. The extent and position of sulfation vary widely between and within GAGs. GAGs are found in the vitreous humor of the eye and synovial fluid of joints, as well as in connective tissue like tendons, cartilage, etc., and skin. They are found in the extracellular matrix and are often covalently attached to proteins to form proteoglycans. From a bird's eye view, they are all elongated polyanions.

They and their structures are very complicated and exceedingly diverse. This makes it difficult for those who want unambiguous structures. From a biological perspective, they present in their local environment an incredibly diverse array of potential binding sites for ligands (both small and large). Because of these, they also have functions in cell signaling. In addition, some GAGs are free-standing, while others are covalently attached to proteins (like glycogen is attached to glycogenin). These large molecules are called **proteoglycans**. We will discuss this later in Chapter 7.4 when we discuss the "carbohydrate code."

Here are the ring structures and descriptions of important GAGs. The common disaccharide repeat unit is shown twice for each structure, with the knowledge that sulfation patterns may differ for the disaccharide repeats in the actual chains. Note also that the first member of each disaccharide repeat shows the ring flipped vertically (top to bottom) as was shown in the structures for other beta-linked glycans (cellulose, chitin) above.

In a long chain, selecting which is the repeating disaccharide unit is a bit relative, as shown in Figure 9.4.13 for the repeating disaccharide sequence of N-acetylglucosamine (blue square) and N-acetylgalactosamine (yellow square).

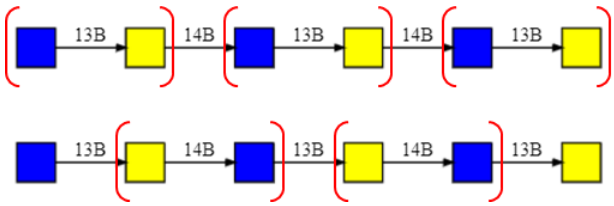


Figure 9.4.13: SNFG representation of a disaccharide sequence of N-acetylglucosamine (blue square) and N-acetylgalactosamine (yellow square)

At the top, the repeating units (blue-yellow) are connected through beta 1,4 links, while at the bottom, the connection of the repeating unit (yellow-blue) is beta 1,3. The best way to annotate the repeating unit without knowing the full chain is elusive. What's most important, however, is to note the alternating acetal/glycosidic links throughout the whole sequence. In the figures below, different disaccharide repeats are highlighted.

Hyaluronic acid

This is a polymer of glucuronate (β 1,3) GlcNAc. It offers a backbone for the attachment of proteins and other GAGs. It's the only GAG without sulfate. Figure 9.4.14 shows a tetrasaccharide fragment with two disaccharide repeats. The illustrated disaccharide repeat's internal acetal/glycosidic link is β 1,3, while the disaccharide's connection is β 1,4. For one last time, the vertical flip of the glucuronic acid is shown to allow a better understanding of its flipped presentation in the actual GAG.

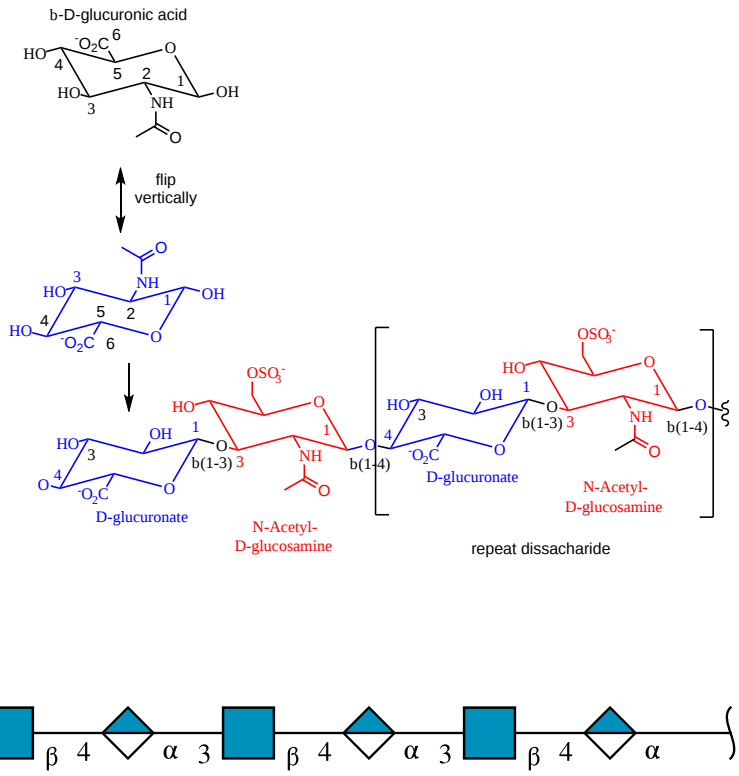


Figure 9.4.14: Hyaluronic acid - tetrasaccharide fragment with two disaccharide repeats

Hyaluronic acids are found in a variety of locations, including synovial fluid, the extracellular matrix, and skin, where they help control skin moisture. It is water-soluble and displays twin antiparallel left-oriented helices. Covalent conjugates of the chemotherapeutic drugs doxorubicin and camptothecin linked to hyaluronic acid, whose overall structure is similar to "worm-like micelles", have been used successfully to treat skin cancers.

Figure 9.4.15 shows an interactive iCn3D model of hyaluronic acid (4HYA). The dotted lines represent hydrogen bonds.

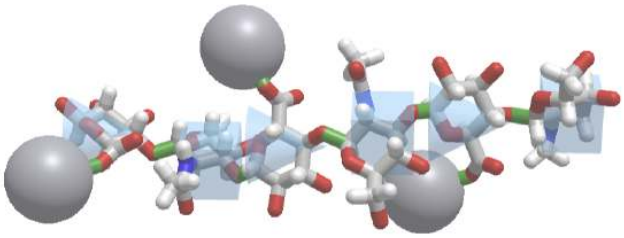


Figure 9.4.15: hyaluronic acid (4HYA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...E7Y2GkhcxRR6n8

Three calcium ions are shown as well. The SNFG cartoon is also illustrated.

Keratan sulfate

This GAG contains repeats of N-acetyl-D-glucosamine-6-phosphate in β 1,3 link to D-galactose or D-galactose-6-sulfate. The link between Gal and the modified glucosamine is β 1,4. Keratin sulfate is most abundant in the cornea of the eye as well as in other connective tissues such as bone, cartilage, and tendon, as well as in the central and peripheral nervous system.

Figure 9.4.16 shows a tetrasaccharide containing two repeating disaccharides.

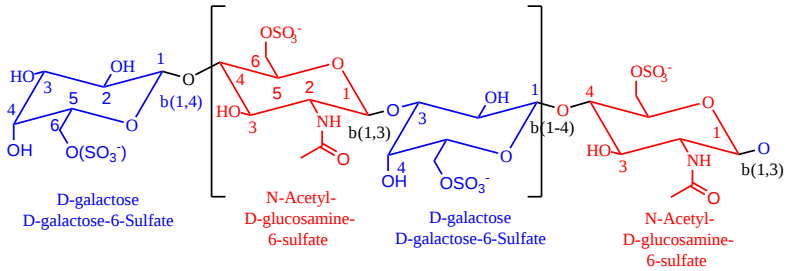
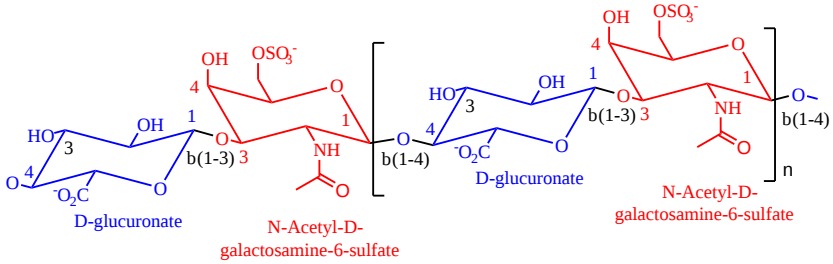


Figure 9.4.16: Keratan sulfate disaccharide repeats

Chondroitin sulfate

The repeat disaccharide unit is D-glucuronate β(1,4) GalNAc-4 or 6-sulfate. It's found in connective tissue matrix, the cell surface (in the form of proteoglycans), basement membranes, and intracellular granules. Figure 9.4.17 shows a tetrasaccharide showing two disaccharide repeats.



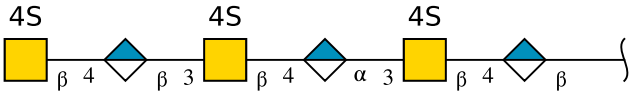


Figure 9.4.17: Chondroitin sulfate disaccharide repeats

Figure 9.4.18 shows an interactive iCn3D model of chondroitin-4-sulfate (1C4S). The dotted lines represent hydrogen bonds.

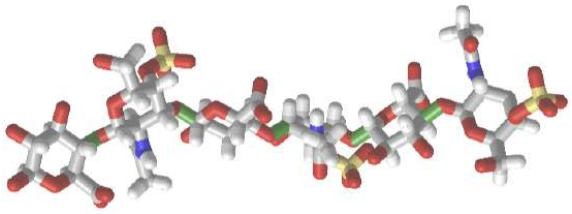


Figure 9.4.18: Chondroitin-4-sulfate (1C4S). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vCXygXckaAVQV8

Dermatan sulfate

This glycosaminoglycan is similar to chondroitin sulfate. It is first made as a polymer of the disaccharide unit of D-gluconic and N-acetyl-D-galactosamine. The gluconic acid is epimerized to L-iduronic acid, followed by sulfation. Its structure is shown in Figure 9.4.19.

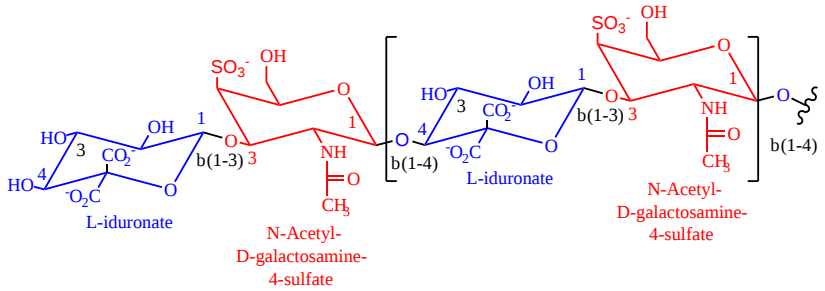


Figure 9.4.19: Dermatan sulfate disaccharide repeats

Heparin

This GAG contains a highly trisulfated disaccharide repeat, as shown in Figure 9.4.20. Note that the molecule can contain glucuronate or iduronate, and the degree of sulfation of the chains varies. Remember, no genetic code specifies these polymers' sequence or sulfation pattern.

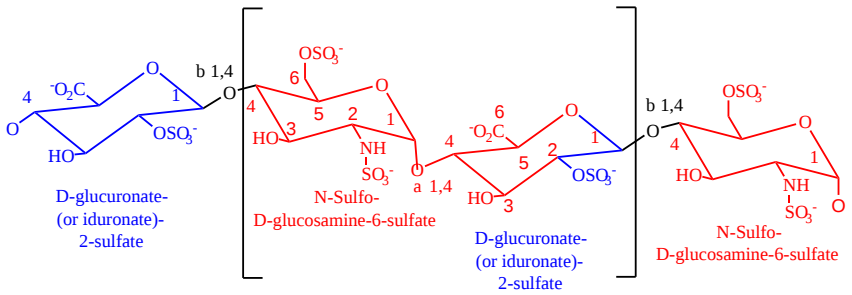


Figure 9.4.20: Heparin disaccharide repeats

Figure 9.4.21 shows an interactive iCn3D model of the solution structure of an 18-mer of heparin (3IRI). The dotted lines represent hydrogen bonds.

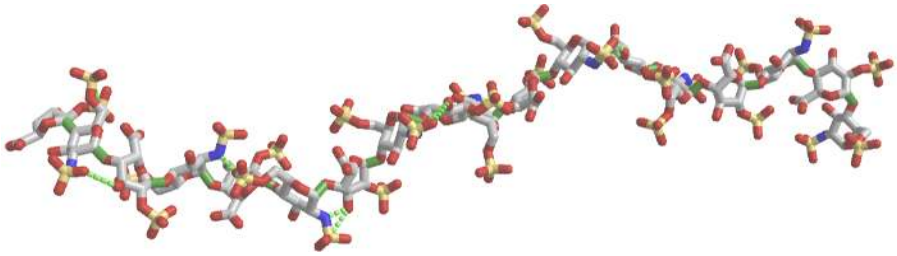


Figure 9.4.21: solution structure of an 18-mer of heparin (3IRI). (Copyright; author via source). Click the image for a popup or use this external link:

Most people are familiar with the anti-clotting properties of heparin administered as a drug. Heparin acts as a "catalyst" to accelerate the inhibition of the enzyme thrombin, which cleaves fibrinogen and activates platelets to form clots by the blood protein antithrombin III. Heparin works in two ways to facilitate thrombin inactivation. It has a specific binding site for antithrombin III, which causes a conformational change in the protein, making it a more effective inhibitor. Thrombin, a positively-charged serine protease, can nonspecifically bind heparin, a polyanion. When it does, it diffuses along the heparin chain, where it can find bound antithrombin III much quicker than if the inhibitor were free in the blood. Heparin effectively changes the search path of thrombin from a 3D to a 1D search.

Figure 9.4.22 shows an interactive iCn3D model of the amino acids in antithrombin III within seven angstroms of a bound heparin 5mer (1NQ9). Dotted lines represent hydrogen bonds and salt bridges between the two. Heparin is highlighted in yellow

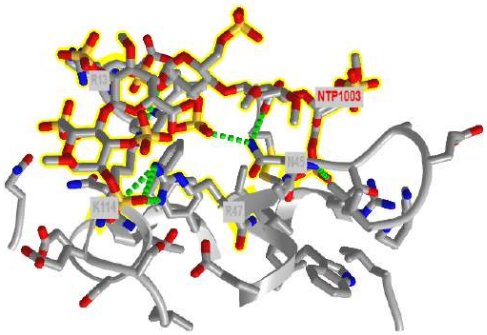


Figure 9.4.22: Heparin 5mer - antithrombin III complex (1NQ9. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fnvt6YfbJovat8

9.4.6: Agarose

Agarose is the main polysaccharide component derived from red algae. Agarose is a polymer of a disaccharide repeat of (1,3)-β-D-galactopyranose-(1,4)-3,6-anhydro-α-L-galactopyranose, is often used for a gelable solid phase for electrophoresis of nucleic acid and as a component of chromatography beads. As with starch, a mixture of amylose and amylopectin, agarose is often found with agaropectin, a sulfated galactan. A tetrasaccharide fragment with two disaccharide repeats is shown in Figure 9.4.23.

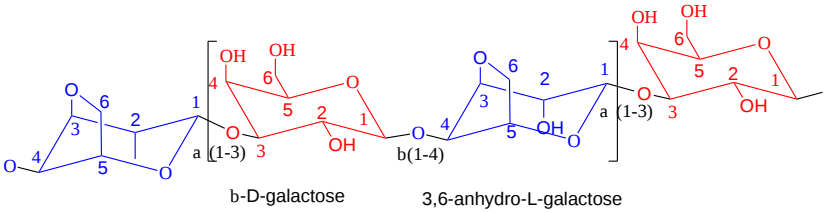


Figure 9.4.23: Agarose disaccharide repeats

Figure 9.4.24 shows an interactive iCn3D model of the agarose double helix (1AGA). The dotted lines represent hydrogen bonds within the polymer.

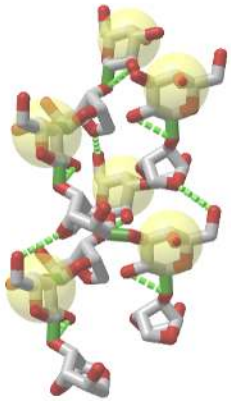


Figure 9.4.24: Agarose double helix (1AGA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5nnzm9XcPogMz7

In the iCn3D model, choose **Style, Glycan, Show Cartoon** to see the yellow sphere for galactose.

9.4.7: Summary

This chapter delves into the intricate world of carbohydrate biochemistry, emphasizing the remarkable complexity and diversity of polysaccharides. It highlights the challenges in studying glycans due to the stereochemical complexity of monosaccharides, the variability of glycosidic linkages, and the lack of a direct genetic template for glycan synthesis. Key themes include:

1. **Fundamental Building Blocks:**
   - The chapter begins by reviewing the structure of monosaccharides, including their linear (Fischer projections) and cyclic forms (Haworth, chair, and wedge/dash representations). Emphasis is placed on the formation of pyranose and furanose rings and the significance of α and β anomers, which influence both structure and function.
2. **Glycosidic Linkages and Polysaccharide Assembly:**
   - Glycosidic bonds are the chemical linkages that join monosaccharide units into polymers. The discussion covers the differences between α and β linkages and their consequences for the overall conformation of the polysaccharide.
   - Homopolymers such as starch, glycogen, dextran, cellulose, and chitin are introduced. Starch and glycogen, composed of glucose units linked by α 1,4 bonds with branching via α 1,6 linkages, serve as energy storage molecules. In contrast, cellulose and chitin, linked by β 1,4 bonds, form rigid, linear structures that provide mechanical support in plant cell walls and exoskeletons, respectively.
3. **Structural Representations and Models:**
   - Various methods for depicting carbohydrate structures are discussed, including two-dimensional diagrams (e.g., Haworth and Fischer projections) and symbolic nomenclature (SNFG) that uses colored shapes to represent different monosaccharides.
   - The chapter utilizes interactive 3D models to illustrate the spatial arrangement of branched glycogen and starch fragments, emphasizing how subtle differences in linkage and branching affect three-dimensional conformation and biological function.
4. **Functional Implications:**
   - The structural diversity of polysaccharides is directly linked to their biological roles. Energy storage polysaccharides like glycogen and starch must be compact and highly branched to allow rapid mobilization of glucose, while structural polysaccharides like cellulose and chitin form extensive networks stabilized by hydrogen bonding and hydrophobic interactions.
   - The chapter also touches on glycosaminoglycans (GAGs), which are heteropolysaccharides with repeating disaccharide units. These molecules, often attached to proteins as proteoglycans, play critical roles in cellular signaling and the formation of the extracellular matrix.
5. **Biological and Clinical Relevance:**
   - Understanding the complexity of carbohydrate structures is essential for insights into cellular function, tissue architecture, and even clinical conditions. For example, variations in sialic acid structures can influence cell recognition and immunity, while dietary components such as glycans impact overall health.

Overall, this chapter provides a comprehensive overview of the chemical, structural, and functional properties of polysaccharides. By integrating detailed structural representations with biological context, it lays the groundwork for further study in glycobiology and its applications in health and disease.

This page titled 9.4: Polysaccharides is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **7.2: Polysaccharides** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls

Learning Goals (ChatGPT o1, 1/30/25)

- **Define Key Terminology and Concepts:**
  - Differentiate among the terms “sugar,” “carbohydrate,” and “glycan” and explain how each applies to both free and protein-conjugated carbohydrates.
  - Understand the structural diversity introduced by glycosylation, and why glycans are considered “information rich” compared to linear biopolymers like DNA and proteins.
- **N-linked Glycoprotein Structure and Diversity:**
  - Describe the core oligosaccharide structure (Man)3(GlcNAc)2(\text{Man})_3(\text{GlcNAc})_2 attached to Asn in the consensus sequence (X-Asn-X-Thr/Ser) and its significance in N-linked glycoprotein formation.
  - Differentiate between high-mannose, complex, and hybrid N-linked glycans by comparing their outer sugar compositions (e.g., presence or absence of mannose, lactosamine repeats, or branching patterns) and discuss their functional implications.
  - Utilize SNFG (Symbolic Nomenclature For Glycans) to interpret and represent glycan structures and variations.
- **O-linked Glycoproteins and Blood Group Antigens:**
  - Explain the process of O-linked glycosylation, where carbohydrates (commonly Gal(β1,3)GalNAc) attach to Ser/Thr residues, and relate this to the structure and function of blood group antigens.
  - Recognize how variations in O-linked glycan structures contribute to cell–cell recognition and immune responses.
- **Proteoglycans and the Extracellular Matrix (ECM):**
  - Describe the structure of proteoglycans, including the linkage of glycosaminoglycans (GAGs) to core proteins, and differentiate between soluble and membrane-bound forms.
  - Explain the roles of proteoglycans in the ECM, including their contribution to tissue structure, cellular adhesion, and signal transduction.
- **Carbohydrate Structures in Cell Walls and Membranes:**
  - Compare and contrast the composition and structure of cell walls in different organisms:
    - For Gram-positive bacteria: Describe the peptidoglycan network, including the roles of NAG and NAM, pentapeptides, and teichoic acids.
    - For Gram-negative bacteria: Explain the arrangement of the peptidoglycan layer, outer membrane, and the role of lipopolysaccharides (LPS) in antigenicity and antibiotic resistance.
    - For plants: Summarize the components of primary and secondary cell walls (cellulose, hemicellulose, pectin, and lignin) and discuss how these polymers contribute to cell rigidity and growth.
  - Discuss the unique features of archaeal cell membranes and walls, including differences in lipid composition and the absence of peptidoglycan.
- **Biological and Clinical Implications of Glycosylation:**
  - Evaluate the functional roles of glycosylation in protein folding, stability, proteolytic protection, and in modulating protein half-life.
  - Analyze how specific glycan structures (e.g., Gal(α1,3)Gal) affect immune recognition and their implications in xenotransplantation and diseases such as alpha-gal syndrome.
  - Discuss how differences in glycosylation patterns contribute to cell signaling and tissue-specific functions, particularly within the ECM and basement membranes.

By achieving these goals, students will gain a comprehensive understanding of the multifaceted roles that carbohydrate modifications play in protein function and cellular organization, and they will be prepared to explore further applications of glycosylation in health, disease, and biotechnology.

Many proteins, especially those destined for secretion or insertion into membranes, are post-translationally modified by the attachment of carbohydrates. They are usually attached through either Asn or Ser side chains. Carbohydrate modifications on the protein appear to be involved in recognizing other binding molecules, preventing aggregation during protein folding, protecting from proteolysis, and increasing the protein's half-life. In contrast to a protein sequence determined by a DNA template, sugars are attached to proteins by enzymes that recognize appropriate sites on proteins and attach the sugars. Since there are many sugars with many functional groups that can serve as potential attachment sites, the structures of the oligosaccharides attached to proteins are enormously varied, complex, and hence "information-rich" compared to linear or folded polymers like DNA and proteins.

9.5.1: N-linked Glycoproteins

These contain carbohydrates attached through either a GlcNAc or GalNAc to an Asn in an X-Asn-X-Thr sequence of the protein. There are three types of N-linked glycoproteins: high mannose, complex, and hybrid. They all contain the same core oligosaccharide - (Man)3(GlcNAc)2 attached to Asn as shown in Figure **9.5.1**.

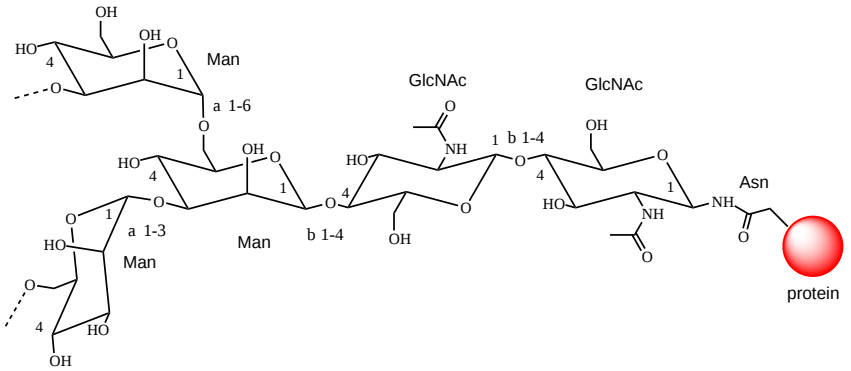


Figure **9.5.1**: Core oligosaccharide - (Man)3(GlcNAc)2 attached to Asn in N-linked glycoproteins

Table 9.5.1 below shows the SNFG representation for the main core and variant glycans of **N-linked glycoproteins**. Note that the designation of α2 implies an α(1→2) linkage. Unless otherwise stated, the linkage is presumed to start from carbon 1.

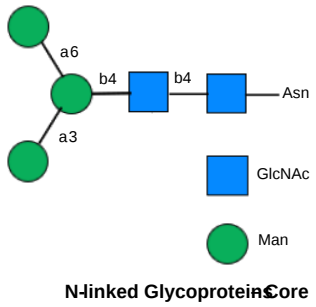
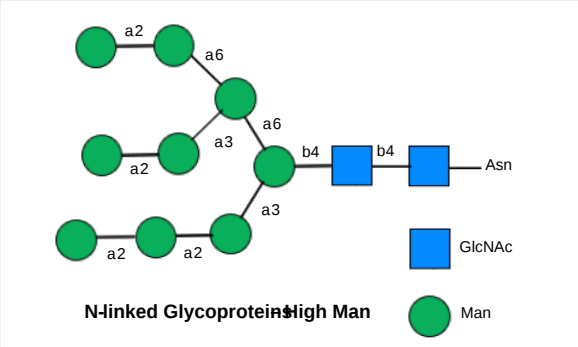
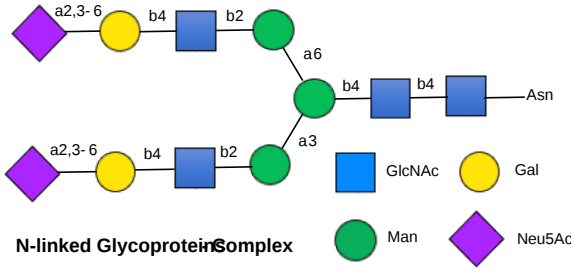
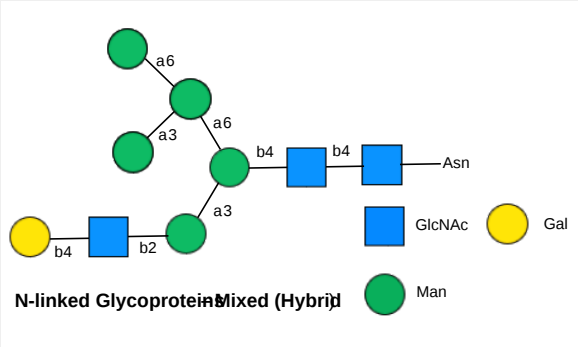
Core	
High mannose	
Complex	
Mixed hybrid	

Table 9.5.1: SNFG representation for the main core and variant glycans in **N-linked glycoproteins**

Complex N-linked glycans don't contain mannose outside the core glycan and have GlcNAc attached to the branching mannoses in the core structure. The complex glycan shown above has a Gal(β1,4)GlcNAc sequence, which could be named the disaccharide lactosamine. Often, lactosamine repeats in the sequence.

Hybrid glycans have both unsubstituted terminal mannoses (as in the high-mannose type) and substituted mannoses with an N-acetylglucosamine attached (as in the complex type. GlcNAc residues added to the core in the hybrid and complex N-glycoproteins are called antennae. Figure 9.5.2 shows an example of a biantennary N-linked glycan with two GlcNAc branches linked to the core. The core is outlined in red, and the two GlcNAcs are labeled 1 and 2.

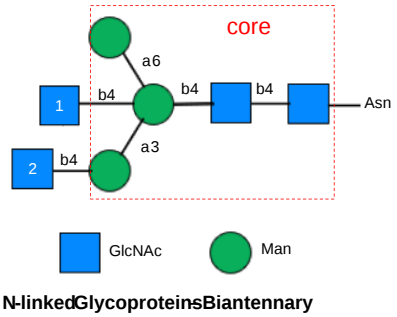


Figure 9.5.2: Biantennary N-linked glycan with two GlcNAc branches linked to the core

Complex glycans also have bi-, tri-, and tetraantennary forms and comprise most N-glycans. As shown in Table 9.5.1 above, complex N-linked glycans usually end with sialic acid residues. About 50% of the surface area of the COVID-19 SARS-CoV-2 spike protein is covered with glycans, as shown in the model structure in Figure 9.5.3. The protein surface is gray, and the glycans (biantennary LacNAc N-glycans) in spacefill CPK with carbon in cyan.

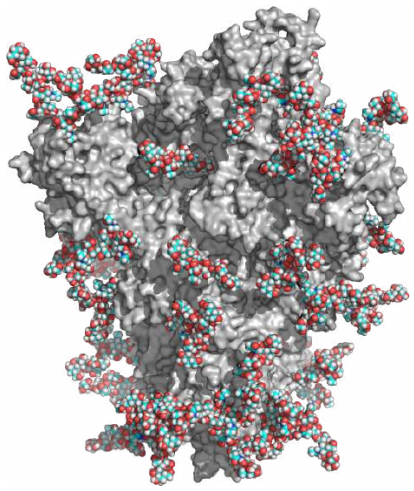


Figure 9.5.3: COVID SARS-CoV-2 spike protein is covered with glycans (colored spacefill). (PDB file coordinates 5.Swiss.3.M3F1.CYX.TER from Analysis of the SARS-CoV-2 spike protein glycan shield: implications for immune recognition. Oliver C. Grant, David Montgomery, Keigo Ito, Robert J. Woods. **doi:** https://doi.org/10.1101/2020.04.07.030445)

In the hybrid oligosaccharide shown above, one terminus contains Gal(β1,4)GlcNAc. However, in all other mammals except man, apes, and Old World monkeys, an additional Gal is often connected in an α1,3 link to the Gal to give a terminus of Gal(α1,3)Gal(β1,4)GlcNAc. These animals have an additional enzyme, an α1,3 Gal transferase. Bacteria also have this enzyme, and since we have been exposed to this link through bacterial infection, we mount an immune response against it. Why is this important? Pig hearts are similar to human hearts so that they might be good candidates for human transplantation (xenotransplants). However, the Gal-α1,3-Gal link is recognized as foreign, and we mount a significant immune response against it. Several biotech firms are trying to delete the pig α1,3 Gal transferase, which would prevent the addition of the terminal Gal and make them good donors for transplanted hearts.

Figure 9.5.4 shows an interactive iCn3D model of an N-linked glycoprotein, human beta-2-glycoprotein-I (Apolipoprotein-H) (1C1Z).

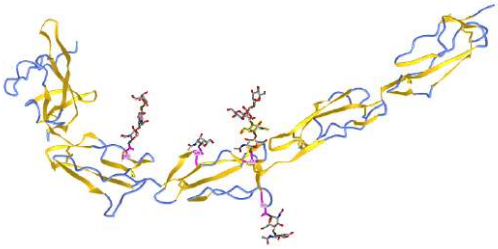
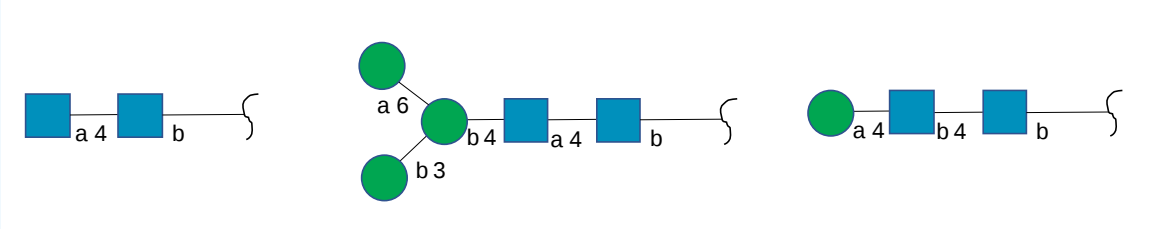


Figure 9.5.4: N-linked glycoprotein, human beta-2-glycoprotein-I (Apolipoprotein-H) (1C1Z). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...GkZSVFdciXiBT7

? Exercise 9.5.1

The glycan structures for the beta-2-glycoprotein-I are shown below. Identify the monosaccharides in each and specify to which asparagine they are linked.



Answer

Here is an interactive iCn3D image showing the protein and attached glycans using SNFG notation.

Figure 9.5.5 shows an interactive iCn3D model of the GP120 HIV protein that contains high mannose, complex, and hybrid N-linked glycans. Most glycoproteins in the Protein Data Bank do not contain attached glycans. The glycans here were added with the program GlyProt at 3 of 17 possible Asn residues that would presumably have attached glycans. Use your mouse or keypad to hover over the monomers in the attached glycans. Abbreviations for the given residues in the model are: adm = alpha-D-Man, bdg= beta-D-Glc or Gal, adn = alpha-D-neuraminidase.

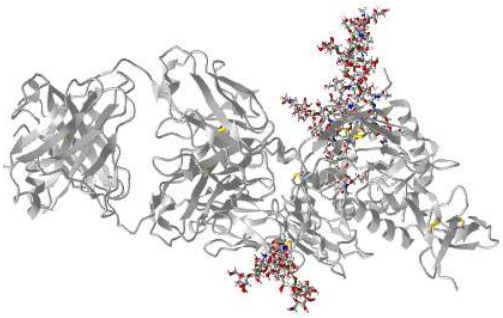


Figure 9.5.5: GP120 HIV protein with high mannose, complex, and hybrid N-linked glycans. (Copyright; author via source). No external link is available for this structure.

The coronavirus pandemic has been deadly (over 1.1 million deaths in the USA alone and 19-36 million around the world). However, the 1918 influenza pandemic was far worse per capita, with an estimated 650,000 deaths in the USA and 50 million around the world in a population less than 1/3 of the present. An additional 3 million deaths in the USA were probably prevented by vaccination as of December 2022. Many deaths in the developing world would have been prevented if wealthy nations had allocated more resources to produce and distribute the vaccines. The evolution of the virus in nonvaccinated areas might come back to haunt wealthy countries if present vaccines become ineffective against the mutants. A worse pandemic might await us. An avian version of the influenza virus (H5N1), presently endemic in wild birds and now found in mink populations, has infected 240 people as of January 2023 and killed 56% of them. A quick note: the 1918 pandemic affected youth the most.

Influenza and the Avian Flu

Figure 9.5.6 shows the simple yet deadly influenza virus. It interacts with human cells through a surface protein, hemagglutinin (HA).

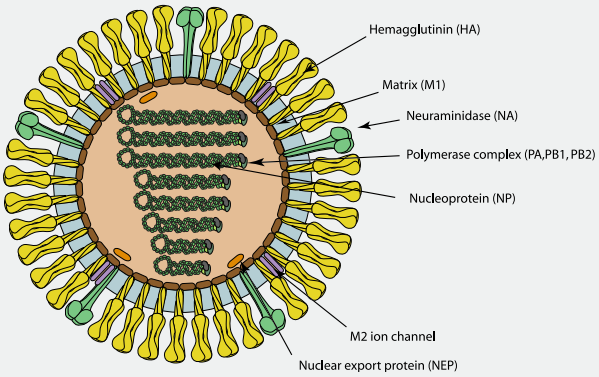
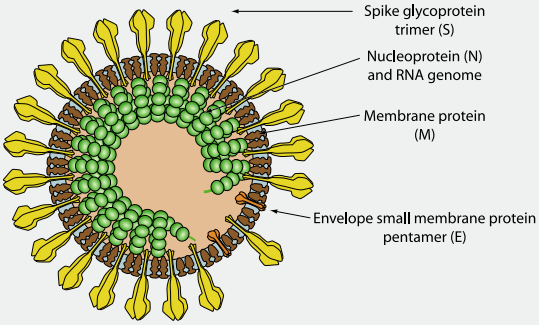


Figure 9.5.6: Alpha Influenza virus. https://viralzone.expasy.org/6. Creative Commons Attribution 4.0 International (CC BY 4.0) License.

Note the similarities and differences with the SARS-CoV-2 virus, shown below.



SARS-Covid-2 virus. https://viralzone.expasy.org/764. Creative Commons Attribution 4.0 International (CC BY 4.0) License.

The virus binds to host cells through the interaction of HA with cell surface carbohydrates. Once bound, the virus internalizes, ultimately releasing the viral RNA genome into the host cell.

The hemagglutinin protein is the most abundant protein on the viral surface. 15 avian and mammalian variants have been identified (based on antibody studies). Only three have adapted to humans in the last 100 years, giving pandemic strains H1 (1918), H2 (957), and H3 (1968). Three recent avian variants (H5, H7, and H9) have jumped directly to humans recently, but have low human-to-human transmissibility.

The influenza hemagglutinin protein has the following characteristics:

- the mature form is a homotrimer (3 identical protein subunits), MW 220,000 with multiple sites for covalent attachment of sugars. Hemagglutinin is a glycoprotein.
- each monomer is synthesized as a single polypeptide chain precursor (HA0), which is cleaved into HA1 and HA2 subunits by the protease trypsin in epithelial cells of the lung.
- structure known for human (H3), swine (H9), avian (H5) subtypes.

Figure 9.5.7 below shows an interactive iCn3D model of an influenza hemagglutinin trimer (6HJQ). The white backbone traces on the membrane's extracellular (red) side are antibody molecules used to stabilize the structure for crystallization.

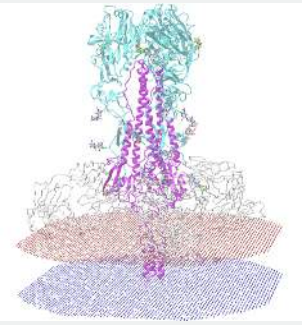


Figure 9.5.7: Influenza Hemagglutinin Trimer (6HJQ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...EhePsF4BiEjwA8

The protein is a trimer of a heterodimer, $(HA1:HA2)_3$. Each heterodimer has two separate chains, HA1 and HA2. Together they form a globular head and stalk region, the latter crossing the membrane. HA1 interacts primarily with the sialic acid on host cells' surface proteins, while HA2 is more involved in the fusion of the virus with host cells and internalization.

Hemagglutinin binds to sialic acid (Sia) covalently attached to many cell membrane glycoproteins. The sialic acid is usually connected through an α(2,3) or α(2,6) link to galactose on N-linked glycoproteins. The subtypes found in avian (and equine) influenza isolates bind preferentially to Sia (α2,3) Gal, which predominates in the avian GI tract where viruses replicate. Human influenza isolates prefer Sia α(2,6)

Sia (α2,3) Gal predominates in the avian GI tract, where viruses replicate. Human influenza isolates prefer Sia α(2,6) Gal. The human virus of H1, H2, and H3 subtypes (causes of the 1918, 1957, and 1968 pandemics) recognizes Sia α(2,6) Gal, the major form in the human respiratory tract. The swine influenza HA binds to Sia α(2,6) Gal and some Sia (α2,3) Gal, both found in swine. The structures of the Sia-Gal disaccharide are shown in Table 9.5.2 below.

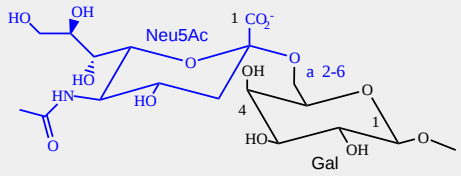
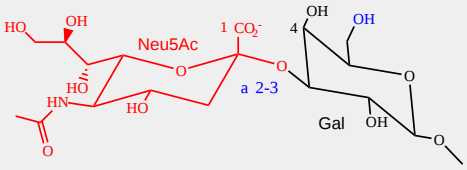
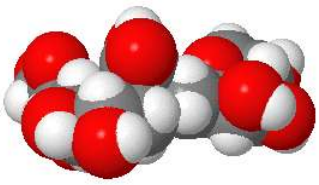
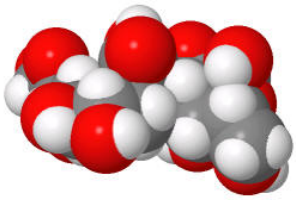
Sia α(2,6) Gal (Human)	Sia α(2,3) Gal (Avian and some Swine)
	
 (made with Sweet, with an OH, not AcNH on sialic acid on C5)	 (made with Sweet, with an OH, not AcNH on sialic acid on C5)

Table 9.5.2: Structures of Sia α(2,6) Gal (human) and Sia α(2,3) Gal (avian/swine)

The H5N1 avian flu H5N1) virus is deadly but presently lacks human-to-human transmissibility. Why? One reason is that it appears to bind deep in the lungs and is not released easily on coughing or sneezing. It appears that cell surface glycoproteins deeper in the respiratory tract have Sia (α2,3) Gal which accounts for this pathology.

Before it leaves the cell, the virus forms a bud on the intracellular side of the cell with the HA and NA in the cell membrane of the host cell. The virus in this state would not leave the cell since its HA molecules would interact with sialic acid residues in

the host cell membrane, holding the virus in the membrane. Neuraminidase hydrolyzes sialic acid from cell surface glycoproteins, allowing the virus to complete the budding process and be released from the cell as new viruses. The drugs Oseltamivir (Tamiflu) and Zanamivir (Relenza) bind to and inhibit neuraminidase, whose activity is necessary for viral release from infected cells. Tamiflu appears to work against N1 of the present H5N1 avian influenza viruses. Governments across the world are hopefully stockpiling this drug in case of a pandemic caused by the avian virus jumping directly to humans and becoming transmissible from human to human.

9.5.2: O-linked Glycoproteins

The CHOs are usually attached from a Gal (β 1,3) GalNAc to a Ser or Thr of a protein, as shown in Figure 9.5.8.

Figure: O-linked Glycoproteins

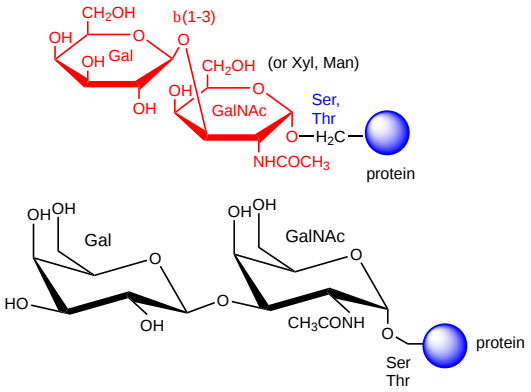


Figure 9.5.8: O-linked Glycoproteins

The blood group antigens (CHOs on cells attached to either proteins or lipids) are examples. The sugars shown as chairs (in contrast to structures found in many texts) in Figure 9.5.9 are the blood group antigens. They are attached to a core heterosaccharide (a red ellipse below), which is connected to either a membrane glycoprotein or glycolipid.

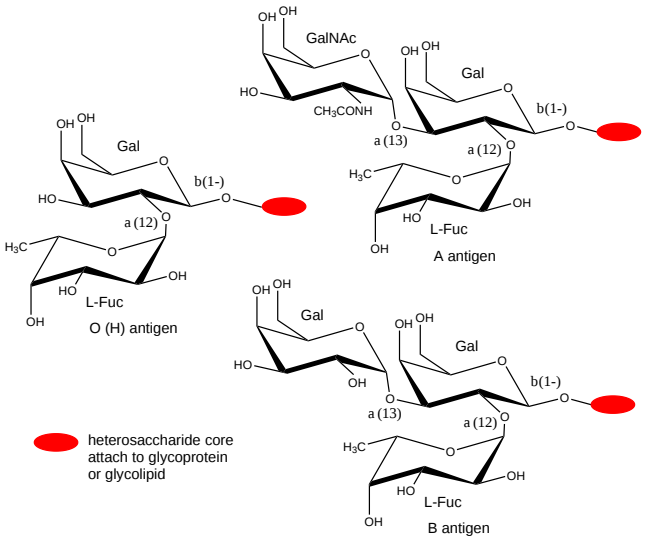


Figure 9.5.9 : Blood group glycans

Figure 9.5.10 shows the SNFG representation for the A antigen in the glycolipid form.

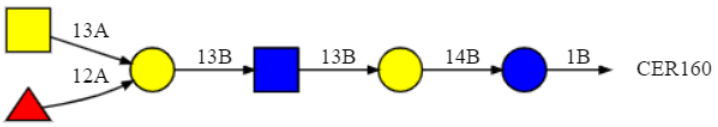


Figure 9.5.10: SNFG representation for the A antigen in the glycolipid form

The trimeric branched residues on the left-hand side represent the A antigen shown above. The red triangle is L-fucose. Yellow represents galactose or GalNac, while blue is glucose or GlcNAc.

9.5.3: Proteoglycans

Some proteins are so modified with CHOs that they contain more CHOs than amino acids. Proteins linked to glycosaminoglycans are called proteoglycans (PGs). They consist of a core protein linked to one or more glycosaminoglycans. GAGs are linear sulfated glycans, which we described earlier. The structures of a few proteoglycans are known. The GAGs are O-linked to the protein, typically to a Ser of a Ser-Gly dipeptide, often repeated in the protein. Some of the proteoglycans also contained N-linked oligosaccharide groups. Figure 9.5.10 represents a proteoglycan structure.

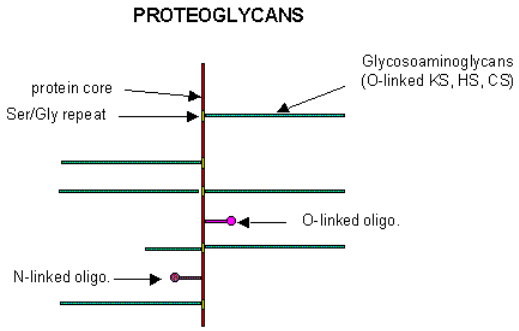


Figure 9.5.10: Representation of proteoglycan structure.

PGs can be soluble and are found in the extracellular matrix or as integral membrane proteins. There are about 43 genes for proteoglycans. Differential splicing of the RNA transcripts gives rise to soluble and transmembrane forms. Given the diversity of sugars and the varying extent of sulfation, the CHO part of PGs provides an incredible variety of binding structures at or near the cell surface. Figure 9.5.11 shows the variety of proteoglycans found in mammalian cells. PGs help form the extracellular matrix, which provides a rich binding environment between cells.

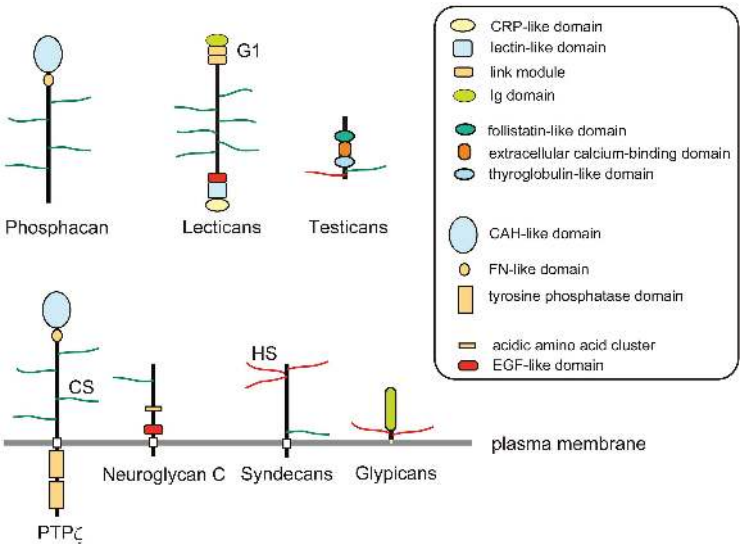


Figure 9.5.11: Proteoglycans found in mammalian cells

One PG, syndecan, binds through its intracellular domain to the internal cytoskeleton of the cell while interacting with another protein - fibronectin - in the extracellular matrix. Fibronectin also binds other molecules, regulating cellular growth and other

interactions. PGs act like glue in connecting the extracellular and intracellular functions of the cell. There are four different core syndecan proteins (SDCs 1–4), with SDC4 lacking the cytoplasmic and transmembrane domains and is a soluble form in the intracellular matrix. The glycan components of syndecans are mostly heparan sulfate, while SDC 1 and 3 also have two chondroitin sulfate chains.

Most proteins bind PGs through a PG binding motif of BBXB or BBBXXB, where B is a basic amino acid. Some proteins bind to specific sequences in specific GAGs. For instance, antithrombin 3, an inhibitor of blood clotting, binds specifically to heparin. This enhances its interaction with clotting proteins such as thrombin and Factor Xa. Figure **9.5.12** shows an interactive iCn3D model of a five-residue fragment of heparin interacting with the key amino acids' side chains of Factor Xa (2gd4).

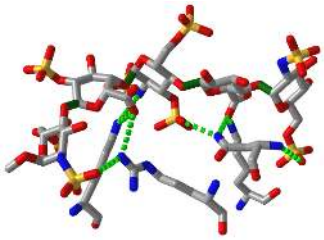


Figure **9.5.12**: Heparin-5mer interactions with antithrombin III in a complex with Factor Xa. (2gd4) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5nZ99mUhULUi59

The extracellular matrix (ECM) might appear to be a nondescript mess for those more chemically oriented, since chemists are used to well-defined structures. Figure **9.5.13** shows a cartoon of the ECM and may clarify the components. Few structure files exist for them, given the inherent flexibility of the glycan components.

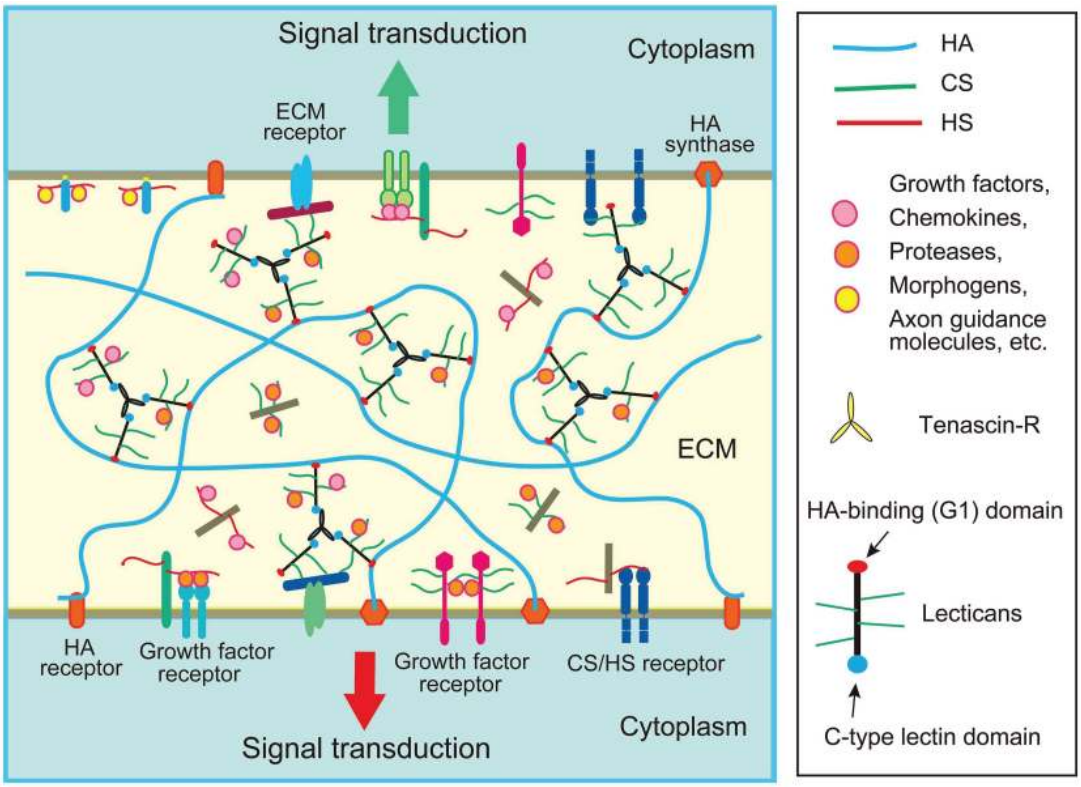


Figure **9.5.13**: Cartoon representation of the extracellular matrix. Frontiers in Neuroscience 9(50):98 (2015). DOI: 10.3389/fnins.2015.00098. CC BY 4.0

9.5.4: Cell Walls and Glycolipids

In contrast to eukaryotic cells, bacteria, and plant cells have a cell wall in addition to a lipid bilayer membrane. These are essentially carbohydrate polymers that determine cell shape, affording protection from exterior pathogens, hypotonic conditions, and high internal osmotic pressures, preventing swelling and bursting of the cells. This is especially important in plants, which need

strength and rigidity against the "turgor" pressure of the aqueous cytoplasm against the cell membrane. This prevents wilting in plants. The cell walls in plants and probably bacteria are involved in cell signaling across the cell membrane.

9.5.4.1: Bacterial Cell Walls

Two types of cell walls occur.

9.5.4.1.1: a. Gram-positive bacteria-

These bacteria can be stained with a Gram stain. The wall consists of a GlcNAc (β 1,4) MurNAc repeat. (GlcNAc is often abbreviated as NAG, while MurNAc is abbreviated as NAM.) This is similar to the GlcNAc (β 1,4) GlcNAc homopolymer chitin, except that every other GlcNAc contains a lactate molecule covalently attached in an ether-linkage to the C3 hydroxyl to form the monomer N-Acetylmuramic acid. A pentapeptide (Ala-D-isoGlu-Lys-D-Ala-D-Ala) is attached through an amide link to the carboxyl group of the lactate in MurNAc. A pentaglycine bridge covalently connects the GlcNAc (β 1,4) MurNAc strands through the epsilon amino group of the pentapeptide Lys on one strand and the terminal D-Ala of a pentapeptide on another strand. A small part of the structure of a gram-positive bacterial cell wall is shown in Figure 9.5.14. It shows one repeating GlcNAc-MurNAc disaccharide unit in front (darker) and one in the back (lighter) connected through the peptides shown.

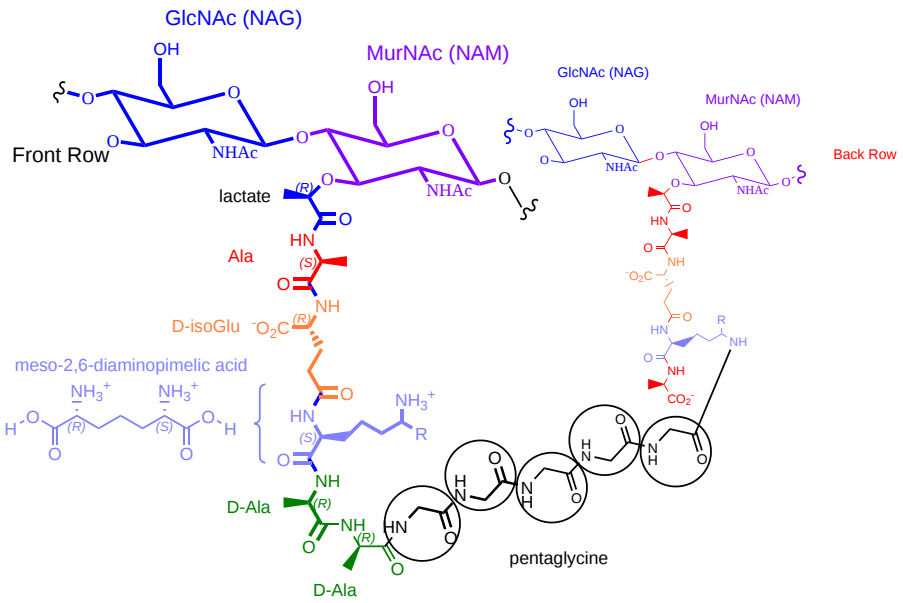


Figure 9.5.14: Part of the structure of a gram-positive bacterial cell wall

The SNFG representation of a larger section of the gram-positive cell well is shown in Figure 9.5.15.

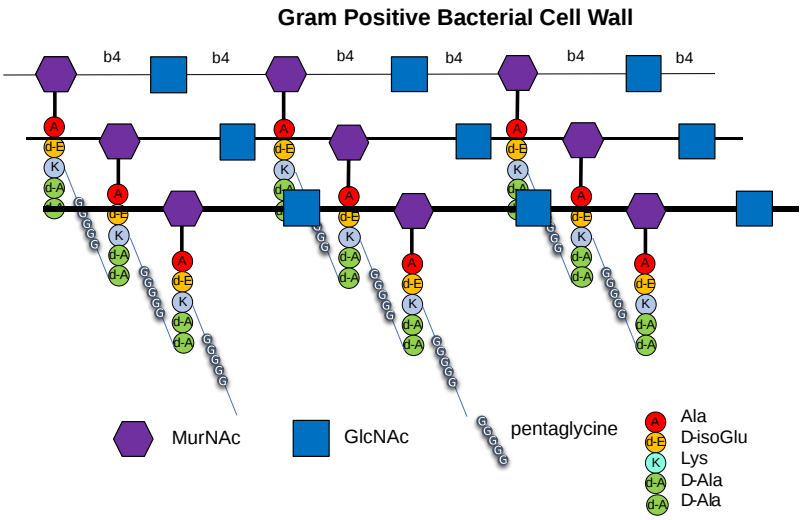


Figure 9.5.15: SNFG representation of a larger section of the gram-positive cell wall.

One final structure is found in Gram-positive peptidoglycan cell walls. Teichoic acids are often attached to carbon 6 of MurNAc. Teichoic acid is a polymer of glycerol or ribitol with alternative GlcNAc and D-Ala linked to the middle C of the glycerol. Multiple glycerols are linked through phosphodiester bonds. These teichoic acids often make up 50% of the dry weight of the cell wall and present a foreign (or antigenic) surface to infected hosts. These often serve as receptors for viruses that infect bacteria (called bacteriophages). Its structure is illustrated in Figure 9.5.16.

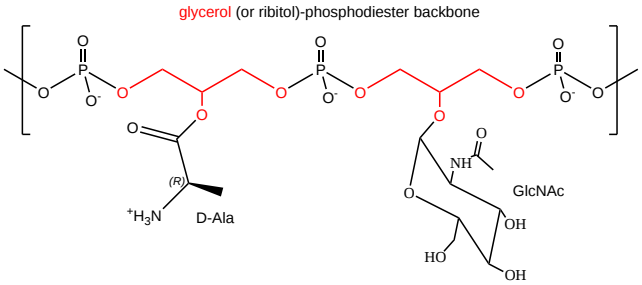


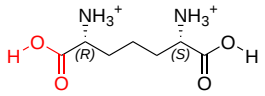
Figure 9.5.16 : Teichoic acid

Notice that all monomeric units of peptidoglycan and attached teichoic acid derivatives are covalently attached to form one large molecule comprising the entire cell wall! This structure, along with the Gram-negative cell wall structures, is the largest single macromolecule in nature.

9.5.4.1.2: *b. Gram-negative bacteria*

These bacteria can NOT be stained with a Gram stain. The wall consists of the same structure as in Gram-positive bacteria. However, the GlcNAc (β 1,4) MurNAc strands are covalently connected through a direct amide bond between a derivative of Lys, meso-diaminopimelic acid (m-A2pm), on one peptide strand and to the last D-Ala of a pentapeptide on another strand. (i.e., there is no penta-Gly spacer). The connector peptide is Ala-D-isoGlu-m-A2pm-D-Ala-D-Ala

m-A2pm replaces Lys 3 of the peptide in most Gram-negative species and Gram-positive bacteria of the genus *Bacillus* and *mycobacteria*. *The stereochemistries at each chiral center are different (R and S)*, but because the molecule has a plane of symmetry, it is an example of a meso-compound, a diastereoisomer of a molecule that does not have a different enantiomeric version. The structure is shown in Figure 9.5.17.



meso diaminopimelic acid

Figure 9.5.17: Meso-diaminopimelic acid

A small part of the structure of a Gram-negative bacterial cell wall is shown in Figure 9.5.18.

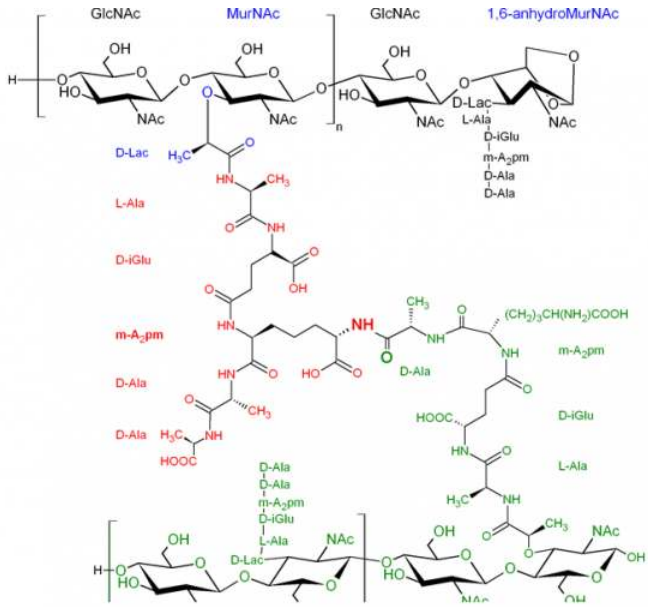


Figure 9.5.18: Part of the structure of a Gram-negative bacterial cell wall. http://www.glycopedia.eu/e-chapters/...ular-structure. https://gagdb.glycopedia.eu/license

Figure 9.5.19 shows an image of a computed model (not a crystal or NMR structure) of the Gram-negative peptidoglycan of E. Coli. Jame Gumbart kindly provided the PDB coordinates. The repeating $(GlcNAc\text{-}MurNAc)_n$ are shown in sticks, alanines in red spheres, m-A2-pm (meso-diaminopimelic acid) in gray spheres, and DiG in orange spheres. The glycine pentapeptide is not shown since it is not found in Gram-negative bacteria.

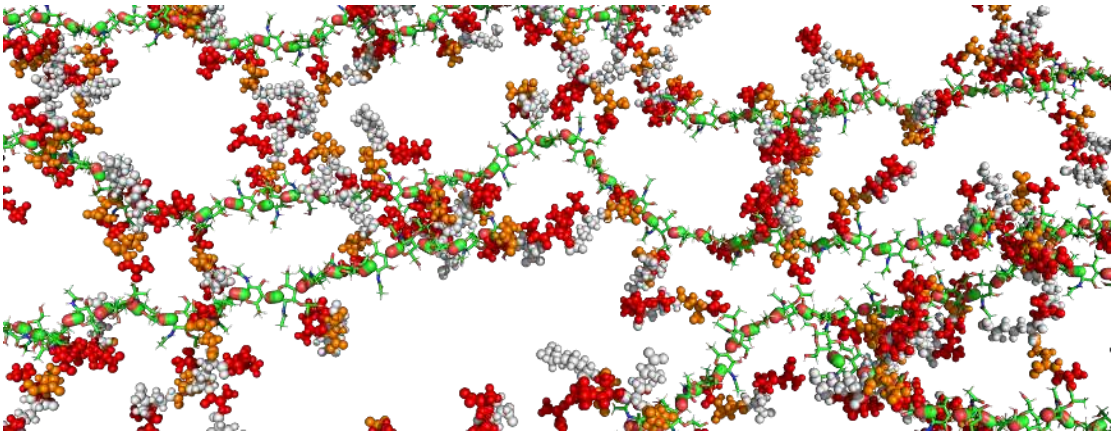


Figure 9.5.19: Part of a Gram-negative peptidoglycan of E. Coli. PDB coordinates kindly provided by Jame Gumbart.

Follow these instructions to get an interactive iCn3D model of the structure.

- open iCn3D

- Download this file to your computer's download folder. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!
- **File, Open File, iCn3D png (appendable),** and choose the downloaded file (it takes a while).

In addition, Gram-negative bacteria don't have teichoic acid polymers. Rather, they have a **second, outer lipid bilayer**. The cell wall peptidoglycan (PG) is sandwiched between the inner and outer bilayers. The space between the lipid bilayers is called the periplasm. The outer leaflet of the outer membrane is coated with a **lipopolysaccharide (LPS)** (a glycolipid) of varying composition. The LPS determines the antigenicity of the bacteria. The different LPS are called the O-antigens. Figure 9.5.20 shows the structure of the Gram-negative bacterial membrane organization. (In the figure, PS is LPS, PG is peptidoglycan). The LPS in the outer leaflet is amphiphilic, with nonpolar acyl chains forming a more classic bilayer, and the inner leaflet phospholipid and polar/charged sugars form the LPS fringe. The extra membrane of Gram-negative bacteria makes them the major source of antibiotic resistance.

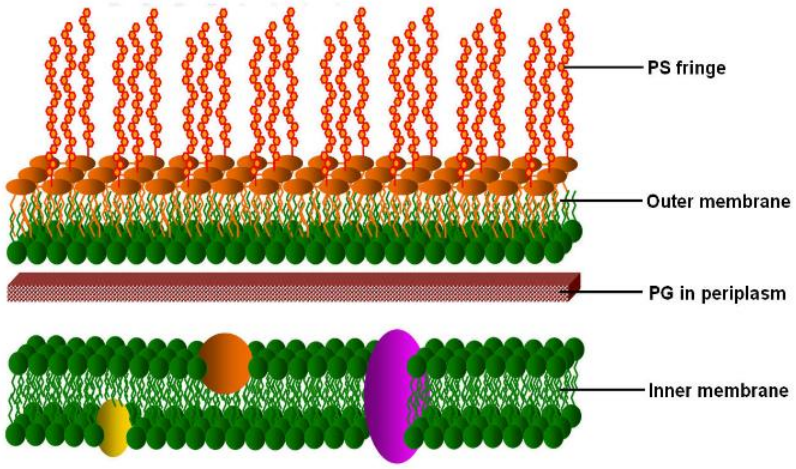


Figure 9.5.20: Overall structure of the Gram-negative bacterial membrane organization. https://cronodon.com/BioTech/Bacteria_envelope.html

A detailed view of the structure of the lipopolysaccharide (LPS) from *Salmonella typhimurium* is shown in Figure 9.5.21 below.

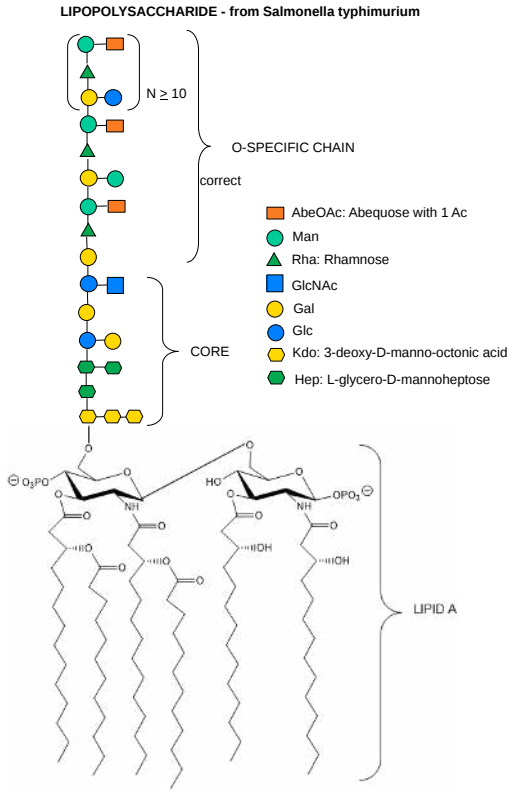


Figure 9.5.21: Lipopolysaccharide (LPS) from *Salmonella Typhimurium*

Recent Updates: 11/8/24

LPS is synthesized in the inner leaflet of the inner membrane and must translocate all the way to the outer leaflet of the outer membrane. Movement requires an ATP-binding cassette transporter $LptB_2FG$. Figure 9.5.22 shows the machinery used to move LPS to the outer membrane. We'll discuss membrane proteins and transport thoroughly in Chapter 11.

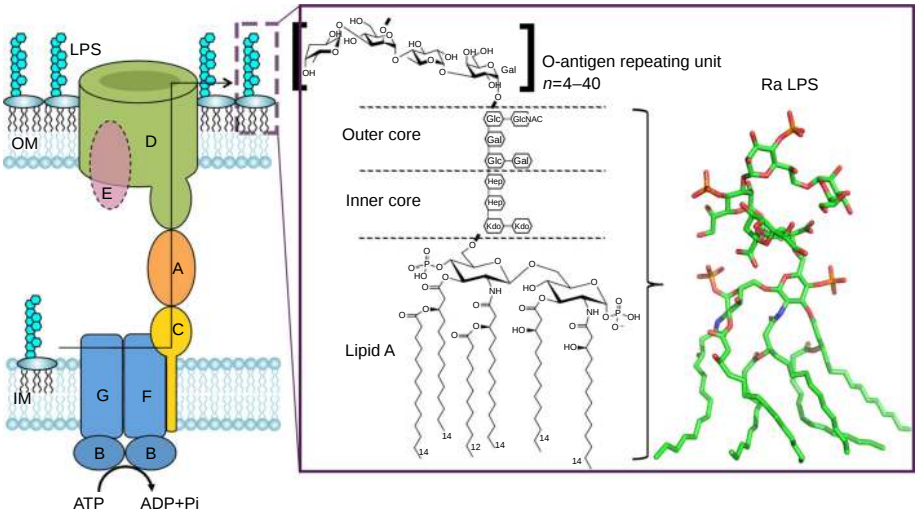
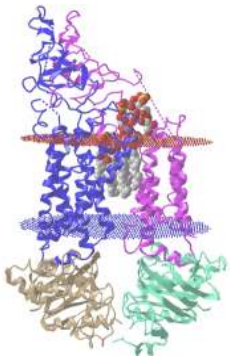


Figure 9.5.22: LPS transport from the IM to the OM by the trans-envelope complex LptABCDEFG. Dong, H., Zhang, Z., Tang, X. *et al*. Structural and functional insights into the lipopolysaccharide ABC transporter $LptB_2FG$. *Nat Commun* **8**, 222 (2017).

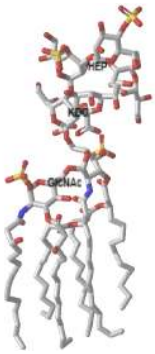
https://doi-org.ezproxy.csbsju.edu/1...67-017-00273-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Legend: LPS is extracted from the periplasmic side of the IM by the ABC transporter $LptB_2FG$, and is delivered to an IM protein LptC, which forms a complex with $LptB_2FG$. LptC comprises a single membrane-spanning domain and a large periplasmic domain, forming a periplasmic bridge with LptA and the N-terminal domain of LptD. LPS is then inserted into the OM by the LptD/E complex. LPS contains O-antigen, core oligosaccharide, and lipid A components, of which the O-antigen has 4-40 O-antigen repeat units. Ra-LPS, rough LPS, Kdo, 3-deoxy-D-manno-oct-2-ulosonic acid, Hep, L-glycero-D-manno-heptose, Glc, D-glucose, Gal, D-galactose.

Figure 9.5.23 shows an interactive iCn3D model of the E. Coli lipopolysaccharide ABC transporter LptB2FG (6MHU). The left panel shows the membrane protein that binds and transports the LPS (spacefill) from the inner to the outer membrane. The right panel shows the bound LPS. Just part of the inner core is shown since other parts were likely too flexible to be observed. The outer O antigen is not shown since a mutation in the protein prevented it. Six acyl chains are evident. The phosphorylated GlcNAc, KDO, and HEP "layers" are labeled in the right panel.



Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CoPtdH2RVY82P6. Click Style, Background, Transparent in the iCn3D window for a better view.



Click the image for a popup (instructions below)

Figure 9.5.23: E. Coli lipopolysaccharide ABC transporter LptB2FG (6MHU) (left panel) and the bound LPS (right panel). (Copyright; author via source).

To see an interactive iCn3D model of the E. Coli LPS from the right panel, follow these steps:

- open iCn3D
- Download this file to your computer's download folder. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!
- **File, Open File, iCn3D png (appendable),** and choose the downloaded file.

9.5.4.1.3: c. Archaeal Cell Membranes and Walls

We have already discussed that the lipids in Archaeal cell membranes contain L (instead of D) glycerol derivatives and that ether links (more stable in reactive environments) replace ester links with isoprenoid (sometimes branched) chains, replacing fatty acid chains. The cell wall is also quite different, and some don't have one. The type of cell wall depends on the environmental need for stability. They don't contain peptidoglycans. Figure 9.5.24 shows four different types.

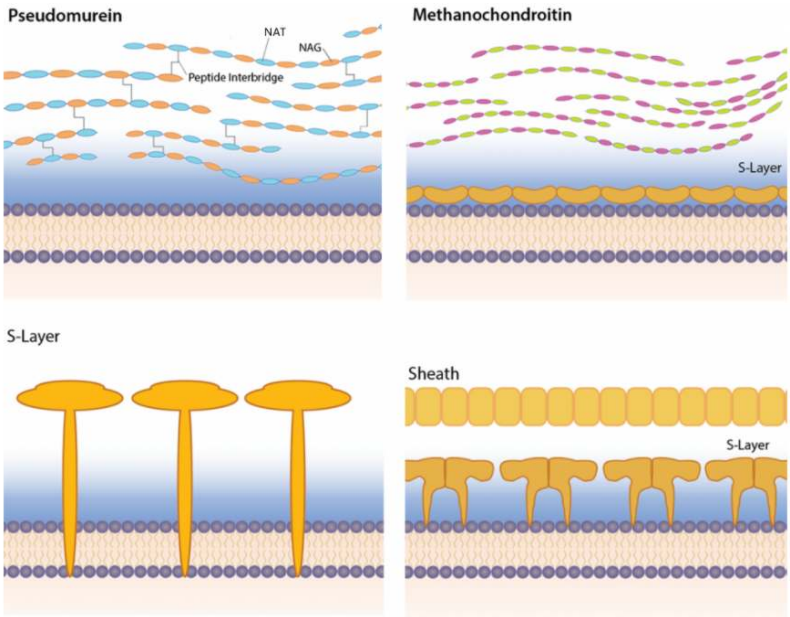


Figure 9.5.24:Archaeal cell walls https://open.oregonstate.education/g...apter/archaea/. https://creativecommons.org/licenses/by-nc/4.0/

Some differences include the presence of

- pseudomurein - This is the closest to the peptidoglycans presented above. Instead of repeating disaccharide units of $(NAM\text{-}NAG)_n$, they have a repeating disaccharide unit of N-acetylalosaminuronic acid (NAT)-NAG. The structure of NAT is shown in Figure 9.5.25.

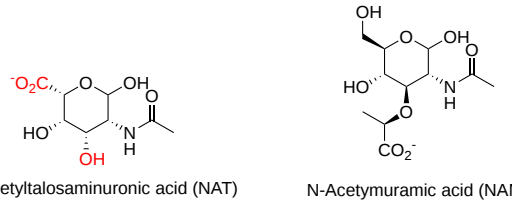


Figure 9.5.25: N-acetylalosaminuronic acid (NAT)compared to N-acetylmuramic acid (NAM)

- methanochondroitin - This is similar to the glycosaminoglycan chondroitin sulfate
- S-Layer
- Sheath/S-Layer

9.5.4.1.4: d. Plant Cell Wall

If you thought bacterial cell walls were complicated, wait until you see plant cell walls! There are about 35 different types of plant cells, and each may have a different cell wall depending on the local needs of a given cell. Cells synthesize thin cell walls that remain thin as the cell grows.

Figure 9.5.26 shows the primary cell wall of plants. The primary cell wall contains cellulose microfibrils (no surprise) and two other polymers, pectin and hemicellulose. The middle lamella, consisting of pectins, is somewhat analogous to the extracellular matrix discussed above.

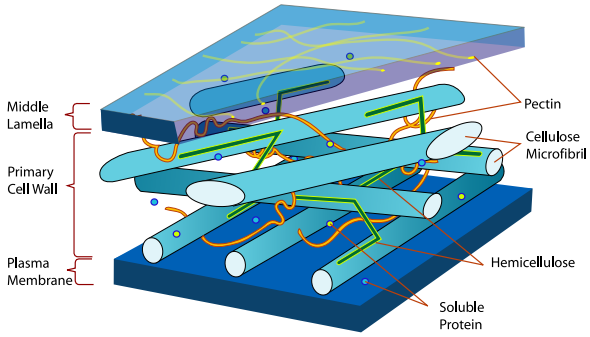


Figure 9.5.26: Primary cell wall of plants. https://commons.wikimedia.org/wiki/F...diagram-en.svg

After cell growth, the cell often synthesizes a secondary cell wall thicker than the first for extra rigidity. The enzymatic machinery for its synthesis is in the cytoplasm and the cell membrane. It is deposited between the cell membrane and the primary cell wall, as shown in the animated image in Figure 9.5.27.

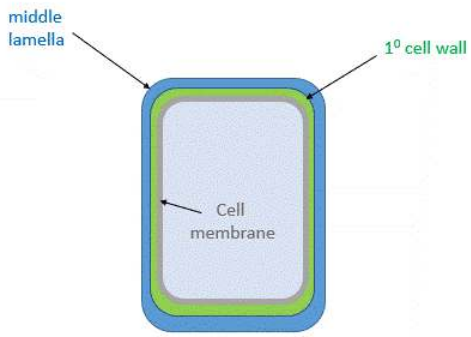


Figure 9.5.27: Primary and secondary cell wall of plants

Figure 9.5.28 shows a structural representation of both the primary and secondary cell walls.

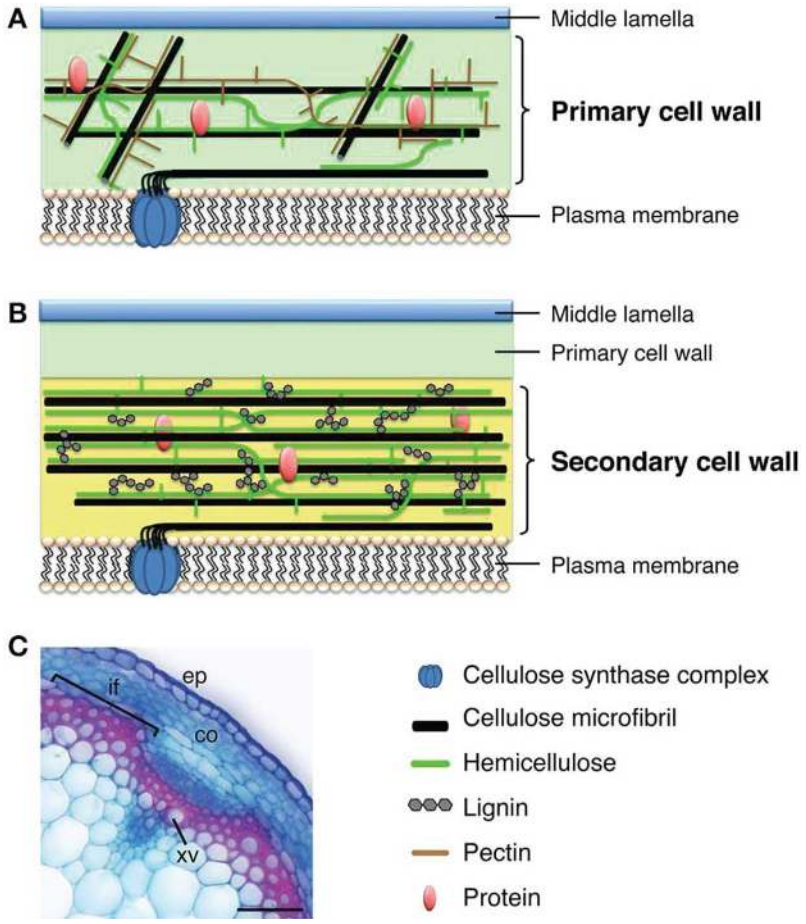


Figure 9.5.28: A structural representation of both the primary and secondary cell walls of plants. Nakano Yoshimi et al. Frontiers in Plant Science (6), 288 (2015) https://www.frontiersin.org/article/...015.00288/full. Creative Commons Attribution License (CC BY).

The middle lamella, which contains pectins, lignins and some proteins, helps "glue together" the primary cell walls of surrounding plants.

Primary Cell Wall:

The main component of the primary plant wall is the homopolymer cellulose (40% -60% mass) in which the glucose monomers are linked β(1→4)-linked into strands that collect into microfibrils through hydrogen bond interactions. Two other groups of polymers, **hemicellulose** and **pectin,** make up the plant cell wall.

Hemicellulose can make up to 20-40% of the mass. These polymers have β(1,4) backbones of glucose, mannose, or xylose (called xyloglucans, xylans, mannans, galactomannans, glucomannans, and galactoglucomannans along with some β(1,3 and 1,4)-glucans. The most abundant hemicellulose in higher plants are the xyloglucans with a cellulose backbone linked at O6 to α-D-xylose. **Pectin** consists of linked galacturonic acids forming homogalacturonans, rhamnogalacturonans, and rhamnogalacturonans II (RGII) [12] [13]. Homogalacturonans (α1→4) linked D-GalA, making up more than 50% of the pectin

Figure 9.5.29 shows some variants of the cell wall components of a plant.

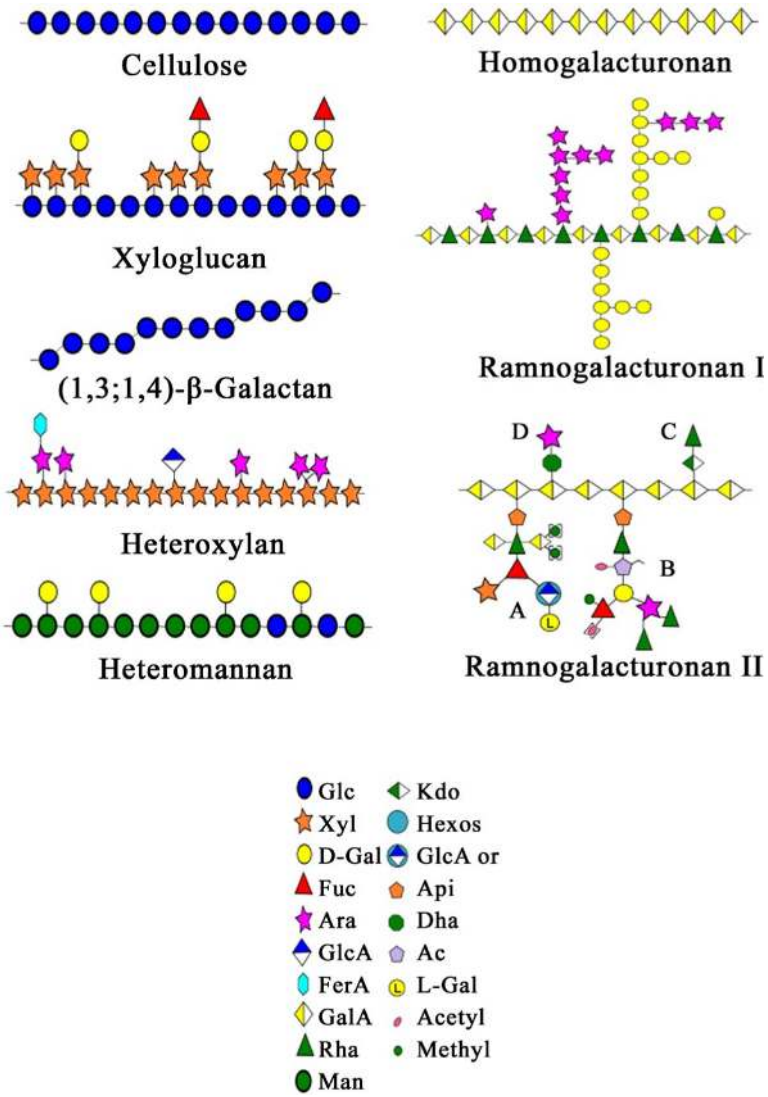


Figure 9.5.29: Variant of the cell wall components of a plant. Costa and Plazanet. Advances in Biological Chemistry 06(03):70-105. DOI: 10.4236/abc.2016.63008. License CC BY 4.0

Secondary Cell Wall

The secondary cell wall's structure depends on the cell's function and environment. It contains cellulose fibers, hemicellulose, and, in addition, a new polymer, **lignin**. It is abundant in xylem vessels and fiber cells of woody plants. It gives the plant extra stability and new functions, including the transport of fluids within the plant through channels.

Lignins, which can make up to 25% of the biomass weight, are made from phenylalanine derivatives but more directly from cinnamic acid. This derivative is made from phenylalanine, which is hydroxylated and converted through other steps to hydroxycinnamyl alcohols called monolignols, as shown in Figure 9.5.30. Three common monomeric (M) derivatives, p-coumaryl, coniferyl, and sinapyl alcohols, can polymerize into lignins, with the units in the polymer (P) named hydroxyphenyl, guaiacyl, and syringyl, respectively.

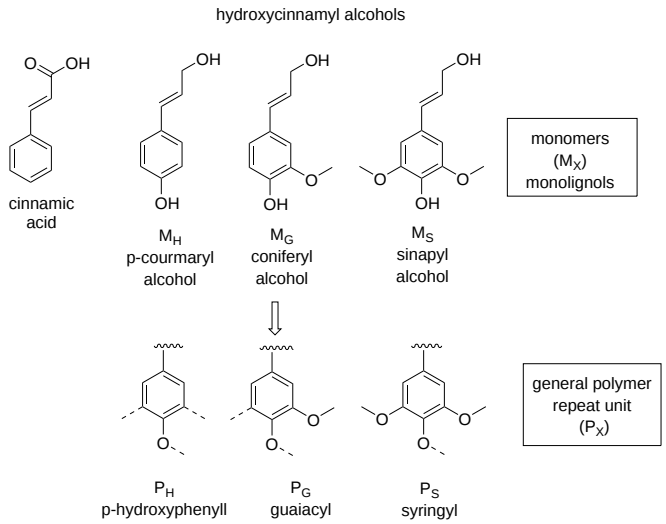


Figure 9.5.30: Monolignols and their polymers

Lignols are activated phenolic compounds that form phenoxide free radicals (catalyzed by peroxidase enzymes), which can attack other lignols to form covalent dimers. Reaction mechanisms for dimerizing the M_S sinapyl alcohol free radical are shown as an example in Figure 9.5.31.

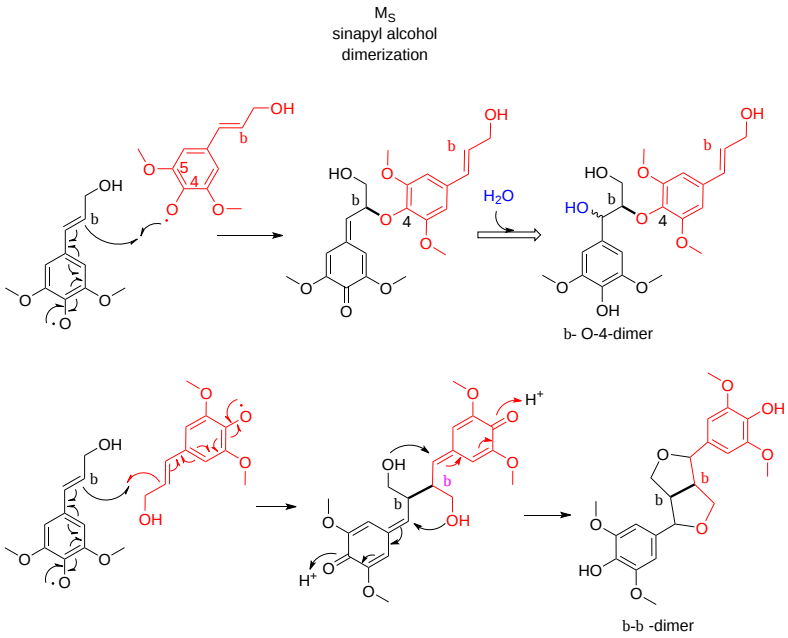


Figure 9.5.31: Dimers of lignols

Now imagine this polymerization continuing by forming additional phenolic free radicals and coupling at many sites to form a huge covalent lignin polymer. Figure 9.5.32 shows one example of a larger lignin.

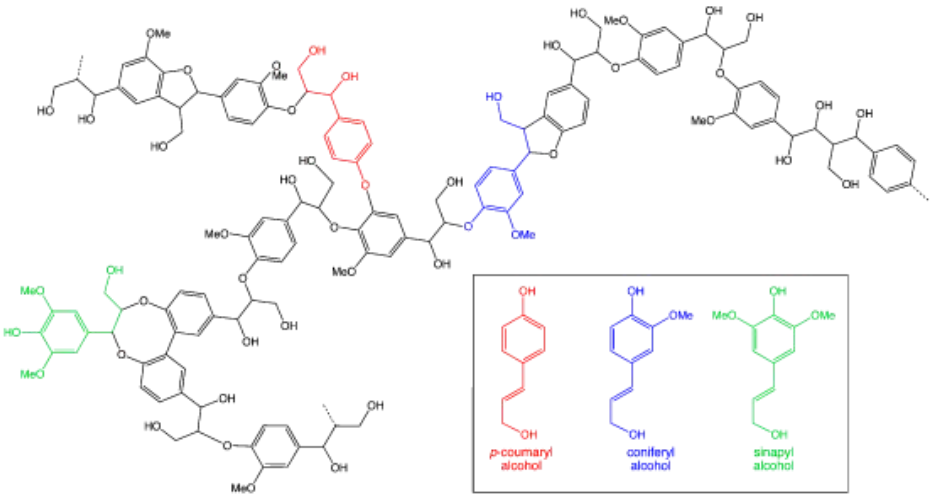


Figure 9.5.32: A larger lignin. https://commons.wikimedia.org/wiki/C...ile:Lignin.png . By Smokefoot - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16022799

Finally, Figure 9.5.33 shows an image of a poplar tree cell wall, made using surface Raman scattering, showing lignin, cellulose, and lipids in secondary xylem cell walls.

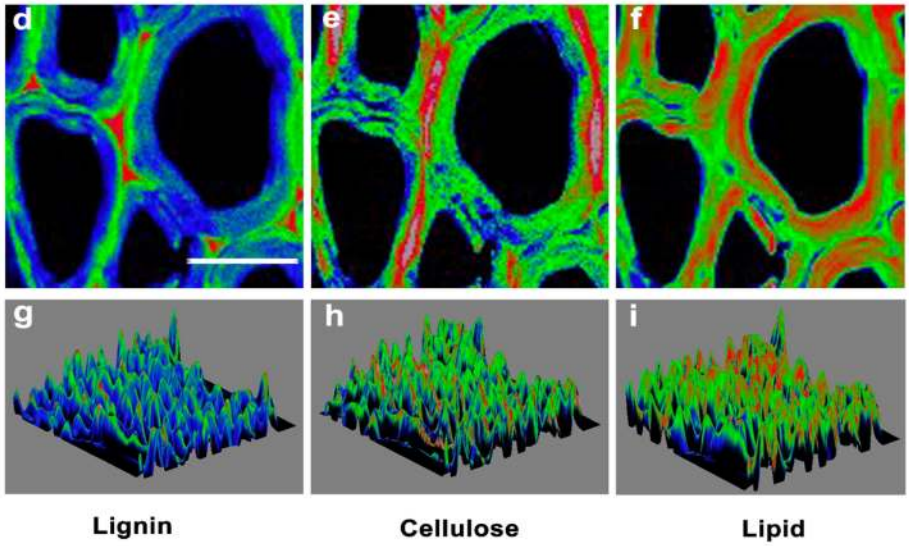


Figure 9.5.33: Surface Raman scattering (SRS) images of lignin, cellulose, and lipids in xylem cell walls. **d-e are** SRS images of lignin, cellulose, and lipids in the secondary xylem cell walls of poplar, respectively. The 3D surface plots are shown in **g–i**. Xu, H., *et al*. A label-free, fast and high-specificity technique for plant cell wall imaging and composition analysis. *Plant Methods* **17,** 29 (2021). https://doi.org/10.1186/s13007-021-00730-9. http://creativecommons.org/licenses/by/4.0/.

9.5.5: The Extracellular Matrix (ECM) and Basement Membranes

We won't formally discuss cell membranes until Chapter 11, but since anyone reading this book has previously seen biological membranes (including the Gram-negative and positive bilayers discussed above), let's explore a term that most chemistry students, but perhaps not biology students, will find very confusing. That topic is the basement membrane. The basement membrane is encountered so often that we will explore its overall structure here, even though it is not a lipid bilayer. It fits well here since it is a complex structure consisting of proteins and proteoglycans. It's very amorphous, making its structure difficult for those hoping for crystal structures or complex bilayers. It is somewhat similar to the cell wall in functionality. We will offer a cursory explanation. Please visit Introduction to Extracellular Matrix and Cell Adhesion in BioLibre texts for a great overall introduction. Some of the images (when noted) below come from that Cell Biology book chapter.

The extracellular matrix (ECM) is a general term for the large protein and polysaccharide network formed on secretion by some cells in a multicellular organism. They act as connective material to hold cells in a defined space. Cell density can vary greatly

between different tissues of an animal, from tightly-packed muscle cells with many direct cell-to-cell contacts to liver tissue, in which some of the cells are only loosely organized, suspended in a web of extracellular matrix, shown in Figure 9.5.34.

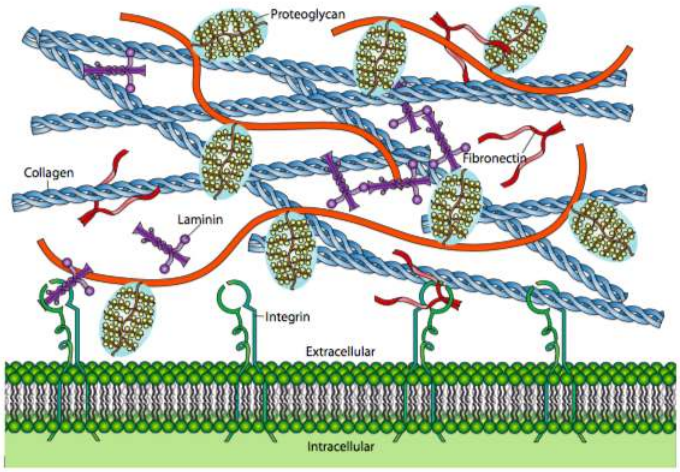


Figure 9.5.34: Extracellular matrix (ECM). Typical components include collagen, proteoglycans (with hydration shells depicted around sugars), bronectin, and laminin. The cellular receptors for several of these ECM components are integrins, although the exact integrin αβ *pair may differ.*

The ECM is a generic term encompassing mixtures of polysaccharides and proteins, including collagens, bronectins, laminins, and proteoglycans, all secreted by the cell. The proportions of these components can vary greatly depending on tissue type. Two quite different examples of ECM are the basement membrane underlying the epidermis of the skin, a thin, almost two-dimensional layer that helps to organize the skin cells into a nearly impenetrable barrier to most simple biological insults, and the massive three-dimensional matrix surrounding each chondrocyte in cartilaginous tissue. The ability of the cartilage in your knee to withstand the repeated shock of your footsteps is due to the ECM proteins in which the cells are embedded, not to the actual cells that are, instead, few and sparsely distributed. Although both types of ECM share some components in common, they are distinguishable not just in function or appearance but in the proportions and identity of the constituent molecules

Figure 9.5.35 shows a general structure of the basement membrane. Think of it as an amorphous polymer mixture (somewhat similar to a polyacrylamide gel).

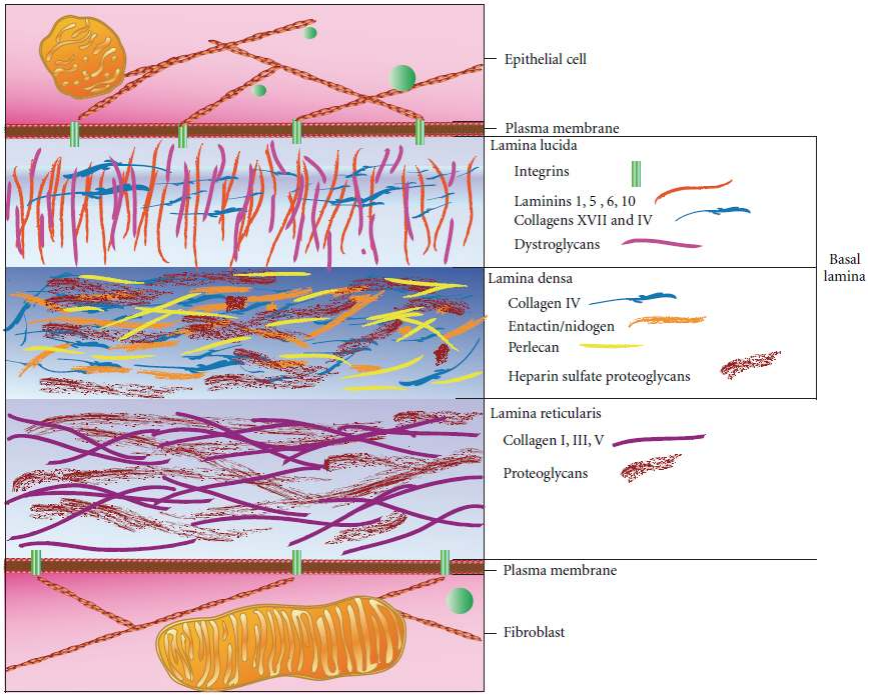


Figure 9.5.35: General structure of the basement membrane. Mentor and DuBois. Journal of Cell Biology · February 2012. DOI: 10.1155/2012/723419 · Source: PubMedCreative Commons License. https://creativecommons.org/licenses/by/3.0/

9.5.6: Summary

This chapter explores the multifaceted world of protein glycosylation, highlighting the complex nature and functional significance of carbohydrates attached to proteins. It emphasizes how post-translational modifications by glycans greatly expand the informational content and functional diversity of the proteome, far beyond what is encoded in the genome.

Key Themes:

1. **Glycosylation and Its Biological Importance:**
   The chapter begins by explaining that many proteins destined for secretion or membrane insertion undergo glycosylation. This modification, typically attached to asparagine (N-linked) or serine/threonine (O-linked) side chains, plays critical roles in:
   - Mediating recognition and binding interactions.
   - Preventing aggregation during protein folding.
   - Protecting proteins from proteolysis.
   - Increasing the half-life of proteins in circulation.
2. **N-linked Glycoproteins:**
   A central focus is placed on N-linked glycosylation, where a core oligosaccharide (Man)3(GlcNAc)2(\text{Man})_3(\text{GlcNAc})_2 is covalently attached to asparagine in a specific consensus sequence. Variants of N-linked glycans include:
   - **High Mannose:** Retaining terminal mannose residues.
   - **Complex:** Featuring GlcNAc branches and terminal sialic acid residues.
   - **Hybrid:** Displaying characteristics of both high mannose and complex types. The chapter uses SNFG (Symbolic Nomenclature For Glycans) to depict these structures and discusses their impact on protein function and immune recognition—as seen, for example, on the heavily glycosylated surface of the SARS-CoV-2 spike protein.
3. **O-linked Glycoproteins and Blood Group Antigens:**
   O-linked glycosylation, which attaches sugars to serine or threonine residues (often starting with a Gal(β1,3)GalNAc moiety), is

examined next. This section explains how such modifications give rise to blood group antigens and affect cell–cell recognition and immune responses.

4. **Proteoglycans and the Extracellular Matrix:**
   The chapter outlines how extensive glycosylation can result in proteoglycans—proteins heavily modified with glycosaminoglycan (GAG) chains. Proteoglycans are major components of the extracellular matrix (ECM) and basement membranes, where they contribute to structural integrity, cell adhesion, and signal transduction.
5. **Carbohydrate Structures in Cell Walls and Membranes:**
   The discussion is broadened to cover the roles of carbohydrates in cell wall structures across different organisms:
   - **Bacterial Cell Walls:** Differences between Gram-positive and Gram-negative bacteria are highlighted, including the structure of peptidoglycan, the role of teichoic acids, and the unique organization of the outer membrane in Gram-negative species.
   - **Plant Cell Walls:** The synthesis and structure of primary and secondary cell walls, with components such as cellulose, hemicellulose, pectin, and lignin, are described, underscoring their importance in plant rigidity and growth.
   - **Archaeal Cell Walls:** Distinct features of archaeal membranes and cell walls are noted, emphasizing the unique lipid composition and structural adaptations in extreme environments.
6. **Clinical Relevance and Glycan Diversity:**
   The chapter concludes by discussing how variations in glycan structures, such as the presence of Gal(α1,3)Gal, have profound implications for immune recognition and xenotransplantation. It also touches on the role of glycosylation in pathogen-host interactions and cell signaling.

Overall, this chapter provides a comprehensive overview of glycosylation as a key post-translational modification, illustrating how complex carbohydrate structures enrich protein function and influence a wide range of biological processes—from cellular communication to structural support.

This page titled 9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **7.3: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

9.6: The Sugar Code and Lectin Decoding

Learning Goals (ChatGPT o1, 1/30/25)

- **Foundational Concepts in Glycosylation and Glycan Diversity:**
  - Define key terms (sugar, carbohydrate, glycan) and explain why glycan structures are extraordinarily diverse compared to proteins and nucleic acids.
  - Discuss the absence of a genetic template for glycan synthesis and how this contributes to the structural heterogeneity of glycans.
- **N-linked Glycoprotein Structure and Variation:**
  - Describe the core oligosaccharide (Man)3(GlcNAc)2(\text{Man})_3(\text{GlcNAc})_2 attached to asparagine in a consensus sequence and explain the differences among high-mannose, complex, and hybrid N-linked glycans.
  - Interpret glycan structures using the Symbolic Nomenclature for Glycans (SNFG) and explain how variations in branching and terminal modifications (e.g., sialic acid, additional galactose) affect function and immunogenicity.
- **O-linked Glycoproteins and Their Roles:**
  - Explain how O-linked glycans are attached to serine or threonine residues and discuss their biological roles, such as in blood group antigens and cell–cell recognition.
- **Proteoglycans and the Extracellular Matrix (ECM):**
  - Define proteoglycans and describe their structure (core protein with covalently attached glycosaminoglycans) and their role in forming the ECM.
  - Analyze how the diversity of glycan structures on proteoglycans contributes to the binding, signaling, and structural integrity of the ECM.
- **Carbohydrate Structures in Cell Walls and Membranes:**
  - Compare and contrast the carbohydrate-based cell walls of bacteria (Gram-positive and Gram-negative) and plants, including the roles of peptidoglycan, teichoic acids, and lignin.
  - Describe the unique features of archaeal cell walls and plant cell walls, and explain how glycan structure underlies functions such as rigidity, protection, and signaling.
- **Glycan-Binding Proteins (GBPs) and the Glycan Code:**
  - Explain the concept of the glycan code and describe how glycan-binding proteins (lectins) serve as “readers” of this code.
  - Differentiate among various classes of GBPs, including viral capsid proteins, bacterial adhesins, and animal lectins, and provide examples such as C-type lectins, galectins, and siglecs.
- **Mechanisms of Glycan Recognition:**
  - Analyze how GBPs interact with specific glycan epitopes through both continuous (linear) and discontinuous (conformational) binding sites.
  - Interpret structural data (including interactive 3D models) to identify key domains (e.g., C-type lectin domains, immunoglobulin-like domains) and motifs (e.g., EPN or WND) involved in glycan binding.
- **Biological and Clinical Relevance:**
  - Discuss the role of glycan modifications in mediating immune responses, cell adhesion, and pathogen recognition (e.g., the role of sialylated glycans in influenza virus binding or the recognition of Gal(α1,3)Gal in xenotransplant rejection).
  - Explore how alterations in glycosylation patterns influence processes such as inflammation, cancer progression, and cell signaling in the ECM and basement membranes.

By achieving these goals, students will develop a detailed understanding of the structural and functional diversity of glycans, the mechanisms by which glycan-binding proteins decode these structures, and the critical roles these processes play in cellular

communication, immune responses, and disease.

9.6.1: Introduction

By now, you should be convinced that the structures of glycans are extraordinarily complex and, in many ways, much more complicated than proteins and nucleic acids. Their structural diversity is staggering, given the number of different sugar monomers, stereocenters, linkages, lengths, conformers, dynamic flexibility, and chemical modifications. Yet evolution has allowed this astronomical diversity, which must serve more than just simple functions such as protecting proteins from degradation. Much of the diversity derives from a lack of an equivalent genetic code for glycan synthesis.

Since all events in biology start with a binding interaction, let's ponder the binding interaction of glycans with partner "ligands" such as proteins, lipids, nucleic acids, etc. A binding site on a glycan could be a single monosaccharide or a much larger and more complex interface. Figure 9.6.1 shows an interactive iCn3D model of one of the few glycoproteins with pdb coordinates, the unliganded simian immunodeficiency virus (SIV) gp120 core glycoprotein (3fus).

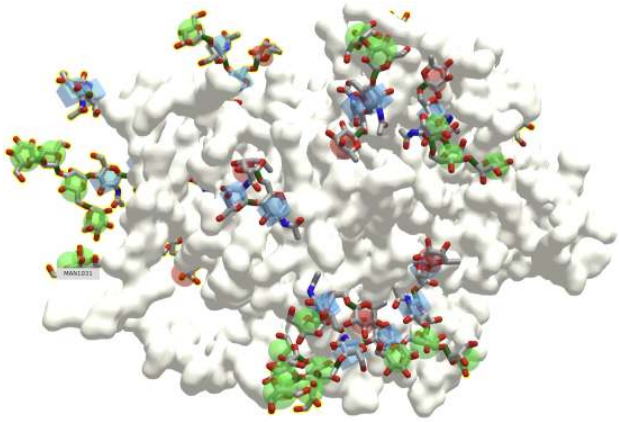


Figure 9.6.1: Simian immunodeficiency virus (SIV) gp120 core glycoprotein (3fus). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?DzZYhLp8JL1vPYYw6

The protein surface is shown in ivory, and the glycans are shown in color sticks with the correct symbolic color-coded spheres or cubes around them.

Now let's convert in our imagination an image file showing one face of the protein to a black and white QR code as shown in Figure 9.6.2.

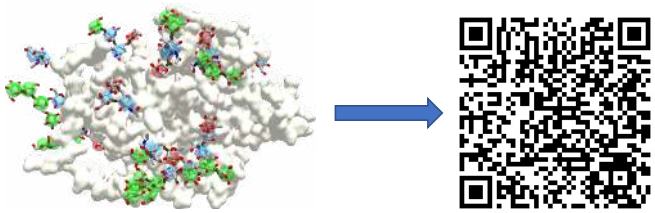


Figure 9.6.2: An imaginary QR code for the surface of a glycoprotein.

Computers can recognize information encoded in the QR codes and decode it into another form of information, such as a menu at a restaurant. Likewise, organisms have evolved "readers" to decode the **glycan code** written by enzymes (glycan synthases, hydrolases, and modifying enzymes). The glycan code is written onto the 3D surfaces of polysaccharides, glycoproteins, glycolipids, and proteoglycans. It should be no surprise that the biological readers of the glycan code are mostly proteins, which locate and bind to the correct "QR" code displayed on the glycan surface.

Luckily, the QR code metaphor for the glycan code is a bit exaggerated since the readers of the glycan code, glycan-binding proteins, seem to recognize just small sections of a glycan. They can be compared to protein antibodies that bind to foreign molecules, such as proteins. The binding site on a foreign protein recognized by an antibody is called an **epitope**. Epitopes can be continuous (linear) stretches of the foreign protein sequence or discontinuous (conformational), made of some continuous stretches of amino acid and some further away in the sequence but close in the 3D folded protein. The average continuous epitope is often 5-

6 amino acids long. Yet that might be an underestimation since an analysis of all contact residues (within a conservative 4 A° distance) for target proteins and their bound antibodies found in the Protein Data Bank is around 18-19 amino acids each (Stave and Lindpaintner). Glycan-binding proteins presumably also bind a mixture of continuous and discontinuous glycan sequences. Linear one would be much easier to determine and study.

Now, let's explore the family of these glycan-binding proteins (GBP), the readers of the glycan code.

9.6.2: Glycan-Binding Proteins (GBPs)

There appear to be nine types of glycan-binding proteins (GBPs). These include nonenveloped capsid virus GBPs,enveloped-virus GBPs (ex., influenza and coronaviruses), eukaryotic microbial GBPs (ex, yeast), and bacterial toxin GBPs (ex, botulinum toxin). Bacterial **adhesins** (parts of organelles like flagella), **lectins** (soluble proteins), and lectin domain-containing proteins are also examples of **glycan-binding proteins (GBP).** We will discuss in more detail three other types: **C-type lectins, galectins, and siglecs**.

In the broadest sense, if a **lectin** is a protein that binds a specific carbohydrate motif (i.e., a glycan code) without modifying the motif, then any glycan-binding protein could be called a lectin. This excludes enzymes that synthesize, degrade, or modify glycans and antibodies that recognize foreign or self-glycan sequences. Table **9.6.1** below shows some lectins and their target glycan ligand from plants, animals, viruses, and bacteria.

Lectin Family/Lectin	Abbreviation	Ligand(s)
	Plants	
Concanavalin A	ConA	Man α1- OCH3
Griffonia simplicifolia lectin 4	GS4	Lewis b (Leb) tetrasaccharide
Wheat germ agglutinin	WGA	Ner5Ac(α 2,3)Gal(β 1,4)GlcGlcNAc(β1,4)GlcNAc
Ricin		Gal(β 1,4)Glc
	Animals	
Galectin-1		Gal(β 1,4)Glc
Mannose-binding protein	MBP-A	High Mannose Octasaccahride
	Viral	
Influenza Virus hemagglutinin	HA	Neu5Ac(α 2,6)Gal(β1,4)Glc
Polyoma virus protein 1	VP1	Neu5Ac(α 2,3)Gal(β1,4)Glc
	Bacterial	
Enterotoxin	LT	Gal
Cholera toxin	CT	GM1 pentasaccharide

Table **9.6.1**: Lectin families and their ligands.

In animals, lectins facilitate cell-cell interactions by forming multiple but weak interactions (also called **multivalent interactions**) between the protein and many sugars on the ligand to which it binds.

Now, let's consider the other three classes of glycan-binding proteins (or lectins), **C-type lectins, galectins,** and **siglecs,** in more detail. Focus on the very different structures of the carbohydrate-binding domains.

9.6.3: C-Type Lectins

C-type lectins comprise the largest number of glycan-binding proteins. These proteins have a glycan or carbohydrate recognition domain that depends on the Ca^{2+} ion. They bind self-glycans and those on pathogens, which can target viruses to specific cells. Many are on the surface of immune cells. They have an N-terminal glycan-binding domain called a C-lectin (CLECT) or carbohydrate recognition domain (CRD). However, some proteins with the domain do not appear to bind either Ca^{2+} or glycans. They serve as adhesion molecules and are also involved in cell signaling. Some residues in the lectin-binding domain appear critical for binding lectins. These include an EPN motif, which interacts with Man, GlcNAc, Fuc, and Glc, and a WND motif involved in binding to Gal and GalNAc.

Let's look now at one example of a C-type lectin, the **selectins**.

P-Selectins

These are involved in the interaction of immune cells in the blood with endothelial cells that line the blood vessel wall. Consider the challenges an immune cell faces as it moves from the blood into a tissue where an infection might occur! Blood flow in vessels is at a rate inversely proportional to the total cross-sectional area of the blood vessel. That rate is about 5-20 cm/sec in arteries, 1.5-7 cm/sec in veins, and about 1 mm/sec (1000 µm)/sec) in capillaries. Assuming a lymphocyte's average diameter of 10 µm, the cell would move about 100 cell lengths per second. An equivalent speed for a human with an arm span of 6 feet (approximately fitting into a circle of diameter = 6 feet as drawn by Leonardo da Vinci) would be around 600 feet/second. The cell must go from its typical circulating speeds to a stop before moving through the blood vessel wall into tissues. Nature has solved this by providing a way to slow down the moving cell until its final capture. The cells roll along the endothelial cells, making transient low-affinity interactions, which slow it down enough until high-affinity ones effectively stop it (unless it dissociates first).

Also, you wouldn't want immune cells to stop and move into tissue without an infection signaled by mediator molecules. Another problem solved! P-selectins are stored in the intracellular granules of platelets (the source of the name P-selectin) and endothelial cells, so moving immune cells are not spuriously captured without some signal. In the presence of the right chemical signal, endothelial cells and platelets get active, and P-selectin is transported rapidly to the cell surface, where "capture" occurs before the cell can move into the underlying tissue. P-selectin mediates the first transient interactions and subsequent rolling of immune cells on activated platelets and endothelial cells.

Figure **9.6.3** is a video animating the rolling and "capture" of a lymphocyte by endothelial cells. (See the video for the reference.) Note that cancer cells can also move through the endothelial cells of blood vessels in forming metastases.

Figure **9.6.3**: Video animation of a lymphocyte rolling and being captured by endothelial cells lining blood vessel walls

P-selectins, hence, are receptors for molecules on immune cells. They bind Ca^{2+} ions, which helps create an active conformation. Their binding ligands are glycan codes and nearby sections of protein connected to the glycan. The glycan ligand on the surface of a circulating immune cell is the **sialyl-Lewis X (SiaLew^X)** glycan or a derivative of it. One of the immune membrane proteins with SiaLew^X is the **P-selectin glycoprotein ligand 1** (**PSGL-1**, the gene name), also called **SELPLG**. It mediates rapid rolling of leukocytes over vascular surfaces during the initial steps in inflammation through interaction with SELPLG"

P-selectin is a mediator of cell adhesion (to other cells). As such, it could also be classified as an adhesion protein. The three main types of selectins:

1. L-selectins: found on leukocytes ("white" blood cells that are circulating immune cells).
2. P-selectins: found on activated platelets (which can aggregate to form a type of blood clot) and activated endothelial cells. Activation occurs during the inflammatory response, which can lead to the quick movement of pre-formed selectins stored within the cytoplasm to the membrane. In addition, their expression can be induced.
3. E-selectins: found on activated endothelial cells only after the cells have been induced to form them by certain immune hormones called cytokines, released by immune cells during an inflammatory response.

Figure 9.6.4 shows the domain structure of human P-selectin.

Figure 9.6.4: Domain structure of human P-selectin (from https://smart.embl.de/)

It contains an N-terminal C-Lectin (CLECT) domain, which is also called the carbohydrate-recognition domain (CRD) or the C-type lectin domain (CTLD). In addition, it has an epidermal growth factor domain (EGF), nine complement control protein (CCP) domains, and the blue transmembrane domain.

Figure 9.6.5 shows the structure of the SLew^x glycan along with its symbol nomenclature for glycans (SNFG) representation.

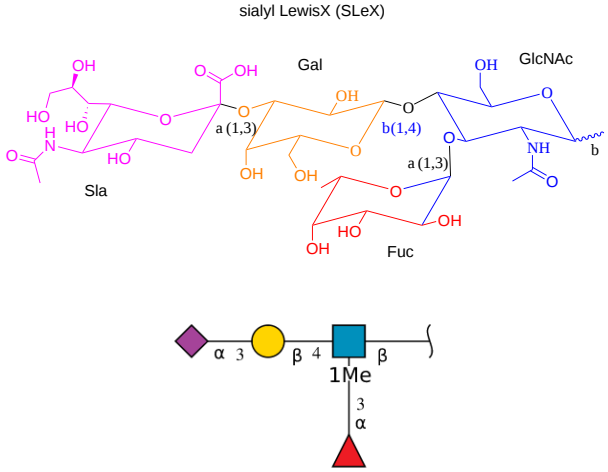


Figure 9.6.5: Structure of the SLew^x glycan

Figure 9.6.6 shows an interactive iCn3D model of the crystal structure of P-selectin lectin/EGF domains complexed with SLeX (1g1r) to which P-selectin binds with weak affinity. Fucose interacts with the Ca^{2+} ion. The glycan interacts with the CLECT domain.

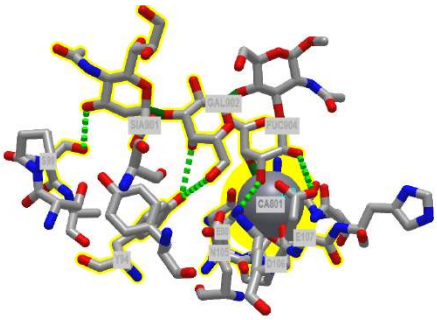


Figure 9.6.6: P-selectin lectin/EGF domains complexed with SLeX (1g1r). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..s1GYbAfXfnuSz9

SLew^x is not present in isolation but rather attached to a membrane protein on an immune cell, which serves as a ligand for the P-selectin on activated endothelial cells or platelets. (The SLew^x can also be part of a glycolipid.) Now let's contrast the interactions of the P-selectin LE domain with "naked" SLew^x with those present between P-selectin LE with a higher affinity natural binding ligand, human P-selectin glycoprotein ligand 1 (PSGL-1), an immune cell integral membrane protein. PSGL-1 is expressed on neutrophils, monocytes, and most lymphocytes. The P-selectin:PSGL-1 complex has a much lower K_D (higher affinity) than the binding of the unmodified SLe^X (1g1s). PSGL-1 is a disulfide-linked homodimer. When sulfated on a specific Tyr (48), the protein displays high affinity for P-selectin. In contrast, when sulfated on a different Tyr (51), it displays a high affinity for L-selectin instead.

The SLex^X type glycan O-linked to the peptide is a bit more complicated than the simple SLex^X ligand, as it is connected to a protein through an O-linked bond at a threonine. The SNFG is shown in Figure 9.6.7.

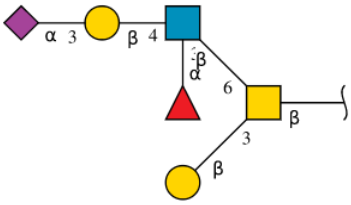


Figure 9.6.7: SNFG if SLex^X type glycan O-linked to a peptide

The crystal structure of a trisulfated, SLew^x-modified peptide from the N-terminal region of PSGL-1 (1G1S) bound to P-selectin lectin and EGF domains (P-LE) has been solved. Figure 9.6.8 shows an interactive iCn3D model that shows some of the interactions between the PSGL-1 peptide (green backbone) and P-LE (magenta backbone).

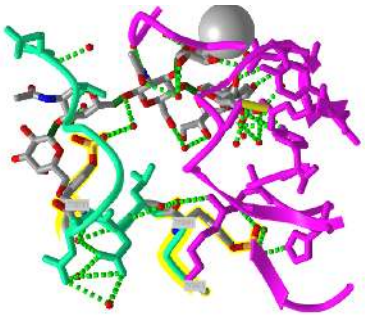


Figure 9.6.8: P-selectin Lectin/EGF domains and bound PSGL-1 peptide (1G1S). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..E2mVULZc8WqVR9

In the crystal structure, the peptide from the P-selectin ligand (which again is a membrane protein) contains three sulfated tyrosine residues (605, 607, and 610), which correspond to amino acids 5, 7, and 10 in the peptide. No electron density for the side chain of Tyr 605 was seen. Tyr 607 binds through multiple interactions to the P-selectin LE domain and is most likely responsible for the high-affinity interaction of P-selectin with the P-selectin glycoprotein ligand (again represented by the green chain). In contrast, Tyr 610 interacts through an intermediary water molecule with the glycan SLex^X of the peptide.

Figure 9.6.9 shows the electrostatic surface potential map of one of the dimers of the P-select LE domains. Blue represents positive potential and red represents negative. The backbone of the P-selectin ligand peptide is green with all of the negatively charged side chains (Tyr, Asp, and Glu) shown in stick with CPK colors. Note that these amino acids are all bound in blue (positive) regions of P-selection. The glycan portion attached to the peptide (stick, CPK color) is positioned mostly over negative potential, allowing hydrogen bonding between the sugar OHs and the protein.

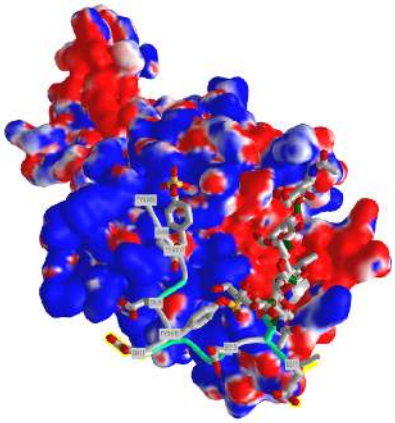


Figure 9.6.9: Electrostatic surface potential map of one monomer of the P-selectin LE domains with bound P-selectin ligand peptide.

You could surmise that the blue region of positive potential could also bind other strongly negatively charged ligands (such as heparin and other glycosaminoglycans), which could inhibit the function of this protein, as it would prevent binding of the PSGL-1.

Figure 9.6.10 shows an interactive iCn3D model, which shows the surface electrostatic potential of the P-selectin Lectin/EGF domains and bound PSGL-1 peptide

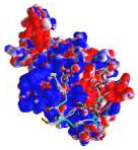


Figure 9.6.10: Electrostatic potential of the P-selectin Lectin/EGF domains and bound PSGL-1 peptide (1G1S). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..X2CSQQzgbFn2VA

The blue represents positive potential, and the red represents negative. The backbone of the P-selectin ligand peptide is green, with all of the negatively charged side chains (Tys, Asp, and Glu) shown in stick with CPK colors.

There are also nonpolar interactions that are not shown in the figure and model above. The aromatic ring of Tyr 607 (7) interacts with the nonpolar parts of a Ser ($-CH_2$) and Lys $(-CH_2)_4$ side chains, and the ring of Tyr 610 (10) interacts with two leucine side chains.

The selectins are also part of a class of molecules called **adhesion** molecules. As described for the selectins, adhesion molecules contain

- an extracellular CHO binding domain (the lectin domain), which mediates binding to adjacent cells or the extracellular matrix;
- a transmembrane domain;
- and a cytoplasmic domain, which often interacts with the cytoskeleton within the cell.

This initial binding mediated by selectin-CHO interactions activates the expression of another adhesion molecule on the leukocyte, **integrin**, a heterodimer with an a and b chain. These cause strong leukocyte-endothelial cell interactions, leading to the movement of the leukocytes through the vessel wall. Other classes of adhesion molecules (in addition to selectins and integrins) are **cadherins** (calcium-dependent adhesion molecules) and the immunoglobulin-like superfamily (ICAM1, ICAM2, VCAM). VCAM (Vascular Adhesion Molecule) binds to integrin expressed on activated lymphocytes, leading to the lymphocyte's passage from the vessel's lumen into the tissues. Integrins appear to bind proteins in the extracellular matrix through RGD (Arg-Gly-Asp) and also through LDV (Leu-Asp-Val) motifs on the

proteins, including fibronectin (RGD), thrombospondin (RGD & LDV), fibrinogen (RGD & LDV), von Willebrand Factor (RGD), and vitronectin (RGD). They also bind other matrix proteins with an "alpha domain," including collagen and laminin. Integrin/Adhesion molecule interactions involve protein/protein interactions.

A fertilized egg (in the blastocyst stage, which is ready for implantation in the uterine cell wall) expresses L-selectin, which allows a low-affinity (rolling-type) interaction of the fertilized egg with the uterine **epithelial** cells. These cells expressed the CHO ligands on their surface, which bind to the L-selectin on the blastocyst. The CHO ligands are only transiently expressed on the surface of the epithelial cells of the uterus, presumably only when the uterus is primed for implantation. After the initial interaction of the blastocyst and epithelial cells, further expression of integrins on the blastocyst surface might result. Problems in any of these molecular steps could result in infertility. Figure **9.6.11** shows endothelial cell/leukocyte interactions mediated through selectins, integrins, and ICAMs.

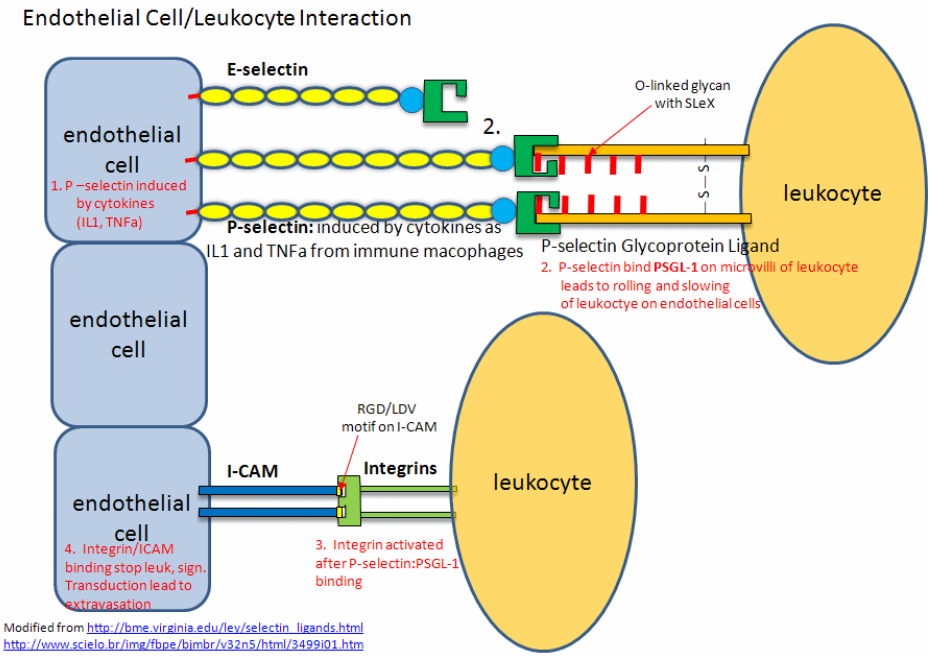


Figure **9.6.11**: Endothelial cell/leukocyte Interactions mediated through selectins, integrins, and ICAMs

Post-translational modifications of protein modification (like glycosylation) can confer new binding and biological functions to a protein. Site-directed mutagenesis can replace surface amino acids with cysteine or methionine with nonnatural amino acid analogs that contain azide or alkyne groups. These modified groups could then direct the location of chemical-modifying reagents (such as sugars) to these sites. A protein completely unrelated to PSGL-1 has been selectively modified to contain covalently attached glycans and sulfated tyrosine side chains—the unrelated protein bound to P-selectin.

9.6.4: Mannose Receptor.

What do you do with a protein that no longer has the correct structure(s) to perform its designed function(s)? Proteins, as with any molecule, undergo chemical changes during their biological lifetime. They must be recognized as aberrant and then removed from "service," ultimately being degraded into component amino acids for reuse. There are no repair enzymes for proteins as there are for DNA. One modification that changes glycoproteins and signals the need for their removal is the removal of the terminal sialic acid residue, forming asialoglycoproteins, whose glycans end in galactose, as you can envision from Figure **9.6.12**, which shows a typical structure of an N-linked glycoprotein.

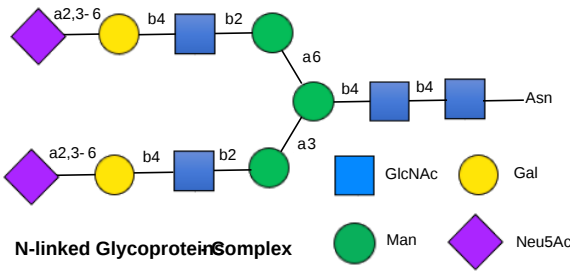


Figure 9.6.12: Typical structure of a N-linked glycoprotein

The **asialoglycoprotein receptor**, a member of the C-type lectin family, is a transmembrane protein that binds terminal galactose and N-acetylgalactosamine sugars on the end of circulating asialoglycoproteins, leading to their endocytosis into the cell. It is expressed on the surface of hepatocytes (liver cells). Receptors of this type are also called scavenger receptors, as they remove proteins from circulation.

The **mannose receptor** (also called CD206), also expressed in liver endothelial cells, is another C-type lectin involved in the binding and removal of glycoproteins from the circulation. It binds both sulfated and non-sulfated glycans. It is also a receptor that allows the binding and phagocytosis of bacterial and fungal pathogens by a type of immune cell, namely macrophages or dendritic cells. Unfortunately, tumor cells can use the same process for uptake into macrophages, promoting tumor cell growth. The protein binds and scavenges sulfated glycoprotein hormones, mannose-bearing glycoproteins released during inflammation, lysosomal enzymes released from cells on injury, and collagen fragments.

Figure 9.6.13 shows the domain structure of the human mannose receptor.

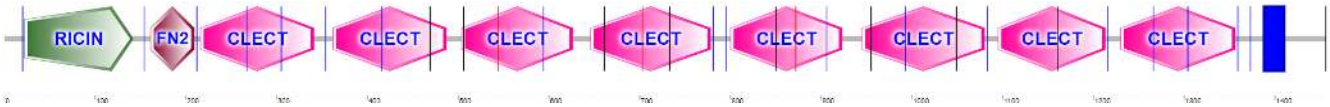


Figure 9.6.13: Domain structure of the human mannose receptor.

Given the large number of CLECT domains, you might surmise that this protein could bind many different target glycans from both self and pathogens. What is different about the domain structure compared to P-selectin is the presence of an N-terminal Ricin and a Fibronectin type 2 (FN2) domain. The FN2 domain has two cystines from the four conserved cysteines involved in disulfide bonds. What's so interesting about the mannose receptor is that it binds glycans both in the CLECT domains and in the FN2 domain.

Glycan binding at the CLECT domain: The CLECT domain binds targets containing mannose, fucose, and N-acetylglucosamine with a preference for Man(α1,2)Man or fucose. Figure 9.6.14 shows an interactive iCn3D model. of the **CLECT** 4 domain of the mannose receptor complexed with Man(α1,2)Man (7jue). Interactions of fucose lin ligands, such as Lewis-a-trisaccharide, strengthen the binding.

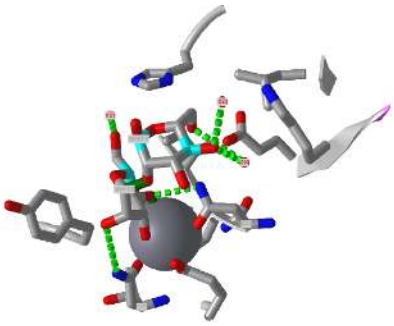


Figure 9.6.14: **CLECT** 4 domain of the mannose receptor complexed with Man(α1,2)Man (7jue). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..jhn5eMLMrG6mX7

The receptor can bind a variety of glycans. Both mannose and N-acetylglucosamine interact with bound Ca^{2+} through equatorial OHs on carbon 3 and 4 of the ring, while fucose uses OHs on carbon 2 and 3, or 3 and 4.

The interactions with fungal pathogens are medically important. Fungi, like yeast, have an outer structure composed of a membrane bilayer and a mixture of glycans, which deploys an incredibly complex "glycan code" to host cells infected by them, as illustrated in Figure **9.6.15**.

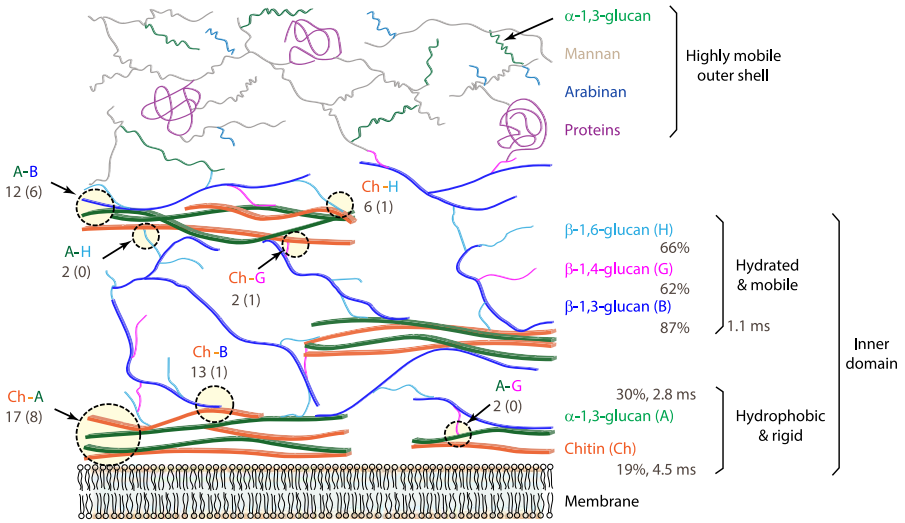


Figure **9.6.15**: Structure of fungal cell wall. Kang, X., Kirui, A., Muszyński, A. *et al*. Molecular architecture of fungal cell walls revealed by solid-state NMR. *Nat Commun* **9,** 2747 (2018). https://doi.org/10.1038/s41467-018-05199-0. Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0/

Mannans, polymers of just mannose, differ widely in structure. Their main backbone can be DMan(α-1,6)DMan or DMan(β-1,4)DMan with many branches.

Glcyan binding at the FN2 (Cysteine-Rich) Domain (1FWU): The mannose receptor can also bind non-mannose sulfated glycans, such as 3-SO4-LEWIS(X), for which the SNFG representation is shown in Figure **9.6.16**.

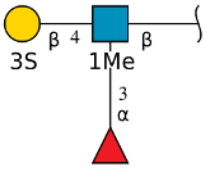


Figure **9.6.16**: 3-SO4-LEWIS(X) non-mannose sulfated glycans for mannose receptor

The mannose receptor binds this glycan, which does not even contain mannose, through the FN2 domain (which contains four disulfide bonds) and not the CLECT calcium-dependent carbohydrate-binding domain. Hence, the protein can bind both sulfated and nonsulfated glycans.

Figure **9.6.17** shows an interactive iCn3D model of the complex of the FN2 domain of the mannose receptor with the non-mannose containing 3-SO4-LEWIS(X) glycan (1fwu).

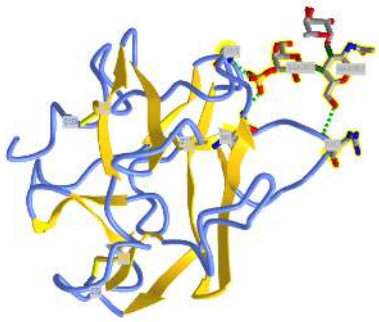


Figure 9.6.17: FN2 domain of the mannose receptor with the non-mannose containing 3-SO4-LEWIS(X) glycan (1fwu). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..kXn2Rz8qMzebo9

Look at the number of CLECT domains in the domain structure diagram for the mannose receptor above. Along with interactions of sulfated glycans at the FN2 domain, these would enable the binding of widely diverse glycan structures. Reported ligands for the mannose receptor include those with high mannose content released during inflammation (lysosomal hydrolases, collagen peptides, and tissue plasminogen activator), and sulfated ones (including the pituitary hormones lutropin and thyrotropin.

9.6.5: Galectins

This family of glycan-binding proteins contains a common carbohydrate recognition domain (CRD) of about 130 amino acids, which binds Galβ1,3GlcNAc or Galβ1,4GlcNAc disaccharides (hence the name galectins) as well as other glycan motifs. They are expressed in almost all cells and multicellular organisms. There are 15 different types grouped in how the CRD is functionally expressed (as dimers, tandem repeats, or chimeras), as illustrated in Figure 9.6.18. The figure also shows their role in cancer biology.

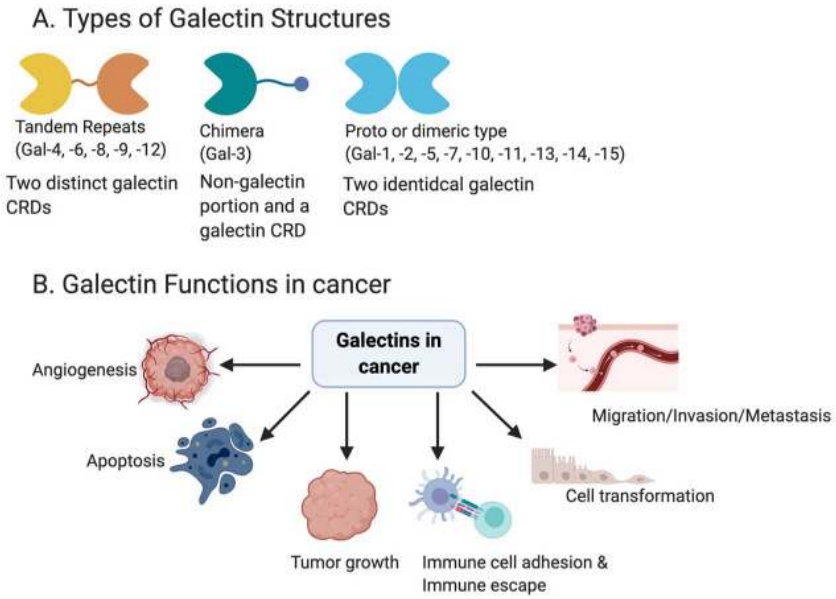


Figure 9.6.18: Galnectins structure and function Shimada, C.; Xu, R.; Al-Alem, L.; Stasenko, M.; Spriggs, D.R.; Rueda, B.R. Galectins and Ovarian Cancer. *Cancers* **2020**, *12*, 1421. https://doi.org/10.3390/cancers12061421. **Creative Commons Attribution License**

The carbohydrate-binding domain of the galactins has a jellyroll-like protein architecture with two anti-parallel β-sheets forming a β-sandwich.

Galectin I

This protein is secreted and is found in the extracellular matrix, as well as in the cytoplasm. It induces apoptosis in T-cells. It binds beta-galactosides as well as other glycans. The main ligand of galectin-1 has a Galβ1-4GlcNAc (or LacNAc) structure. Figure

9.6.19 shows an interactive iCn3D model of Human Galectin-1 in Complex with Type 1 N-acetyllactosamine (Gal(β1,3)GlcNAc), which binds less tightly than Galβ1,4GlcNAc (Type 2)(4XBL)

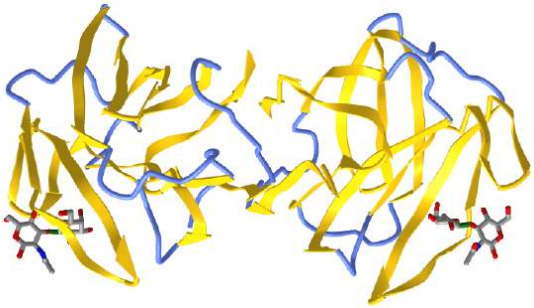


Figure 9.6.19: Human Galectin-1 in Complex with Type 1 N-acetyllactosamine (Gal(β1,3)GlcNAc) (4XBL). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..MTyV7QvPKqqibA

A comparison of the crystal structures shows different phi/psi angles for bound Type I (135°) versus the more tightly bound Type 2 (-108°), which shows the nuance in binding conformations in the interactions of glycans with glycan-binding proteins.

9.6.6: Siglecs

The proteins are sialic acid-binding immunoglobulin (Ig)-like lectins found on immune cells like basophils, macrophages, mast cells, and eosinophils. One type (Siglec-4)is found in myelinated structures in the central and peripheral nervous systems. They all have an N-terminal extracellular immunoglobulin domain (abbreviated as IG or V-Set) and a differing number of IG-like domains, also called C2-set Ig domains. The glycan binding epitope recognized by Siglecs is sialylated oligosaccharides on a protein section containing a conserved arginine. Figure 9.6.20 compares the domain structures of the human Siglec family.

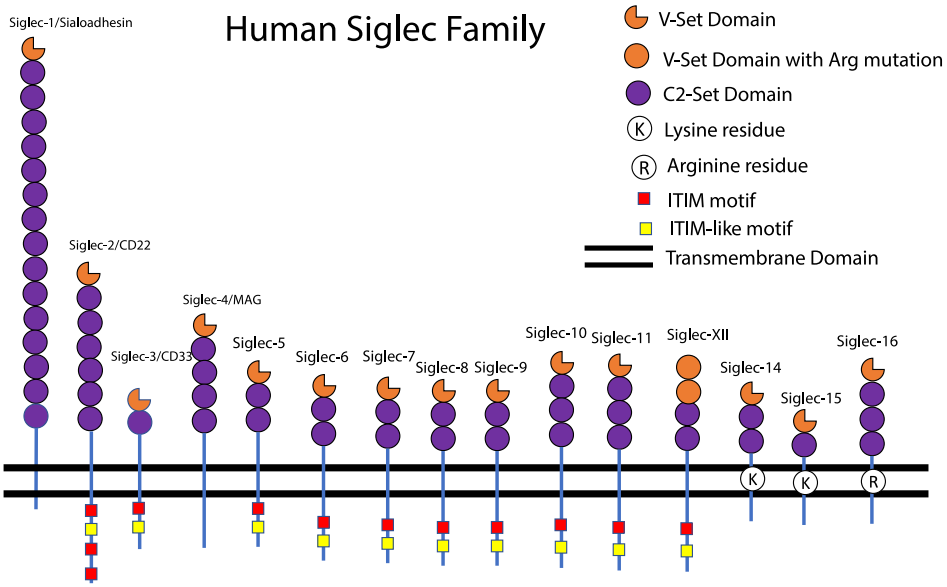


Figure 9.6.20: Domain structures of the human Siglec family Siddiqui, S.S.; Matar, R.; Merheb, M.; Hodeify, R.; Vazhappilly, C.G.; Marton, J.; Shamsuddin, S.A.; Al Zouabi, H. Siglecs in Brain Function and Neurological Disorders. *Cells* **2019**, *8*, 1125. https://doi.org/10.3390/cells8101125. Open access article distributed under the Creative Commons Attribution License

Here is one example of a Siglec.

Siglec-8

This protein is expressed on the surface of immune cells like basophils, mast cells, and eosinophils. When activated by infection and prolonged inflammation, they release the contents of intracellular granules, which have potent physiological effects that can lead to allergic and asthmatic responses. In infection and other inflammatory states, immune cytokines are released that, in a signaling process, lead to the release of sialoglycans that act as ligands, binding to the Siglec-8 on the surface of the immune

cells. One type of sialoglycan released is mucins, which are very large glycoproteins with many 6′S sLe^x glycans attached. These "multivalent" glycan epitopes can bind to Siglec-8, leading to signaling in the cells and ultimate inhibition of cell function (including by death or apoptosis). The mucins in mucus (cross-linked mucins), which cover epithelial cells or airways, also act as a first line of defense as they can bind viruses through multiple-contact (multivalent) binding sites, effectively trapping the viruses. The glycan structure recognized by Siglec-8 is sialic acid and sulfate (NeuAcα2-3[6S]Galβ1-4G[Fucα1-3]GlcNAc-). Given their role in inhibiting and inducing apoptosis in immune cells, the family of siglecs is likely involved as checkpoints, which are important in cancer and inflammatory conditions.

Figure 9.6.21 shows the domain structure of Siglec-8

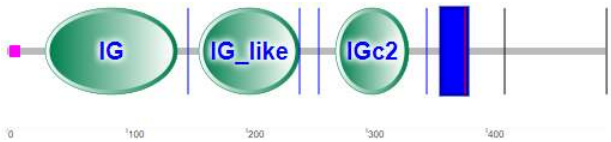


Figure 9.6.20: Domain structure of Siglec-8

Note that there is no CLECT domain, but rather immunoglobulin (IG) or IG-like domains, which seems logical given their role in binding glycan "epitopes". The IG domain is also called the immunoglobulin V-set domain (V-Set). The blue rectangle represents the transmembrane domain (single helix). The cytoplasm contains a tyrosine-inhibitory motif (ITIM) involved in transducing the signal on binding 6′S sLe^x glycans to the IG domains.

As discussed above, humans lack a hydrolase gene necessary for the hydroxylation of Neu5Ac to Neu5Gc, which is found in chimps that possess the enzyme. Chimp's immune systems seem to confer protection from acquiring simian versions of AIDS, cirrhosis, and other diseases which humans acquire when they are infected with the human versions of HIV, hepatitis B or C, or other viruses. These diseases and others associated with overactive T cells (rheumatoid arthritis, asthma, Type 1 diabetes) are uncommon in chimps. It turns out that there is a link between the type of sialic acid and the expression of siglecs that influences the difference in our disease propensity. Varki et al. have shown that chimps and gorillas show much higher levels of expression of siglecs on T cells, which are critical regulatory and effector cells in the immune system. When siglecs on T cells are activated, T-cell responses are down-regulated. Although HIV ultimately kills T helper cells, the virus initially activates them on infection, leading to their proliferation and production of a larger number of cells for the virus to infect.

Figure 9.6.21 shows an interactive iCn3D model of human Siglec-8 lectin domain in complex with 6'sulfo sialyl Lewisx (2N7B)

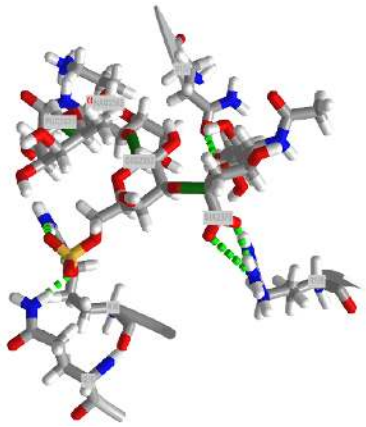


Figure 9.6.19: Human Siglec-8 lectin domain in complex with 6'sulfo sialyl Lewisx (2N7B). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i..Gaa93cu86jMPA9

9.6.7: Summary

This chapter delves into the complex world of protein glycosylation and glycan interactions, highlighting the extraordinary structural diversity and biological significance of carbohydrate modifications. Although the genetic code provides instructions for

protein synthesis, glycan structures are not directly encoded, resulting in a vast array of possible modifications that contribute to the functional complexity of the proteome.

Key topics include:

- **Protein Glycosylation and Its Roles:**
  The chapter introduces how carbohydrates are post-translationally attached to proteins—most commonly via N-linked glycosylation (through Asn in a consensus sequence) and O-linked glycosylation (typically via Ser or Thr). These modifications enhance protein folding, stability, and protection from proteolysis while also mediating critical binding interactions.
- **Diversity in Glycan Structures:**
  N-linked glycans share a common core structure $(Man)_3(GlcNAc)_2$ but diverge into high-mannose, complex, or hybrid types based on further modifications and branching patterns. Similarly, O-linked glycans, exemplified by blood group antigens, show a wide range of structures that affect cell–cell recognition and immune responses.
- **Glycan-Binding Proteins (GBPs):**
  The chapter explains how specific proteins—often termed lectins—serve as "readers" of the glycan code. GBPs, including C-type lectins (like selectins), galectins, and siglecs, bind to defined glycan epitopes and mediate essential functions such as immune cell adhesion, signaling, and pathogen recognition. Structural insights, including interactive 3D models and symbolic nomenclature (SNFG), help illustrate these interactions.
- **Glycans in Cellular Architecture:**
  Beyond glycoproteins, the role of glycans in forming complex extracellular matrices, bacterial cell walls, and plant cell walls is examined. The unique composition and organization of these carbohydrate-rich structures underpin critical functions—from maintaining cell shape and rigidity to modulating cellular interactions and immune responses.
- **Biological and Clinical Implications:**
  The chapter also highlights the clinical importance of glycan structures. For example, variations in terminal glycan epitopes can trigger immune responses (e.g., in xenotransplantation or alpha-gal syndrome) and influence the infectivity of viruses such as influenza and SARS-CoV-2.

Overall, this chapter underscores that glycosylation is a fundamental, yet highly complex, modification that plays vital roles in protein function, cell signaling, and organismal physiology. It prepares the reader to appreciate how the dynamic "glycan code"—decoded by specialized glycan-binding proteins—contributes to both normal biological processes and disease states.

This page titled 9.6: The Sugar Code and Lectin Decoding is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **7.4: The Sugar Code and Lectin Decoding** by Henry Jakubowski and Patricia Flatt has no license indicated.

9.7: Chapter 7 Problems

7.07: Chapter 7 Problems

This page titled 9.7: Chapter 7 Problems is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

9.8: Chapter 7 Problems - Answer Key

7.06: Chapter 7 Problems - Answer Key

This page titled 9.8: Chapter 7 Problems - Answer Key is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

CHAPTER OVERVIEW

10: Nucleic Acids

10.1: Structure and Function - Nucleic Acids
10.2: Nucleic Acids - Structure and Function
10.3: Nucleic Acids - RNA Structure and Function
10.4: Nucleic Acids - Comparison of DNA and RNA
10.5: Enzymes for Genetic modifications
10.6: DNA Isolation, Sequencing, and Synthesis
10.7: Bioinformatics

10: Nucleic Acids is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.1: Structure and Function - Nucleic Acids

Source: BiochemFFA_2_5.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

The nucleic acids, DNA and RNA, may be thought of as the information molecules of the cell. In this section, we will examine the structures of DNA and RNA, and how these structures are related to the functions these molecules perform.

We will begin with DNA, which is the hereditary information in every cell, that is copied and passed on from generation to generation. The race to elucidate the structure of DNA was one of the greatest stories of 20th century science. Discovered in 1869 by Friedrich Miescher, DNA was identified as the genetic material in experiments in the 1940s led by Oswald Avery, Colin MacLeod, and Maclyn McCarty. X-ray diffraction work of Rosalind Franklin and the observations of Erwin Chargaff were combined by James Watson and Francis Crick to form a model of DNA that we are familiar with today. Their famous paper, in the April 25, 1953 issue of Nature, opened the modern era of molecular biology. Arguably, that one-page paper has had more scientific impact per word than any other research article ever published. Today, every high school biology student is familiar with the double helical structure of DNA and knows that G pairs with C and A with T.

The double helix, made up of a pair of DNA strands, has at its core, bases joined by hydrogen bonds to form base pairs - adenine always paired with thymine, and guanine invariably paired with cytosine. Two hydrogen bonds are formed between adenine and thymine, but three hydrogen bonds hold together guanine and cytosine (Figure 2.127).

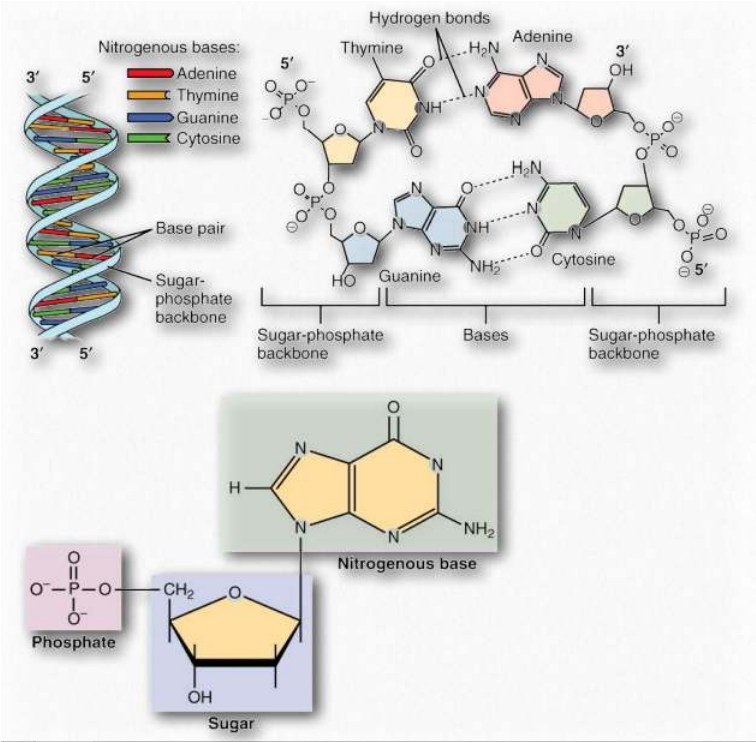


Figure 2.127 - A DNA duplex with base pairs, a closeup of base pairing, and a closeup of a nucleotide Wikipedia

The complementary structure immediately suggested to Watson and Crick how DNA might be (and in fact, is) replicated and it further explains how information is DNA is transmitted to RNA for the synthesis of proteins. In addition to the hydrogen bonds between bases of each strand, the double helix is held together by hydrophobic interactions of the stacked, non-polar bases. Crucially, the sequence of the bases in DNA carry the information for making proteins. Read in groups of three, the sequence of the bases directly specifies the sequence of the amino acids in the encoded protein.

Structure

A DNA strand is a polymer of nucleoside monophosphates held together by phosphodiester bonds. Two such paired strands make up the DNA molecule, which is then twisted into a helix. In the most common Bform, the DNA helix has a repeat of 10.5 base pairs per turn, with sugars and phosphate forming the covalent phosphodiester “backbone” of the molecule and the adenine, guanine, cytosine, and thymine bases oriented in the middle where they form the now familiar base pairs that look like the rungs of a ladder.

Building blocks

The term nucleotide refers to the building blocks of both DNA (deoxyribonucleoside triphosphates, dNTPs) and RNA (ribonucleoside triphosphates, NTPs). In order to discuss this important group of molecules, it is necessary to define some terms.

Nucleotides contain three primary structural components. These are a nitrogenous base, a pentose sugar, and at least one phosphate. Molecules that contain only a sugar and a nitrogenous base (no phosphate) are called nucleosides. The nitrogenous bases found in nucleic acids include adenine and guanine (called purines) and cytosine, uracil, or thymine (called pyrimidines). There are two sugars found in nucleotides - deoxyribose and ribose (Figure 2.128). By convention, the carbons on these sugars are labeled 1’ to 5’. (This is to distinguish the carbons on the sugars from those on the bases, which have their carbons simply labeled as 1, 2, 3, etc.) Deoxyribose differs from ribose at the 2’ position, with ribose having an OH group, where deoxyribose has H.

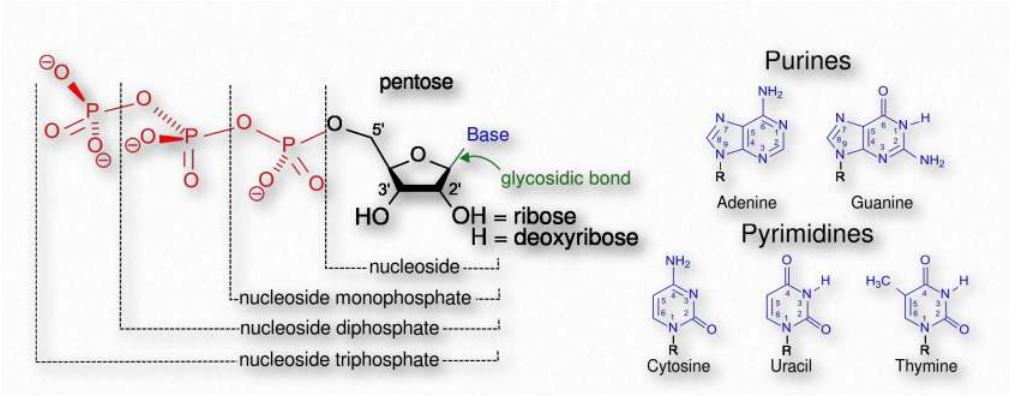


Figure 2.128 - Nucleotides, nucleosides, and bases

Nucleotides containing deoxyribose are called deoxyribonucleotides and are the forms found in DNA. Nucleotides containing ribose are called ribonucleotides and are found in RNA. Both DNA and RNA contain nucleotides with adenine, guanine, and cytosine, but with very minor exceptions, RNA contains uracil nucleotides, whereas DNA contains thymine nucleotides. When a base is attached to a sugar, the product, a nucleoside, gains a new name.

- **uracil-containing** = uridine (attached to ribose) / deoxyuridine (attached to deoxyribose)
- **thymine-containing** = ribothymidine (attached to ribose) / thymidine (attached to deoxyribose)
- **cytosine-containing** = cytidine (attached to ribose - Figure 2.129) / deoxycytidine (attached to deoxyribose)
- **guanine-containing** = guanosine (attached to ribose) / deoxyguanosine (attached to deoxyribose)
- **adenine-containing** = adenosine (attached to ribose) / deoxyadenosine (attached to deoxyribose)

Of these, deoxyuridine and ribothymidine are the least common. The addition of one or more phosphates to a nucleoside makes it a nucleotide. Nucleotides are often referred to as nucleoside phosphates, for this reason. The number of phosphates in the nucleotide is indicated by the appropriate prefixes (mono, di or tri).

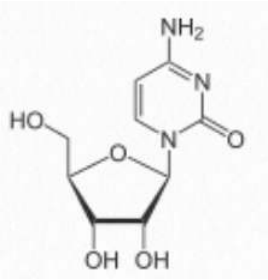


Figure 2.129 Cytidine

Thus, cytidine, for example, refers to a nucleoside (no phosphate), but cytidine monophosphate refers to a nucleotide (with one phosphate). Addition of second and third phosphates to a nucleoside monophosphate requires energy, due to the repulsion of negatively charged phosphates and this chemical energy is the basis of the high energy triphosphate nucleotides (such as ATP) that fuel cells.

Note: Ribonucleotides as Energy Sources

Though ATP is the most common and best known cellular energy source, each of the four ribonucleotides plays important roles in providing energy. GTP, for example, is the energy source for protein synthesis (translation) as well as for a handful of metabolic reactions. A bond between UDP and glucose makes UDP-glucose, the building block for making glycogen. CDP is similarly linked to several different molecular building blocks important for glycerophospholipid synthesis (such as CDP-diacylglycerol).

The bulk of ATP made in cells is not from directly coupled biochemical metabolism, but rather by the combined processes of electron transport and oxidative phosphorylation in mitochondria and/or photophosphorylation that occurs in the chloroplasts of photosynthetic organisms. Triphosphate energy in ATP is transferred to the other nucleosides/nucleotides by action of enzymes called kinases. For example, nucleoside diphosphokinase (NDPK) catalyzes the following reaction

$$\text{ATP} + \text{NDP} \longleftrightarrow \text{ADP} + \text{NTP} \tag{10.1.1}$$

where ‘N’ of “NDP” and “NTP corresponds to any base. Other kinases can put single phosphates onto nucleosides or onto nucleoside monophosphates using energy from ATP.

Deoxyribonucleotides

Individual deoxyribonucleotides are derived from corresponding ribonucleoside diphosphates via catalysis by the enzyme known as ribonucleotide reductase (RNR). The deoxyribonucleoside diphosphates are then converted to the corresponding triphosphates (dNTPs) by the addition of a phosphate group. Synthesis of nucleotides containing thymine is distinct from synthesis of all of the other nucleotides and will be discussed later.

Building DNA strands

Each DNA strand is built from dNTPs by the formation of a phosphodiester bond, catalyzed by DNA polymerase, between the 3’OH of one nucleotide and the 5’ phosphate of the next. The result of this directional growth of the strand is that the one end of the strand has a free 5’ phosphate and the other a free 3’ hydroxyl group (Figure 2.130). These are designated as the 5’ and 3’ ends of the strand.

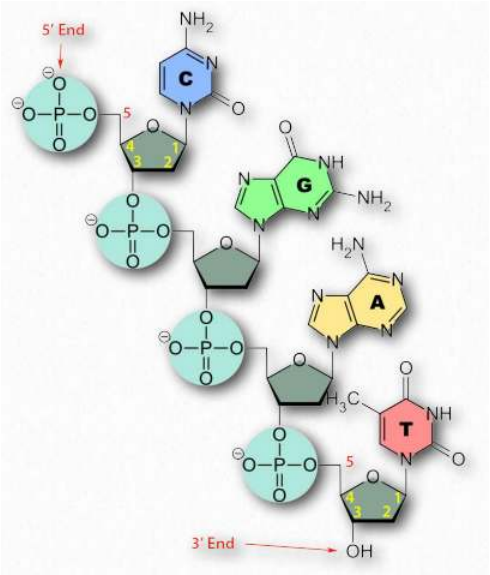


Figure 2.130 - 5’-3’ Polarity of a DNA strand

Figure 2.131 shows two strands of DNA (left and right). The strand on the left, from 5’ to 3’ reads T-C-G-A, whereas the strand on the right, reading from 5’ to 3’ is T-C-G-A. The strands in a double-stranded DNA are arranged in an anti-parallel fashion with the 5’ end of one strand across from the 3’ end of the other.

Hydrogen bonds

Hydrogen bonds between the base pairs hold a nucleic acid duplex together, with two hydrogen bonds per A-T pair (or per A-U pair in RNA) and three hydrogen bonds per G-C pair. The B-form of DNA has a prominent major groove and a minor groove tracing the path of the helix (Figure 2.132). Proteins, such as transcription factors bind in these grooves and access the hydrogen bonds of the base pairs to “read” the sequence therein.

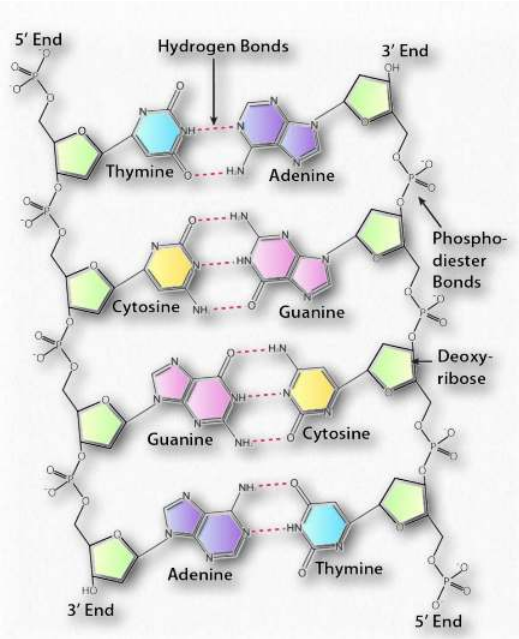


Figure 2.131 - Anti-parallel orientation of a DNA duplex, phosphodiester backbone, and base pairing Image by Aleia Kim

Other forms of DNA besides the B-form (Movie 2.5) are known (Figure 2.133). One of these, the A-form, was identified by Rosalind Franklin in the same issue of Nature as Watson and Crick's paper. Though the A-form structure is a relatively minor form of DNA and resembles the B-form, it turns out to be important in the duplex form of RNA and in RNA-DNA hybrids. Both the A form and the B-form of DNA have the helix oriented in what is termed the right-handed form.

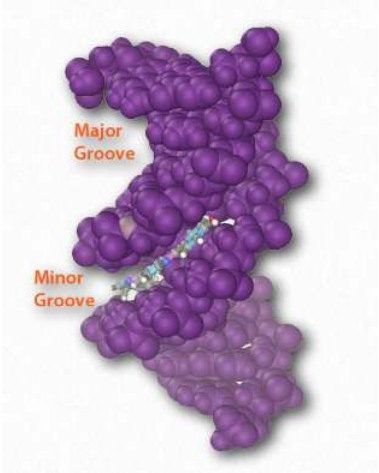
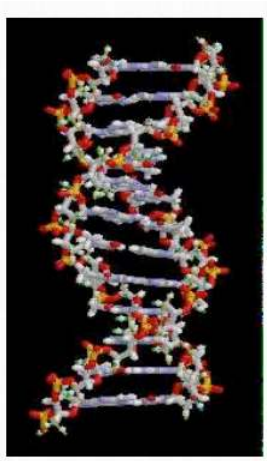


Figure 2.132 - Major and minor grooves of DNA. The minor groove has been bound by a dye Wikipedia



Movie 2.5 - B-form DNA duplex rotating in space Wikipedia

Z-DNA

The A-form and the B-form stand in contrast to another form of DNA, known as the Z-form. ZDNA, as it is known, has the same base-pairing rules as the B and A forms, but instead has the helices twisted in the opposite direction, making a left-handed helix (Figure 2.133). The Z-form has a sort of zig-zag shape, giving rise to the name Z-DNA.

In addition, the helix is rather stretched out compared to the A- and B-forms. Why are there different topological forms of DNA? The answer relates to both superhelical tension and sequence bias. Sequence bias means that certain sequences tend to favor the "flipping" of Bform DNA into other forms. ZDNA forms are favored by long stretches of alternating Gs and Cs. Superhelical tension is discussed below.

Superhelicity

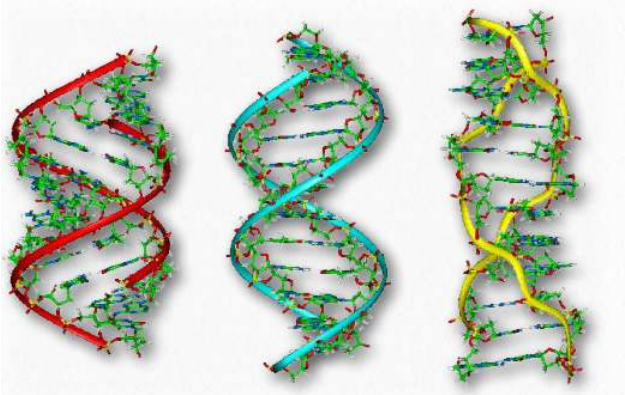


Figure 2.133 - From left to right, the A-, B-, and Zforms of DNA

Short stretches of linear DNA duplexes exist in the B-form and have 10.5 base pairs per turn. Double helices of DNA in the cell can vary in the number of base pairs per turn they contain. There are several reasons for this. For example, during DNA replication, strands of DNA at the site of replication get unwound at the rate of 6000 rpm by an enzyme called helicase. The effect of such local unwinding at one place in a DNA has the effect increasing the winding ahead of it. Unrelieved, such 'tension' in a DNA duplex can result in structural obstacles to replication.

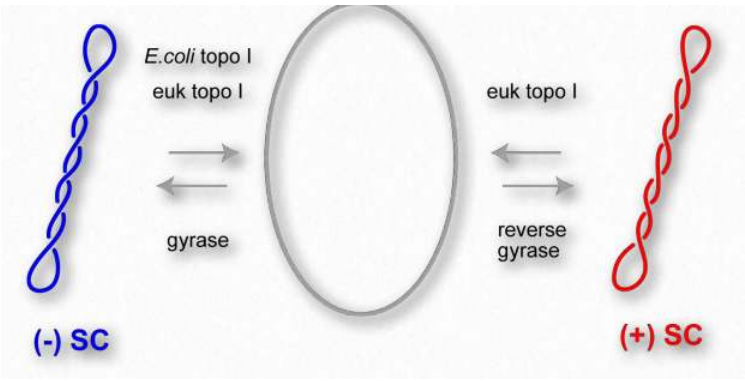


Figure 2.134 - Negatively supercoiled, relaxed circular, and positively supercoiled DNA interconverted by topoisomerase enzymes Wikipedia

Such adjustments can occur in three ways. First, tension can provide the energy for 'flipping' DNA structure. Z-DNA can arise as a means of relieving the tension. Second, DNA can 'supercoil' to relieve the tension (Figures 2.134 & 2.135). In this method, the duplex crosses over itself repeatedly, much like a rubber band will coil up if one holds one section in place and twists another part of it. Third, enzymes called topoisomerases can act to relieve or, in some cases, increase the tension by adding or removing twists in the DNA.

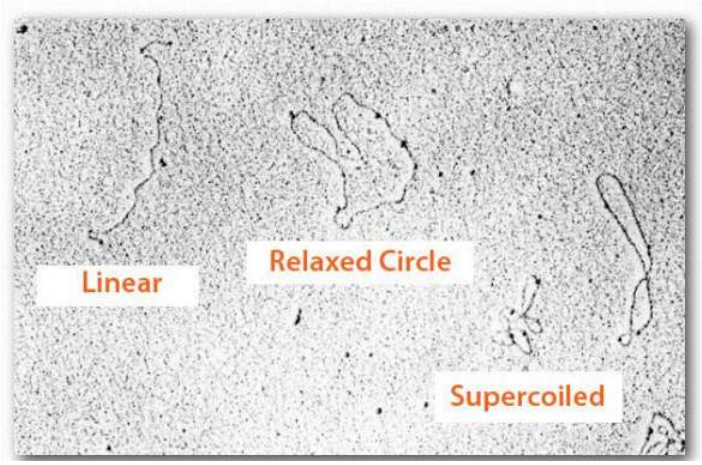


Figure 2.135 - Three topological forms of DNA Wikipedia

Topological isomers

As noted, so-called "relaxed" DNA has 10.5 base pairs per turn. Each turn corresponds to one twist of the DNA. Using enzymes, it is possible to change the number of base pairs per turn. In either the case of increasing or decreasing the twists per turn, tension is introduced into the DNA structure. If the tension cannot be relieved, the DNA duplex will act to relieve the strain, as noted. This is most easily visualized for circular DNA, though long linear DNA (such as found in eukaryotic chromosomes) or DNAs constrained in other ways will exhibit the same behavior.

Parameters

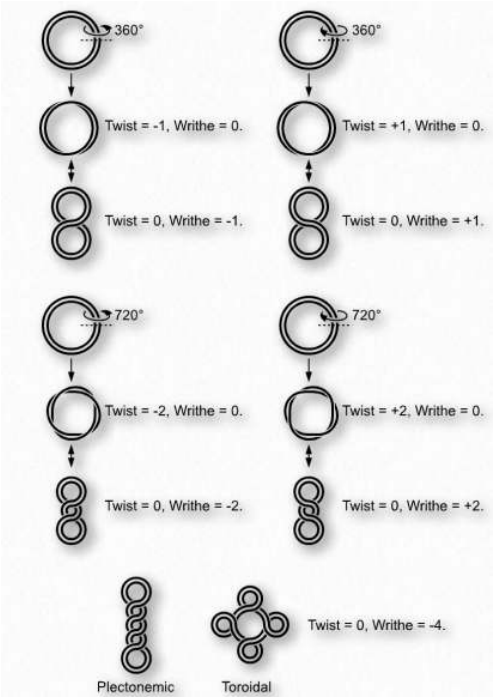


Figure 2.136 - Twists and writhes - DNA duplex coiling omitted for simplicity Wikipedia

To understand topologies, we introduce the concepts of ‘writhe’ and ‘linking number’. First, imagine either opening a closed circle of DNA and either removing one twist or adding one twist and then re-forming the circle. Since the strands have no free ends, they cannot relieve the induced tension by re-adding or removing the twists at their ends, respectively. Instead, the tension is relieved by “superhelices” that form with crossing of the double strands over each other (figure 8 structures in Figure 2.136). Though it is not apparent to visualize, each crossing of the double strands in this way allows twists to be increased or decreased correspondingly. Thus, superhelicity allows the double helix to reassume 10.5 base pairs per turn by adding or subtracting twists as necessary and replacing them with writhes.

We write the equation L= T + W where T is the number of twists in a DNA, W is the number of writhes, and L is the linking number. The linking number is therefore the sum of the twists and writhes. Interestingly, inside of cells, DNAs typically are in a supercoiled form. Supercoiling affects the size of the DNA (compacts it) and also the expression of genes within the DNA, some having enhanced expression and some having reduced expression when supercoiling is present. Enzymes called topoisomerases alter the superhelical density of DNAs and play roles in DNA replication, transcription, and control of gene expression. They work by making cuts in one strand (Type I topoisomerases) or both strands (Type II topoisomerases) and then add or subtract twists as appropriate to the target DNA. After that process is complete, the topoisomerase re-ligates the nick/cut it had made in the DNA in the first step.

Topoisomerases may be the targets of antibiotics. The family of antibiotics known as fluoroquinolones work by interfering with the action of bacterial type II topoisomerases. Ciprofloxacin also preferentially targets bacterial type II topoisomerases. Other topoisomerase inhibitors target eukaryotic topoisomerases and are used in anti-cancer treatments.

RNA

The structure of RNA (Figure 2.137) is very similar to that of a single strand of DNA. Built of ribonucleotides, joined together by the same sort of phosphodiester bonds as in DNA, RNA uses uracil in place of thymine. In cells, RNA is assembled by RNA polymerases, which copy a DNA template in the much same way that DNA polymerases replicate a parental strand. During the synthesis of RNA, uracil is used across from an adenine in the DNA template. The building of messenger RNAs by copying a DNA template is a crucial step in the transfer of the information in DNA to a form that directs the synthesis of protein. Additionally, ribosomal and transfer RNAs serve important roles in “reading” the information in the mRNA codons and in polypeptide synthesis. RNAs are also known to play important roles in the regulation of gene expression.

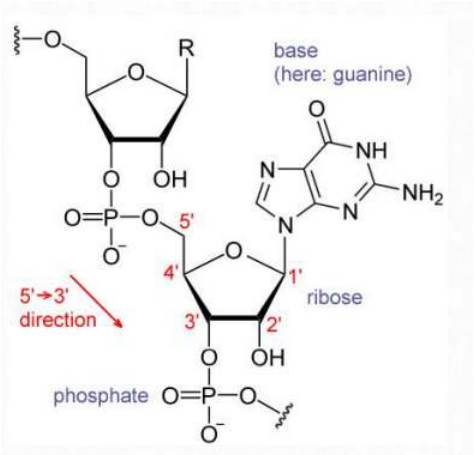


Figure 2.137 A section of an RNA molecule Wikipedia

RNA world

The discovery, in 1990, that RNAs could play a role in catalysis, a function once thought to be solely the domain of proteins, was followed by the discovery of many more so-called ribozymes- RNAs that functioned as enzymes. This suggested the answer to a long-standing chicken or egg puzzle - if DNA encodes proteins, but the replication of DNA requires proteins, how did a replicating system come into being? This problem could be solved if the first replicator was RNA, a molecule that can both encode information and carry out catalysis. This idea, called the “RNA world” hypothesis, suggests that DNA as genetic material and proteins as catalysts arose later, and eventually prevailed because of the advantages they offer. The lack of a 2’OH on deoxyribose makes DNA more stable than RNA. The double-stranded structure of DNA also provides an elegant way to easily replicate it. RNA catalysts, however, remain, as remnants of that early world. In fact, the formation of peptide bonds, essential for the synthesis of proteins, is catalyzed by RNA.

Secondary structure

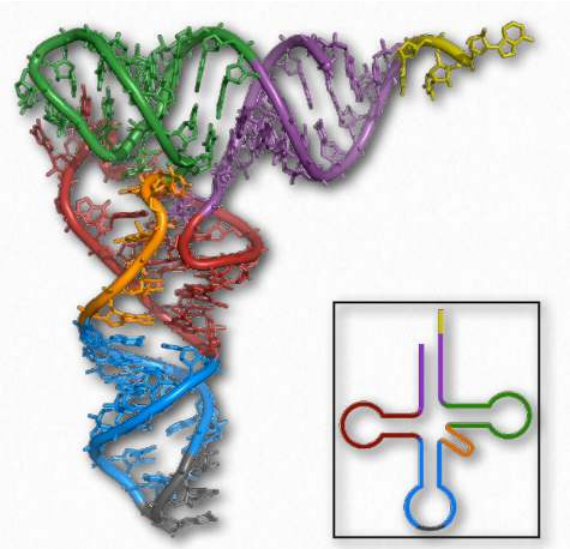


Figure 2.138 - tRNA Images - 3D projection (left) and 2D projection (inset) Wikipedia

With respect to structure, RNAs are more varied than their DNA cousin. Created by copying regions of DNA, cellular RNAs are synthesized as single strands, but they often have self-complementary regions leading to “foldbacks” containing duplex regions. These are most easily visualized in the ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) (Figure 2.138), though other RNAs, including messenger RNAs (mRNAs), small nuclear RNAs (snRNAs), microRNAs (Figure 2.139), and small interfering RNAs (siRNAs) may each have double helical regions as well.

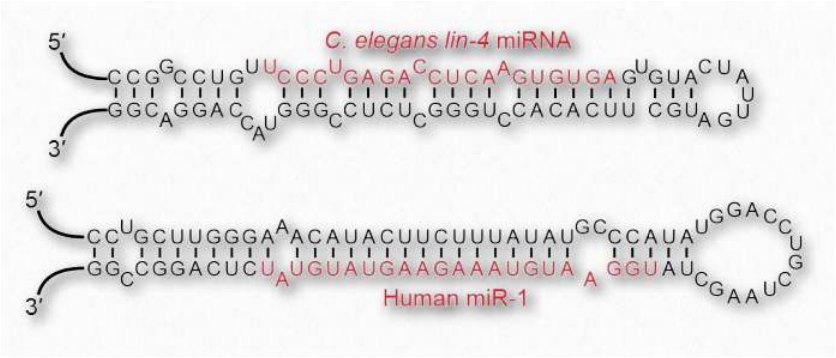


Figure 2.139 - MicroRNA stem loops. Wikipedia

Base pairing

Base pairing in RNA is slightly different than DNA. This is due to the presence of the base uracil in RNA in place of thymine in DNA. Like thymine, uracil forms base pairs with adenine, but unlike thymine, uracil can, to a limited extent, also base pair with guanine, giving rise to many more possibilities for pairing within a single strand of RNA.

These additional base pairing possibilities mean that RNA has many ways it can fold upon itself that single-stranded DNA cannot. Folding, of course, is critical for protein function, and we now know that, like proteins, some RNAs in their folded form can catalyze reactions just like enzymes. Such RNAs are referred to as ribozymes. It is for this reason scientists think that RNA was the first genetic material, because it could not only carry information, but also catalyze reactions. Such a scheme might allow certain RNAs to make copies of themselves, which would, in turn, make more copies of themselves, providing a positive selection.

Stability

RNA is less chemically stable than DNA. The presence of the 2’ hydroxyl on ribose makes RNA much more prone to hydrolysis than DNA, which has a hydrogen instead of a hydroxyl. Further, RNA has uracil instead of thymine. It turns out that cytosine is the least chemically stable base in nucleic acids. It can spontaneously deaminate and in turn is converted to a uracil. This reaction occurs in both DNA and RNA, but since DNA normally has thymine instead of uracil, the presence of uracil in DNA indicates that deamination of cytosine has occurred and that the uracil needs to be replaced with a cytosine. Such an event occurring in RNA would be essentially undetectable, since uracil is a normal component of RNA. Mutations in RNA have much fewer consequences than mutations in DNA because they are not passed between cells in division.

Catalysis

RNA structure, like protein structure, has importance, in some cases, for catalytic function. Like random coils in proteins that give rise to tertiary structure, single-stranded regions of RNA that link duplex regions give these molecules a tertiary structure, as well. Catalytic RNAs, called ribozymes, catalyze important cellular reactions, including the formation of peptide bonds in ribosomes (Figure 2.114). DNA, which is usually present in cells in strictly duplex forms (no tertiary structure, per se), is not known to be involved in catalysis.

RNA structures are important for reasons other than catalysis. The 3D arrangement of tRNAs is necessary for enzymes that attach amino acids to them to do so properly. Further, small RNAs called siRNAs found in the nucleus of cells appear to play roles in both gene regulation and in cellular defenses against viruses. The key to the mechanisms of these actions is the formation of short foldback RNA structures that are recognized by cellular proteins and then chopped into smaller units. One strand is copied and used to base pair with specific mRNAs to prevent the synthesis of proteins from them.

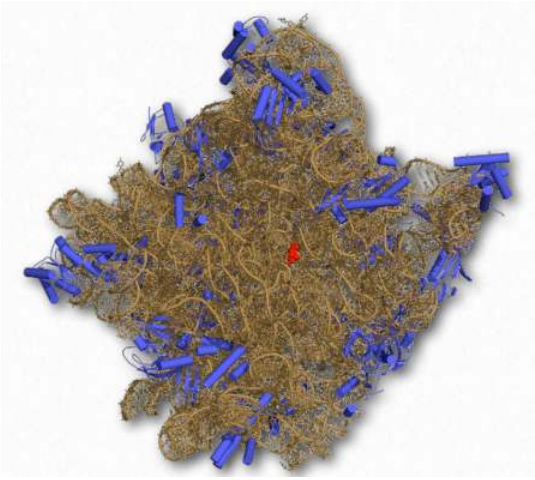


Figure 2.140 - Structure of the 50S ribosomal subunit. rRNA shown in brown. Active site in red Wikipedia

Denaturing nucleic acids

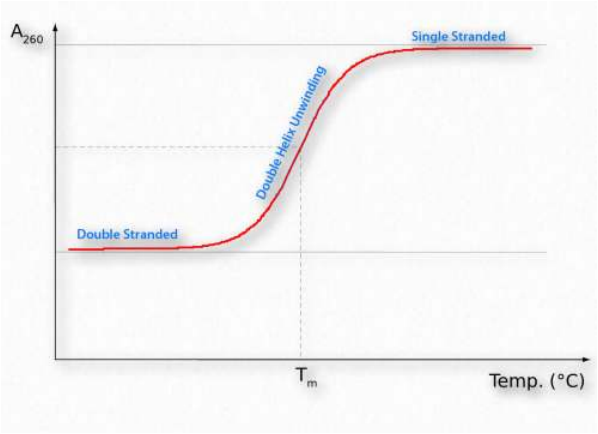


Figure 2.141 - The hyperchromic effect Wikipedia

Like proteins, nucleic acids can be denatured. Forces holding duplexes together include hydrogen bonds between the bases of each strand that, like the hydrogen bonds in proteins, can be broken with heat or urea. (Another important stabilizing force for DNA arises from the stacking interactions between the bases in a strand.) Single strands absorb light at 260 nm more strongly than double strands. This is known as the hyperchromic effect (Figure 2.141)and is a consequence of the disruption of interactions among the stacked bases. The changes in absorbance allow one to easily follow the course of DNA denaturation. Denatured duplexes can readily renature when the temperature is lowered below the “melting temperature” or Tm, the temperature at which half of the DNA strands are in duplex form. Under such conditions, the two strands can re-form hydrogen bonds between the complementary sequences, returning the duplex to its original state. For DNA, strand separation and rehybridization are important for the technique known as the polymerase chain reaction (PCR). Strand separation of DNA duplexes is accomplished in the method by heating them to boiling. Hybridization is an important aspect of the method that requires single stranded primers to “find” matching sequences on the template DNA and form a duplex. Considerations for efficient hybridization (also called annealing) include temperature, salt concentration, strand concentration, and magnesium ion levels (for more on PCR, see HERE).

DNA packaging

DNA is easily the largest macromolecule in a cell. The single chromosome in small bacterial cells, for example, can have a molecular weight of over 1 billion Daltons. If one were to take all of the DNA of human chromosomes from a single cell and lay them end to end, they would be over 7 feet long. Such an enormous molecule demands careful packaging to fit within the confines of a nucleus (eukaryotes) or a tiny cell (bacteria). The chromatin system of eukaryotes is the best known, but bacteria, too, have a system for compacting DNA.

DNA in Bacteria

In bacteria, there is no nucleus for the DNA. Instead, DNA is contained in a structure called a nucleoid (Figure 2.142). It contains about 60% DNA with much of the remainder comprised of RNAs and transcription factors. Bacteria do not have histone proteins that DNA wrap around, but they do have proteins that help organize the DNA in the cell - mostly by making looping structures.

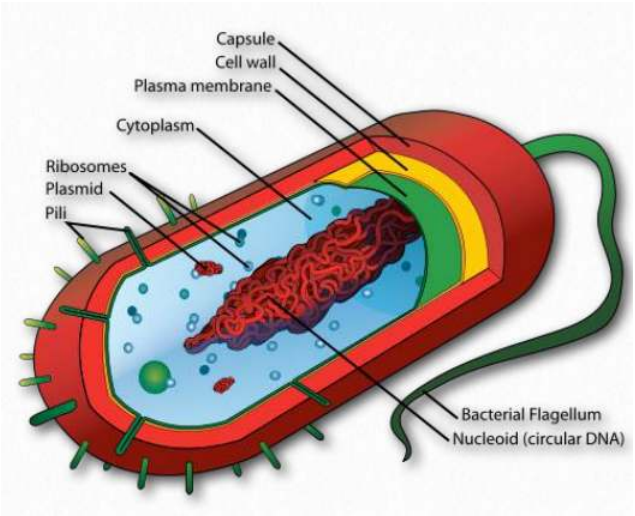


Figure 2.142 - Anatomy of a bacterial cell.

These proteins are known as Nucleoid Associated Proteins and include ones named HU, H-NS, Fis, CbpA, and Dps. Of these, HU most resembles eukaryotic histone H2B and binds to DNA non-specifically. The proteins associate with the DNA and can also cluster, which may be the origin of the loops. It is likely these proteins play a role in helping to regulate transcription and respond to DNA damage. They may also be involved in recombination.

Eukaryotes

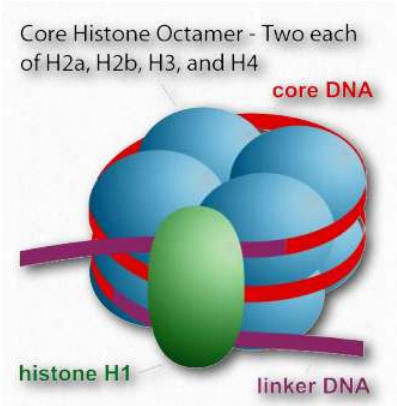


Figure 2.143 - Structure of a nucleosome Wikipedia

The method eukaryotes use for compacting DNA in the nucleus is considerably different, and with good reason - eukaryotic DNAs are typically much larger than prokaryotic DNAs, but must fit into a nucleus that is not much bigger than a prokaryotic cell. Human DNA, for example, is about 1000 times longer than c DNA. The strategy employed in eukaryotic cells is that of spooling - DNA is coiled around positively charged proteins called histones. These proteins, whose sequence is very similar in cells as diverse as yeast and humans, come in four types, dubbed H1, H2a, H2b, H3, and H4. A sixth type, referred to as H5 is actually an isoform of H1 and is rare. Two each of H2a, H2b, H3, and H4 are found in the core structure of what is called the fundamental unit of chromatin - the nucleosome (Figure 2.143).

Octamer

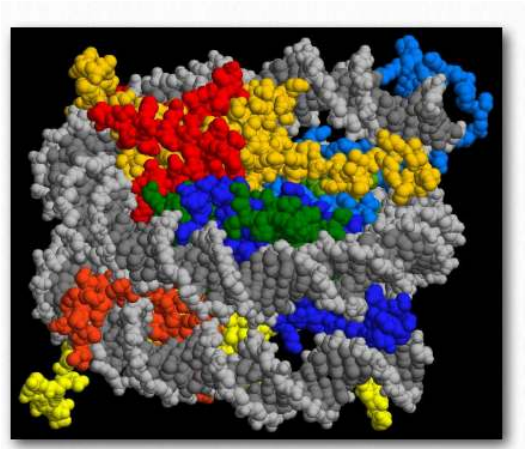


Figure 2.144 - Detailed core particle structure - DNA in gray, histones in red, yellow, green, blue Wikipedia

The core of 8 proteins is called an octamer. The stretch of DNA wrapped around the octamer totals about 147 base pairs and makes 1 2/3 turns around it. This complex is referred to as a core particle (Figure 2.144). A linker region of about 50-80 base pairs separate core particles. The term nucleosome then refers to a a core particle plus a linker region (Figure 2.143). Histone H1 sits near the junction of the incoming DNA and the histone core. It is often referred to as the linker histone. In the absence of H1, non-condensed nucleosomes resemble “beads on a string” when viewed in an electron microscope.

Histones

MARTKQTARKSTGGKAPRKQLASKAA
RKSAPSTGGVKKPHRYKPGTVALREIR
RFQKSTELLIRKLPFQRLVREIAQDFKT
DLRFQSSAIGALQESVEAYLVSLFEDT
NLASIHAKRVTIQKKDIKLARRLRGERS

Figure 2.145 - Amino acid sequence (1- letter code) of histone H3 of S. cerevisiae. Arginines (R) and lysines (K) shown in red.

Histone proteins are similar in structure and are rich in basic amino acids, such as lysine and arginine (Figure 2.145). These amino acids are positively charged at physiological pH, with enables them to form tight ionic bonds with the negatively charged phosphate backbone of DNA.

For DNA, compression comes at different levels (Figure 2.146). The first level is at the nucleosomal level. Nucleosomes are stacked and coiled into higher order structures. 10 nm fibers are the simplest higher order structure (beads on a string) and they grow in complexity. 30 nm fibers consist of stacked nucleosomes and they are packed tightly. Higher level packing produces the metaphase chromosome found in meiosis and mitosis.

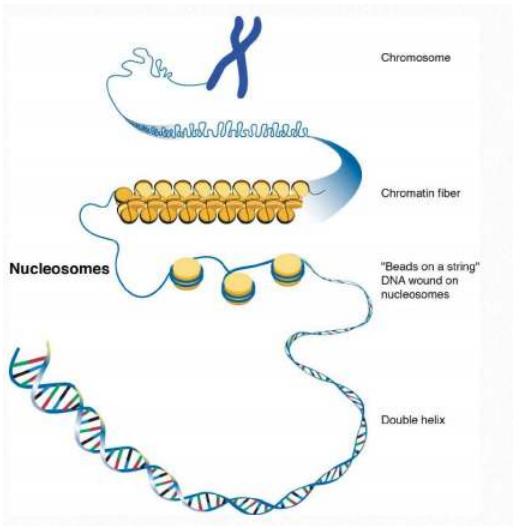


Figure 2.146 Packaging of DNA into a chromosome

The chromatin complex is a logistical concern for the processes of DNA replication and (particularly) gene expression where specific regions of DNA must be transcribed. Altering chromatin structure is therefore an essential function for transcriptional activation in eukaryotes. One strategy involves adding acetyl groups to the positively charged lysine side chains to “loosen their grip” on the negatively charged DNA, thus allowing greater access of proteins involved in activating transcription to gain access to the DNA.

This page titled 10.1: Structure and Function - Nucleic Acids is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **2.6: Structure and Function - Nucleic Acids** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **11.1: Structure and Function - Nucleic Acids** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

10.2: Nucleic Acids - Structure and Function

Learning Goals (ChatGPT o1, 1/30/25)

1. **Differentiate Between DNA and RNA Structures:**
   - Explain the basic composition of nucleic acids (sugar–phosphate backbone and nitrogenous bases).
   - Describe how the sugar (deoxyribose in DNA vs. ribose in RNA) and the specific bases (T in DNA vs. U in RNA) influence overall structure and function.
2. **Describe Canonical and Noncanonical Base Pairing:**
   - Illustrate the principles of Watson–Crick base pairing and how hydrogen bonding between complementary bases stabilizes the double helix.
   - Define and compare alternative (noncanonical) base pairing modes, including reverse Watson–Crick, wobble, Hoogsteen, and reverse Hoogsteen base pairs, and explain how these variations affect nucleic acid function.
3. **Understand DNA Conformations and Structural Dynamics:**
   - Compare the three main double-helical forms of DNA (A-DNA, B-DNA, and Z-DNA), focusing on differences in handedness, base pair spacing, twist angles, and groove dimensions.
   - Analyze how environmental factors (e.g., hydration, salt concentration) and molecular interactions induce transitions between these forms.
4. **Explore Higher-Order Nucleic Acid Structures:**
   - Describe the formation and biological significance of G-quadruplexes, including how guanine tetrads and metal ions (e.g., K^+) stabilize these structures.
   - Explain the mechanisms of triple helix formation, both intermolecular and intramolecular (H-DNA), and discuss their potential roles in gene regulation.
   - Recognize the characteristics of four-way junctions and their role in nucleic acid folding and interactions with small molecules.
5. **Analyze the Role of Base Stacking in Nucleic Acid Stability:**
   - Explain how base stacking and hydrophobic interactions contribute to the stability of the DNA double helix, in addition to hydrogen bonding.
   - Evaluate the effects of nucleotide sequence (e.g., GC vs. AT content) on the melting temperature (Tm) and overall stability of nucleic acids.
6. **Interpret Structural Models and Parameters:**
   - Utilize interactive molecular visualization tools (e.g., iCn3D) to explore nucleic acid structures.
   - Identify key structural parameters such as propeller and twist angles and understand how they influence nucleic acid conformation and dynamics.
7. **Relate Nucleic Acid Structure to Biological Function:**
   - Connect the structural features of DNA (e.g., major and minor grooves) with the mechanisms of protein recognition and binding.
   - Discuss how conformational flexibility (e.g., local “breathing” and dynamic unfolding) enables critical cellular processes like replication, transcription, and repair.
8. **Evaluate Experimental Approaches in Nucleic Acid Research:**
   - Summarize the historical and modern techniques used to elucidate nucleic acid structure, including X-ray crystallography, NMR, and cryo-electron microscopy.
   - Analyze how advances in structural biology have deepened our understanding of nucleic acid dynamics and their roles in cellular function.

These goals will help you integrate knowledge of nucleic acid chemistry with its biological implications, preparing you for advanced study and research in biochemistry.

10.2.1: Introduction to Nucleic Acids

Alongside proteins, lipids, and complex carbohydrates (polysaccharides), *nucleic acids* are one of the four major types of macromolecules that are essential for all known forms of life. The nucleic acids consist of two major macromolecules, *Deoxyribonucleic acid (DNA)* and *ribonucleic acid (RNA)* that carry the genetic instructions for the development, functioning, growth, and reproduction of all known organisms and viruses. Both consist of polymers of a sugar-phosphate-sugar backbone with organic heterocyclic bases attached to the sugars. The sugar in DNA is deoxyribose while in RNA it is ribose. DNA contains four bases, cytosine and thymine (pyrimidine bases) and guanine and adenine (purine bases). DNA in vivo consists of two antiparallel strands intertwined to form the iconic DNA double-stranded helix. RNA is single-stranded but may adopt many secondary and tertiary conformations not unlike that of a protein. Figure 10.2.1 shows a low-resolution comparison of the structure of DNA and RNA.

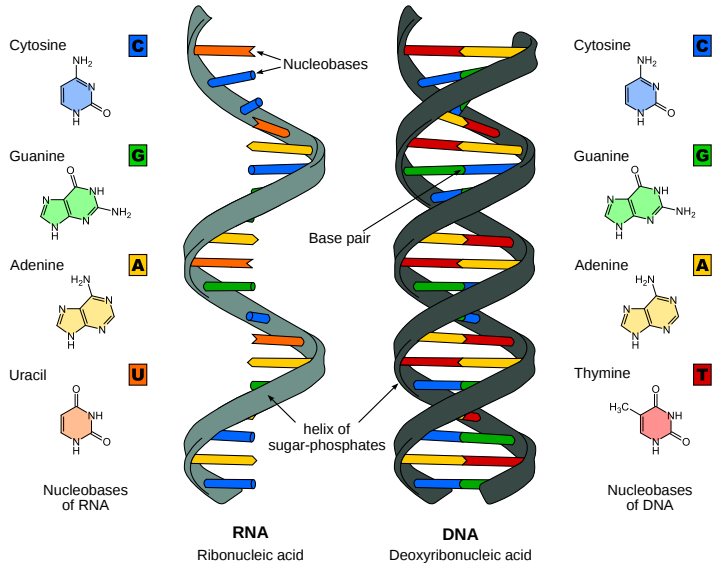


Figure 10.2.1: Low-resolution comparison of the structure of DNA and RNA. https://commons.wikimedia.org/wiki/F...DNA_RNA-EN.svg. Creative Commons Attribution-Share Alike 3.0 Unported license.

The biological function of DNA is quite simple: to carry and protect the genetic code. Its structure serves that purpose well. In the next section, we will study the functions of RNA, which are much more numerous and complicated. The structure of RNA has evolved to serve those added functions.

The core structure of a nucleic acid monomer is the *nucleoside*, which consists of a sugar residue + a nitrogenous base that is attached to the sugar residue at the 1′ position as shown in Figure 10.2.2. The sugar utilized for RNA monomers is ribose, whereas DNA monomers utilize deoxyribose, which has lost the hydroxyl functional group at the 2′ position of ribose. DNA contains four nitrogenous bases, including the *Purines*, Adenine (A), and Guanine (G), and the *Pyrimidines*, Cytosine (C), and Thymine (T). RNA uses the same nitrogenous bases as DNA, except for Thymine. Thymine is replaced with Uracil (U) in the RNA structure.

When one or more phosphate groups are attached to a nucleoside at the 5′ position of the sugar residue, it is called a *nucleotide*. *Nucleotides* come in three flavors depending on how many phosphates are included: the incorporation of one phosphate forms a *nucleoside monophosphate*, the incorporation of two phosphates forms a *nucleoside diphosphate*, and the incorporation of three phosphates forms a nucleoside triphosphate as shown in Figure 10.2.2.

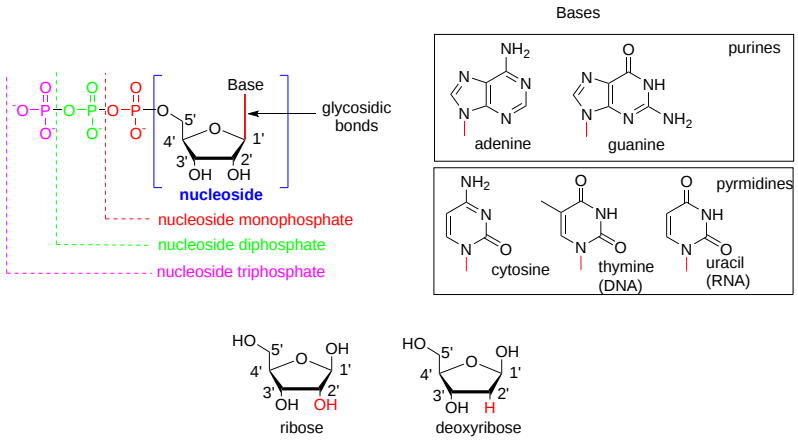


Figure 10.2.2: The Monomer Building Blocks of Nucleic Acids. The site of the nitrogenous base attached to the sugar residue (glycosidic bond) is shown in red.

10.2.2: DNA and RNA Hydrogen-bonded structures

Figure 10.2.3 below shows a "flattened" structure of double-stranded B-DNA that best shows the backbone and hydrogen-bonded base pairs between two antiparallel strands of the DNA. Unlike the protein α-helix, where the R-groups of the amino acids are positioned to the outside of the helix, in the DNA double-stranded helix, the nitrogenous bases are positioned inward and face each other. The backbone of the DNA is made up of repeating sugar-phosphate-sugar-phosphate residues. Bases fit in the double-helical model if pyrimidine on one strand is always paired with purine on the other. From Chargaff's rules, the two strands will pair A with T and G with C. This pairs a keto base with an amino base and a purine with a pyrimidine. Two H-bonds can form between A and T, and three can form between G and C. This third H-bond in the G:C base pair is between the additional exocyclic amino group on G and the C2 keto group on C. The pyrimidine C2 keto group is not involved in hydrogen bonding in the A:T base pair.

Furthermore, the orientation of the sugar molecule within the strand determines the directionality of the strands. The phosphate group that makes up part of the nucleotide monomer is always attached to the 5′ position of the deoxyribose sugar residue. The free end that can accept a new incoming nucleotide is the 3′ hydroxyl position of the deoxyribose sugar. Thus, DNA is directional and is always synthesized in the 5′ to 3′ direction. Interestingly, the two strands of the DNA double helix lie in opposite directions or have a head-to-tail orientation.

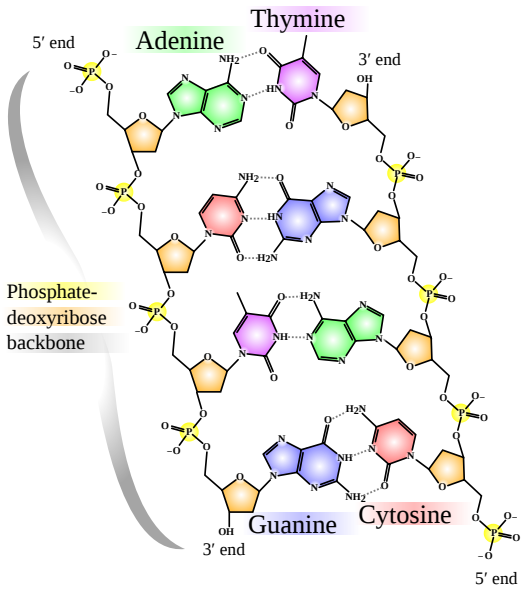


Figure 10.2.3: "Flattened" Structure of DNA Madeleine Price Ball. https://en.Wikipedia.org/wiki/File:D..._structure.svg. Wikimedia Commons

By analogy to proteins, DNA and RNA can be loosely thought to have primary and secondary structures. For a single strand, the primary sequence is just the base sequence read from the 5' to 3' end of the strand, with the bases thought of as "side chains" as illustrated in Figure 10.2.4 for an RNA strand which contains U instead of T.

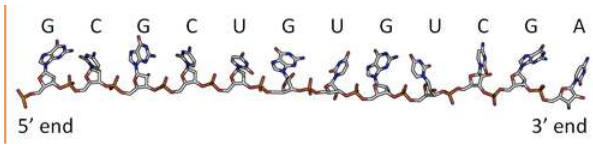


Figure 10.2.4: "Primary" sequence of a single RNA strand. https://en.Wikipedia.org/wiki/Nucleic_acid_sequence. Creative Commons Attribution-ShareAlike License

Since it is found partnered with another molecule (strand) of DNA, the double-stranded DNA, which consists of two molecules held together by hydrogen bonds, might be considered to have secondary structure (analogous to alpha and beta structure in proteins). Of course, the hydrogen bonds are not between backbone atoms but between side chain bases in double-stranded DNA.

Figure 10.2.5 shows an interactive iCn3D model of the iconic structure of a short oligomer of double-stranded DNA (1BNA).

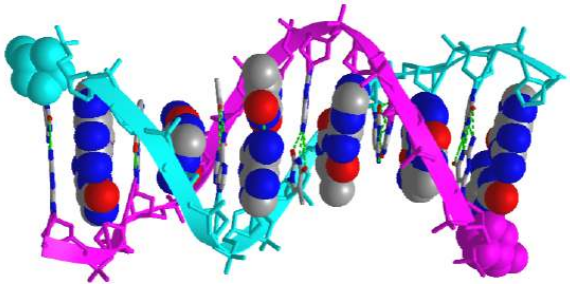


Figure 10.2.5: Iconic structure of a short oligomer of double-stranded DNA (1BNA). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...b5HUbmuQrCobg8

The backbones of the antiparallel strands are magenta (chain A) and cyan (chain B). Each chain's 5' sugar-phosphate end is shown in spacefill and colored magenta (chain A) and cyan (chain B). The hydrogen-bonded interstrand base pairs are shown in two ways: as spacefill and sticks to illustrate how the bases stack over each other.

Figure 10.2.6 shows types of "secondary (flat representations) and their 3D or tertiary representations found in nucleic acids.

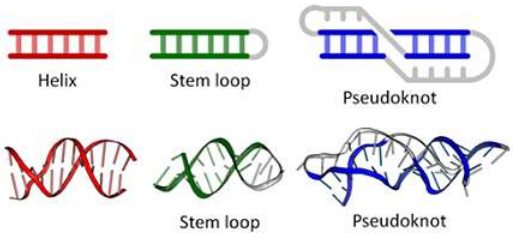


Figure 10.2.6: https://en.Wikipedia.org/wiki/Nucleic_acid_sequence. Creative Commons Attribution-ShareAlike License

Figure 10.2.7 shows an interactive iCn3D model of the tertiary structure of the T4 hairpin loop on a Z-DNA stem (1D16).

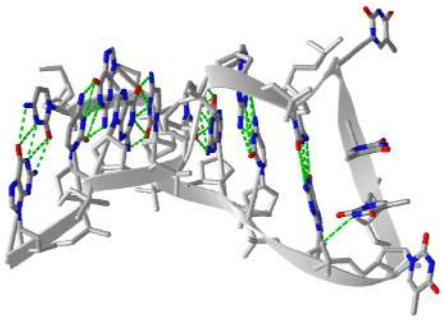


Figure 10.2.7: T4 hairpin loop on a Z-DNA stem (1D16). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...8C7qBqgh8ZTJH9

The hairpin shown is from a synthetic DNA oligomer C-G-C-G-C-G-T-T-T-T-C-G-C-G-C-G, which adopts an alternative Z-DNA conformation (which we will explore below) with a loop at one end. The thymine bases 7, 8, and 9 are generally perpendicular to one another and stack together, along with the ribose of T7.

Figure 10.2.8 shows an interactive iCn3D model of pseudoknot in RNA (437D).

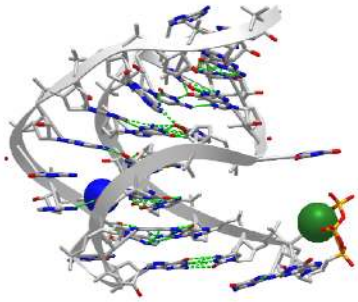


Figure 10.2.8: Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot (437D). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZtdeJqQXvjCKfA

The pseudoknot has two stems that form a "helix" and two loops. The knot consists of a hairpin in the nucleic acid structure, with the loop between the helices paired to another part of the nucleic acid. Pseudoknots can be found in mRNA and ribosomal RNA and affect the translation of the RNA (decoding to instruct the synthesis of a protein sequence). RNA viruses have pseudoknots, which likewise affect protein synthesis and RNA replication. Pseudoknots also occur in DNA.

10.2.3: Synthesis and structure of DNA

The nucleotide that is required as the monomer for the synthesis of both DNA and RNA is *nucleoside triphosphate*. During the incorporation of the nucleotide into the polymeric structure, two phosphate groups (PP_i, called *pyrophosphate*) from each triphosphate are cleaved from the incoming nucleotide and further hydrolyzed during the reaction, leaving a *nucleoside monophosphate* that is incorporated into the growing RNA or DNA chain as shown in **Figure** 10.2.9 below. The nucleophilic attack of the 3′-OH of the growing DNA polymer mediates the attack on the incoming nucleoside triphosphate. Thus, DNA synthesis is directional, only occurring at the 3′-end of the molecule.

The further hydrolysis of the pyrophosphate (P_i-P_i) releases a large amount of energy, ensuring that the overall reaction has a negative ΔG. Hydrolysis of P_i-P_i $\leftrightarrow$ $2P_i$ has a ΔG = -7 kcal/mol (-29 kJ/mol) and is essential to provide the overall negative ΔG (-6.5 kcal/mol, -27 kJ/mol) of the DNA synthesis reaction. Hydrolysis of the pyrophosphate also ensures that the reverse reaction, pyrophosphorolysis, will not occur, removing the newly incorporated nucleotide from the growing DNA chain.

This reaction is mediated in DNA by a family of enzymes known as DNA polymerases. Similarly, RNA polymerases are required for RNA synthesis. A more detailed description of polymerase reaction mechanisms will be covered in Chapter 24, which deals with DNA replication and repair, and DNA transcription in Chapter 25.

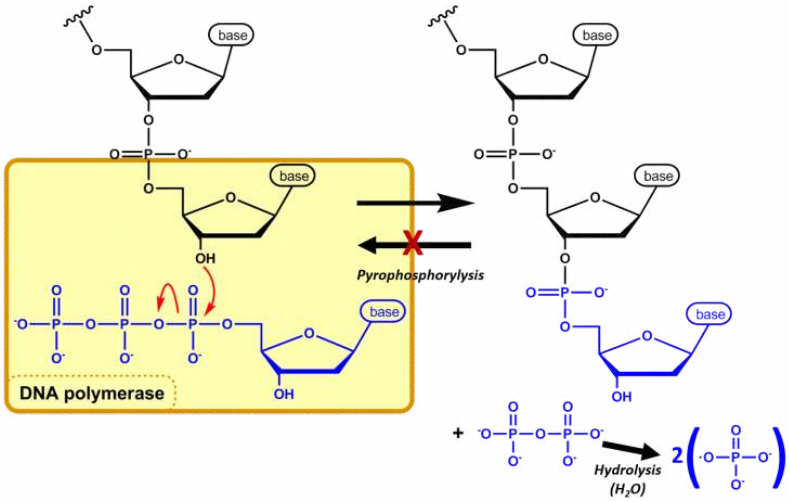


Figure 10.2.9: Nucleic Acid Synthesis: In nucleic acid synthesis, the 3' OH of a growing chain of nucleotides attacks the α-phosphate on the next NTP to be incorporated (blue), resulting in a phosphodiester linkage and the release of pyrophosphate (PPi). The DNA polymerase further mediates the hydrolysis of the pyrophosphate, preventing the reverse reaction from occurring and releasing enough energy to drive the reaction forward. The synthesis of DNA is shown in this diagram. *Image modified from Michal Sobkowski*

DNA was first isolated by Friedrich Miescher in 1869. The double-helix model of DNA structure was first published in the journal *Nature* by James Watson and Francis Crick in 1953 based upon the crucial X-ray diffraction image of DNA from Rosalind Franklin in 1952, followed by her more clarified DNA image with Raymond Gosling, Maurice Wilkins, Alexander Stokes, and Herbert Wilson, and base-pairing chemical and biochemical information by Erwin Chargaff. The prior model was triple-stranded DNA.

The realization that the structure of DNA is that of a double helix elucidated the mechanism of base pairing by which genetic information is stored and copied in living organisms and is widely considered one of the most important scientific discoveries of the 20th century. Crick, Wilkins, and Watson each received one-third of the 1962 Nobel Prize in Physiology or Medicine for their contributions to the discovery. (Franklin, whose breakthrough X-ray diffraction data was used to formulate the DNA structure, died in 1958 and thus was ineligible to be nominated for a Nobel Prize.)

Watson and Crick proposed two strands of DNA – each in a right-hand helix – wound around the same axis. The two strands are held together by H-bonding between the complementary base pairs (A pairs with T and G pairs with C) as shown in **Figure 10.2.10** below. Note that when looking from the top view, down on a DNA base pair, the position where the base pairs attach to the DNA backbone is not equidistant, but that attachment favors one side over the other. This creates unequal gaps or spaces in the DNA known as the *major groove* for the larger gap and the *minor groove* for the smaller gap (Figure 4.5). Based on the DNA sequence within the region, the hydrogen-bond potential created by the nitrogen and oxygen atoms present in the nitrogenous base pairs causes unique recognition features within the major and minor grooves, allowing for specific protein recognition sites to be created.

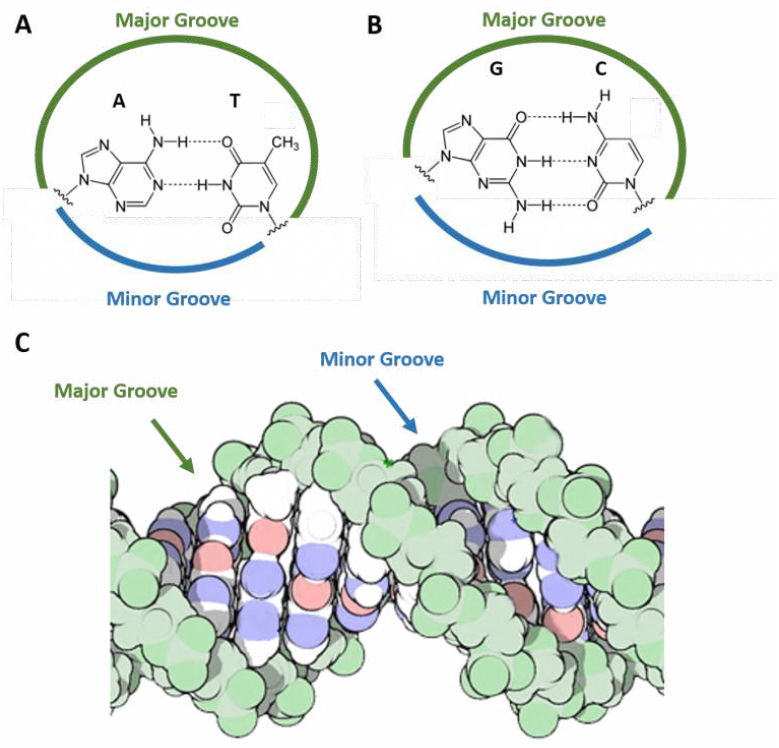


Figure 10.2.10: The Major and Minor Grooves of DNA. Top view of an (A) A-T base pair and a (B) G-C base pair showing the formation of the major and minor groove sides of the DNA. (C) Side view of the DNA double helix with the major and minor grooves indicated. The DNA backbone is green, potential nitrogen hydrogen-bonding locations are indicated in blue, and oxygen hydrogen-bonding locations are red. Figure C modified from *dullhunk*

Figure 10.2.1 shows a schematic representation of available hydrogen bond donors and acceptors in the major and minor grooves for TA and CG base pairs.

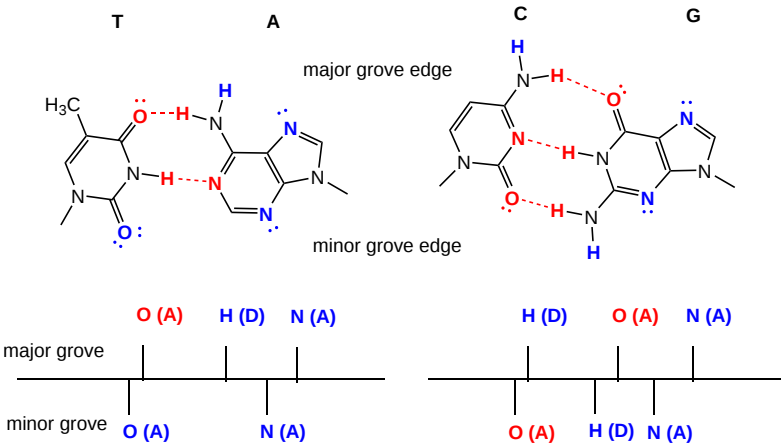


Figure 10.2.11: Available hydrogen bond donors and acceptors in the major and minor groove for TA and CG base pairs

Figure 10.2.12 shows an interactive iCn3D model of DNA showing the major and minor grooves.

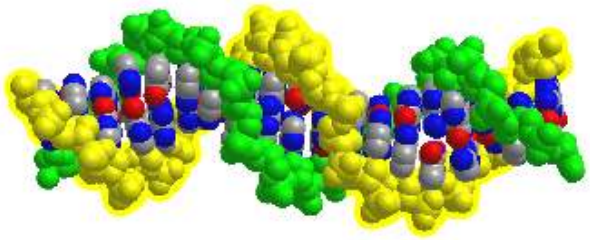


Figure 10.2.12: Major and minor grooves of ds-DNA (1D66). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?3xicqSv9ERCBPHvd6

The two sugar-phosphate backbones are shown in green and yellow. Some of the red (oxygen) and blue (nitrogen) atoms in the major groove (and to a much lesser extent in the minor groove) are not involved in inter-strand G-C and A-T base pairing and so would be available to hydrogen bond donors with specific binding proteins that would display complementary shape and hydrogen bond acceptors and donors. (The white spheres are Cd ions.)

Figure 10.2.13 shows an interactive iCn3D model of the N-terminal fragment of the yeast transcriptional activator GAL4 bound to DNA (1D66).

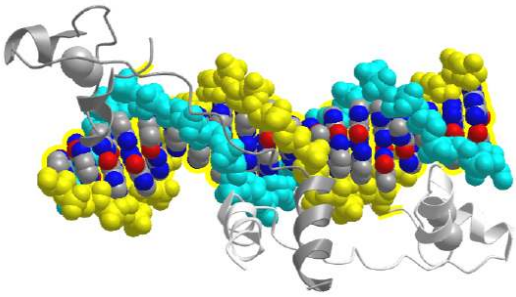


Figure 10.2.13: N-terminal fragment of the yeast transcriptional activator GAL4 bound to DNA (1D66). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5kLYSSfG7rsmS9

The N-terminal fragment binds to conserved CCG triplets found at both ends of the DNA in the major groove. The protein shown is a dimer held together by a short coiled-coil interaction domain, so the site has 2-fold symmetry. A small Zn^{2+}-containing secondary structure motif in each member of the dimer interacts with the major groove. An extended chain connects the DNA-binding and interaction domains of each protein.

In addition to the major and minor grooves providing variation within the double helix structure, the axis alignment of the helix, along with other influencing factors such as the degree of solvation, can give rise to three forms of the double helix, the *A-form (A-DNA)*, the *B-form (B-DNA)*, and the *Z-form (Z-DNA)* as shown in Figure 10.2.14.

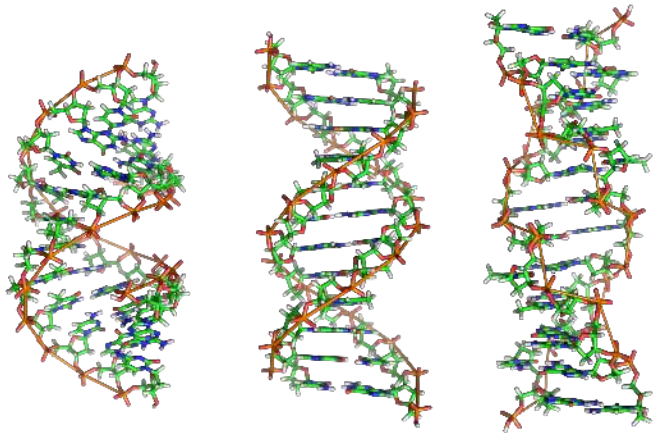


Figure 10.2.14: Structures of A-DNA (left), B-DNA (middle), and Z-DNA (right). https://en.Wikipedia.org/wiki/File:A..._and_Z-DNA.png. Creative Commons Attribution-Share Alike 3.0 Unported

Both the A- and B-forms of the double helix are right-handed spirals, with the B-form being the predominant form found *in vivo*. The A-form helix arises when conditions of dehydration below 75% of normal occur and has mainly been observed *in vitro* during X-ray crystallography experiments when the DNA helix has become desiccated. However, the A-form of the double helix can occur *in vivo* when RNA adopts a double-stranded conformation or when RNA-DNA complexes form. The 2′-OH group of the ribose sugar backbone in the RNA molecule prevents the RNA-DNA hybrid from adopting the B-conformation due to steric hindrance.

The third double helix type formed is a left-handed helical structure known as the Z-form or Z-DNA. Within this structural motif, the phosphates within the backbone appear to zigzag, providing the name Z-DNA. *In vitro*, the Z-form of DNA is adopted in short sequences that alternate pyrimidine and purines when high salinity is present. However, the Z-form has been identified *in vivo* within short regions of the DNA, showing that DNA is quite flexible and can adopt a variety of conformations. A comparison of features between A-, B-, and Z-form DNA is shown in Table 4.1.

Table 4.1 Comparisons of B-DNA, A-DNA, and Z-DNA

	B-DNA	A-DNA	Z-DNA
helix sense	Right Handed	Right Handed	Left Handed
base pairs per turn	10	11	12
vertical rise per bp	3.4 Å	2.56 Å	19 Å
rotation per bp	+36°	+33°	-30°
helical diameter	19 Å	19 Å	19 Å

The double-stranded helix of DNA is not always stable. This is because the attractions between the strands are noncovalent, reversible interactions. Depending on the DNA sequence, denaturation (melting) can be local or widespread, enabling various crucial cellular processes, including DNA replication, transcription, and repair.

Both sequence specificity and interaction (whether covalent or not) with a small compound or a protein can induce tilt, roll, and twist effects that rotate the base pairs in the x, y, or z axis, respectively, as seen in **Figure 10.2.15**, and can therefore change the helix's overall organization. Furthermore, slide or flip effects can also modify the geometrical orientation of the helix. Hence, the flip effects and (to a lesser extent) the other movements described above modulate the double-strand stability within the helix or at its ends. Indeed, under physiological conditions, local DNA 'breathing' has been evidenced at both ends of the DNA helix, and B-DNA to Z-DNA structural transitions have been observed in internal DNA regions. These locally open DNA structures are good substrates for specific proteins, which can also induce the opening of a 'closed' helix. The DNA replication and repair processes will be discussed in more detail in Chapter 24.

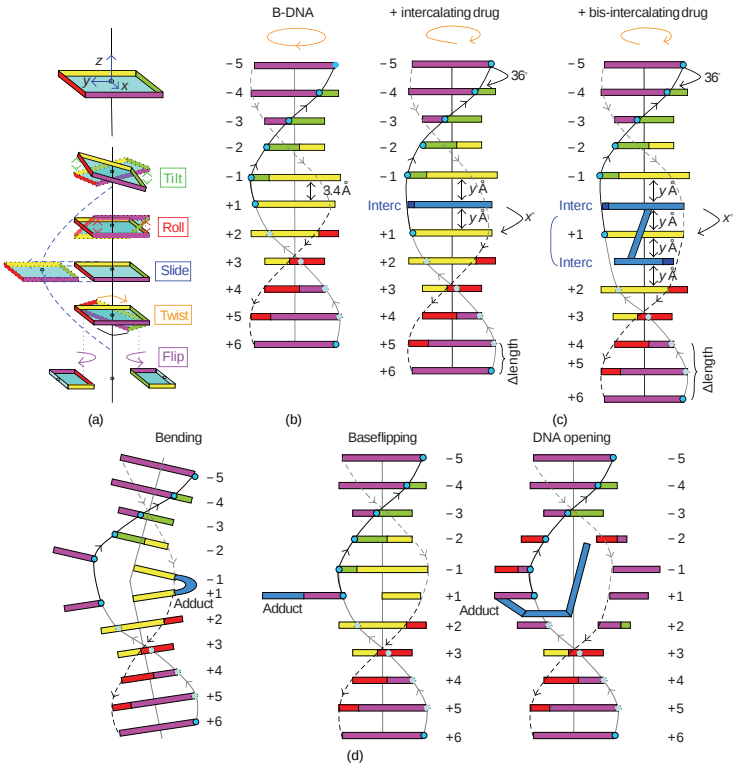


Figure 10.2.15: Localized Structural Modification of the DNA Double Helix. (a) Base pair orientations with the x, y, and z axes result in different kinds of rotation (tilt, roll, or twist) or slipping of the bases (slide, flip) regarding the central helix axis. (b) Mature B-DNA has nearly 11 base pairs within one helical turn. (c) Mono- or bis-intercalation of a small molecule (shown in blue) between adjacent base pairs, resulting in an unwinding of the DNA helix (orange arrow on the top) and a lengthening of the DNA helix (ΔLength) depending on the X and y Å values that are specific for a defined DNA intercalating compound. (d) Representation of the DNA bending, base flipping, or double-strand opening induced by some DNA destabilizing alkylating agents (adducts shown in blue). Adapted from Calladine and Drew's schematic box representation. Lenglet and David-Cordonnier (2010) Journal of Nucleic Acids, http://dx.doi.org/10.4061/2010/290935. Creative Commons Attribution License,

Figure 10.2.16 shows interactive iCn3D models of A-DNA (top), B-DNA (center), and Z-DNA (bottom). (Copyright; author via source). Click the image for a popup or use the external links in column 1.

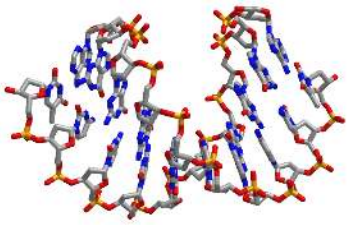
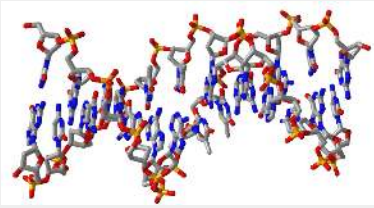
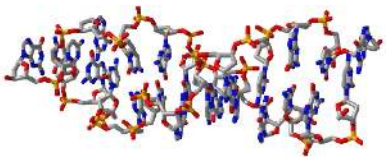
A-DNA (440D)	
B-DNA (1BNA)	
Z-DNA (4OCB)	

Figure 10.2.16: A, B, and Z-DNA. Click the image for a popup or use the links in column 1

We studied the structure of proteins in-depth, discussing resonance in the peptide backbone, allowed backbone angles φ, ψ, and ω, side chain rotamers, Ramachandran plots, and different structural motifs. We also explored them dynamically using molecular dynamics simulations. We also discussed the thermodynamics of protein stability and how stability could be altered by changing environmental factors such as solution composition and temperature.

In contrast, our understanding of the structural parameters and the dynamics of nucleic acids is less advanced. This may seem paradoxical, especially given the apparent simplicity of the iconic structure of DNA presented in textbooks. Yet, we should look at the types of secondary structures of nucleic acids presented and then the complicated tertiary and quaternary structures of RNA.

The nucleic acid backbone has a 5-membered sugar ring, which adds rigidity to the backbone, linked to another sugar ring by $CH_2O(PO_3)O$- connectors, adding some additional conformational freedom. We'll explore the effects of the pentose ring geometry in RNA and DNA in Chapter 8.3. To illustrate a yet unexplored complexity of nucleic acid structure, consider just the orientation of rings in double-stranded DNA and in regions of RNA where double-stranded structures form. The variants in the orientation of the hydrogen-bonded base pairs and the corresponding parameters that define them are shown in Figure 10.2.17.

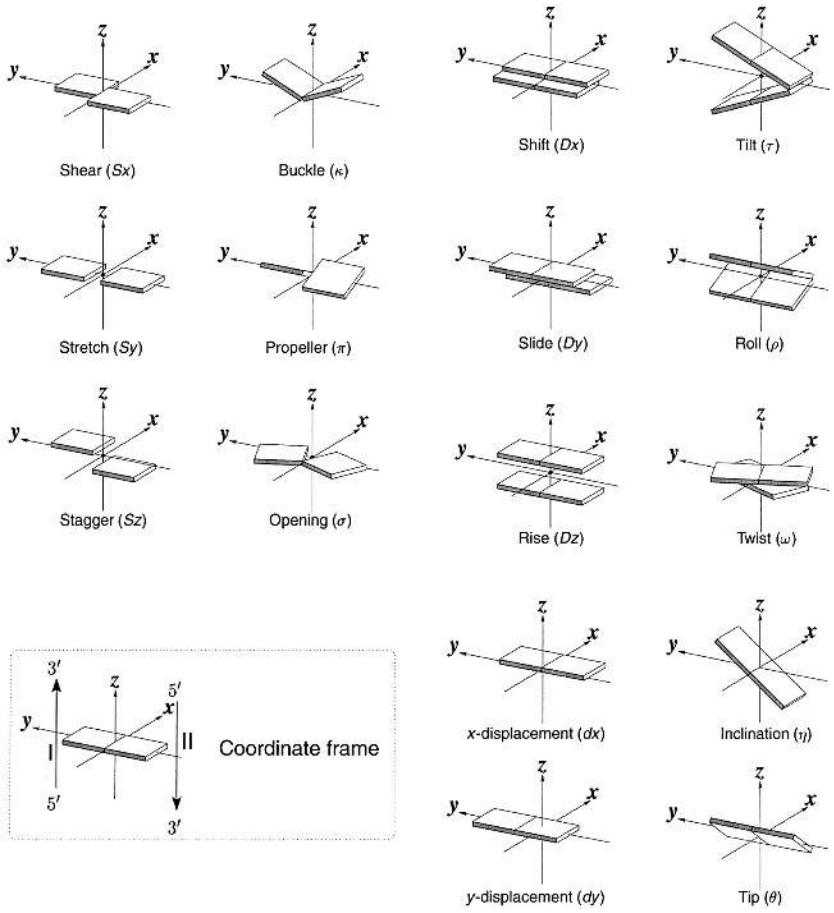


Figure 10.2.17: Base pair orientation and corresponding parameters in nucleic acids. http://x3dna.org/highlights/schemati...air-parameters (with permission). 2008 3DNA *Nature Protocols* paper (NP08), the initial 3DNA *Nucleic Acids Research* paper .

Consider just two of these: the **propeller** and **twist** angles. If you examine the iCn3D models of nucleic acids presented above, you will see that the base pairs are not perfectly flat but are twisted. Larger propeller angles are associated with increased rigidity. The propeller angles for A, B, and Z DNA are + 18°, + 16 +/-7 ^{o}, and about 0 ^{o}, respectively. The twist angles A, B, and Z DNA are +33 ^{o}, +36 ^{o}, and -30 ^{o}, respectively. The lower the twist angle, the higher the number of base pairs per turn. This, of course, affects the pitch of the helix (the length of one complete turn). These terms should be minimized to computationally determine the lowest energy state for a given double-stranded nucleic acid.

10.2.4: Alternative Base Pairing in DNA and RNA

A first glance at a DNA or RNA structure reveals a myriad of possible hydrogen bond donors and acceptors in the nucleic acid bases. Hence, it should come as no surprise that a variety of alternative or **noncanonical** (not in the canon or dogma) intermolecular hydrogen bonds can form between and among bases, leading to alternatives to the classical Watson-Crick base pairing. There are 28 possible base pairs with two hydrogen bonds between them. As structure determines function and activity, these alternative structures influence DNA/RNA function. We will consider four types of noncanonical base pairing: reverse Watson-Crick, wobble, Hoogsteen, and reverse Hoogsteen base pairs.

These noncanonical base pairs can occur in DNA when bases become mismatched in double-stranded regions. In RNA, which we will explore more fully in Chapter 8.2, double-stranded molecules formed by separate RNA molecules aren't common. Instead, the molecule folds on itself in 3D space to form a complex tertiary structure containing regions of helical secondary structure. RNAs also form quaternary structures when bound to other nucleic acids and proteins. Larger RNAs have loops with complex secondary and tertiary structures, which often require noncanonical base pairing, stabilizing the alternative structures. The noncanonical structures are also important for RNA-protein interactions in the RNA region that binds proteins. As with protein-ligand interactions, protein binding to RNA might also induce conformation changes, specifically noncanonical base pairs, in the RNA.

For example, the HIV Rev peptide binds to a target site in the envelope gene of HIV (which has an RNA genome) and leads to the formation of an RNA loop with hydrogen bonding between two purines.

Figure 10.2.18 shows an interactive iCn3D model of the REV Response element RNA complexed with REV peptide (1ETF).

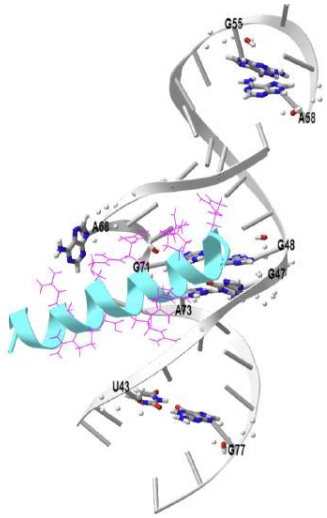


Figure 10.2.18: REV Response element RNA complexed with REV peptide (1ETF). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...T8CJ3pCe986Vx9

The peptide is shown in cyan, and its arginine side chains are shown as cyan lines. An extraordinary number of arginines form ion-ion interactions with the negatively charged phosphates in the major groove of this double-stranded A-RNA. The noncanonical base pairs are shown in CPK-colored sticks. A wobble base, U43-G77 (see below), can be seen, as well as three homopurine base pairs, G47-A73, G55-A58, and G48-G71. The solitary A68 base is shown projecting away from the RNA.

Figure 10.2.19 shows the Watson-Crick model and the first set of alternative non-canonical base pairs.

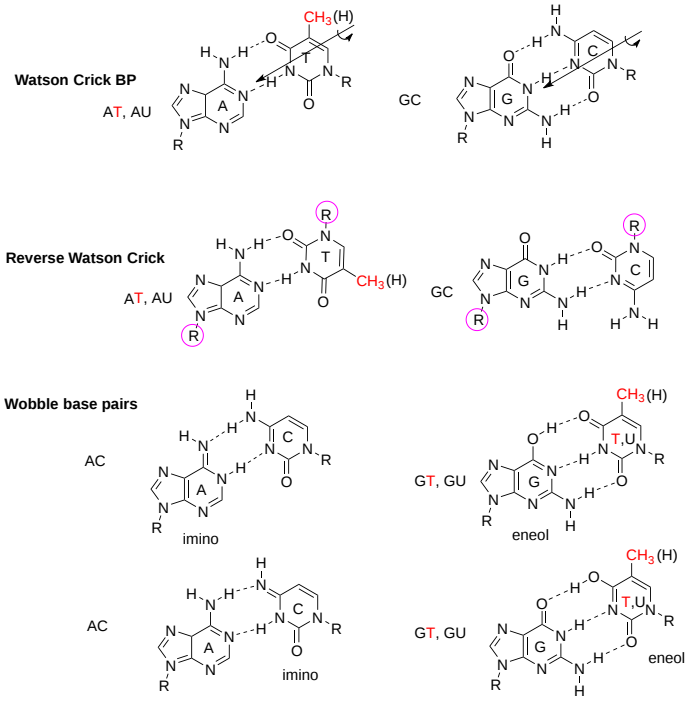


Figure 10.2.19: Some noncanonical base nucleic acid base pairs

Let's look at them in more detail.

Reverse Watson Crick: The reverse Watson-Crick AT (AU) and GC pairs can sometimes be found at the end of DNA strands and RNA. In forming the reverse base pairs, the pyrimidine can rotate 180 ^{o} along the axis shown and then rotate in the plane to align the hydrogen bond donors and acceptors, as shown in the top part of the figure. The glycosidic bond between the N in the base and the sugar (the circled R group) is now in an "antiparallel" arrangement in the reverse base pair.

Wobble Base Pairs

The bases in nucleic acids can undergo tautomerization to produce forms that can base pair noncanonically. They are termed **wobble** base pairs and include G-T(U) base pairs from keto-enol tautomerism and A-C base pairs from amino–imino tautomerism, as illustrated in Figure 18 above.

Figure 10.2.20 shows an interactive iCn3D model of the GT Wobble Base-Pairing in Z-DNA form of d(CGCGTG) (1VTT). Two such GT pairs are found in the structure.

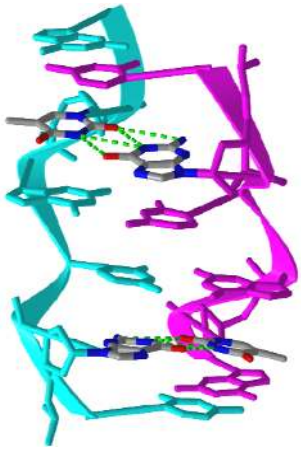


Figure 10.2.20: GT Wobble Base-Pairing in Z-DNA form of d(CGCGTG) (1VTT). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LtwfzyeqDCaPEA

The water around the wobble base pairs can form hydrogen bonds and stabilize the pair if a hydrogen bond is missing.

Figure 10.2.21 shows an interactive iCn3D model of dsRNA with G-U wobble base pairs (6L0Y).

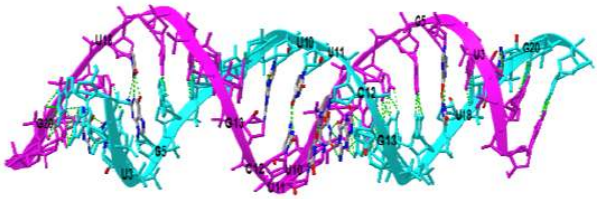


Figure 10.2.21: dsRNA with G-U wobble base pairs (6L0Y). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...reyeD6JQM1djq6

The structure contains many GU wobble base pairs and two CU base pairs between two pyrimidine bases.

Inosine, a variant of the base adenine, can be found in RNA. It is formed by the deamination of adenosine by the enzyme adenosine deaminase. A nucleotide having inosine is named hypoxanthine. Hypoxanthine can form the wobble base pairs I-U, I-A, and I-C when incorporated into RNA, as illustrated in Figure 10.2.22.

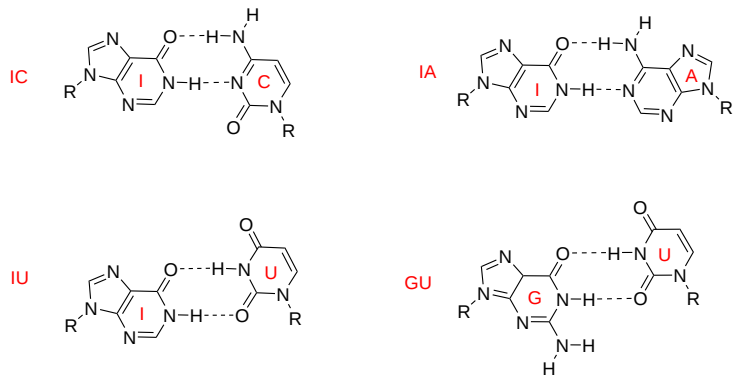


Figure 10.2.22: Wobble bases pairs using hypoxanthine with the base inosine

Wobble base pair interactions are especially important in translation when a protein sequence is made from a messenger RNA template (which will be discussed in Unit III). For that decoding process, two RNA molecules, messenger RNA (mRNA) and a transfer RNA (tRNA) covalently attached to a specific amino acid like glutamic acid, must bind to each other through a 3-base pair interaction. The three bases on the mRNA are called the codon, and the three complementary bases on the tRNA are called the anticodon. The triplet base pairs are antiparallel to each other. The interaction between mRNA and tRNA is illustrated in Figure 10.2.23.

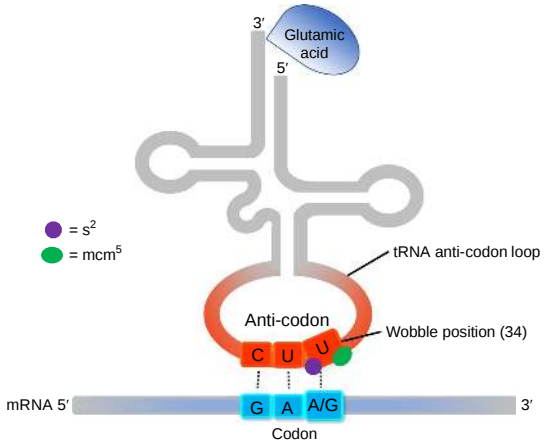


Figure 10.2.23: The wobble uridine (U34) of tRNA molecules that recognize both AA and AG-ending codons for Lys, Gln, and Glu, is modified by the addition of both a thiol (s2) and a methoxy-carbonyl-methyl (mcm5). This double modification enhances the translational efficiency of AA-ending codons. Goffena, J et al. *Nat Commun* **9,** 889 (2018). https://doi.org/10.1038/s41467-018-03221-z. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The third 3' base in the mRNA is less restricted and can form noncanonical, specifically, wobble base pairs, with the 5' base in the anti-codon triplet of tRNA. The term wobble arises from the subtle conformational changes used to optimize the pairing of the triplets. Wobble bases occur much more in tRNA than in other nucleic acids.

Hoogsteen base pairing

Flexibility in DNA allows rotation around the C1'-N glycosidic bond connecting the deoxyribose and base in DNA, allowing different orientations of AT and GC base pairs with each other. The normal "anti" orientation allows "Watson-Crick" (WC) base pairing between AT and GC base pairs, while the altered rotation allows "Hoogsteen" base pairs. Figure 10.2.24 shows the different orientations for an AT base pair.

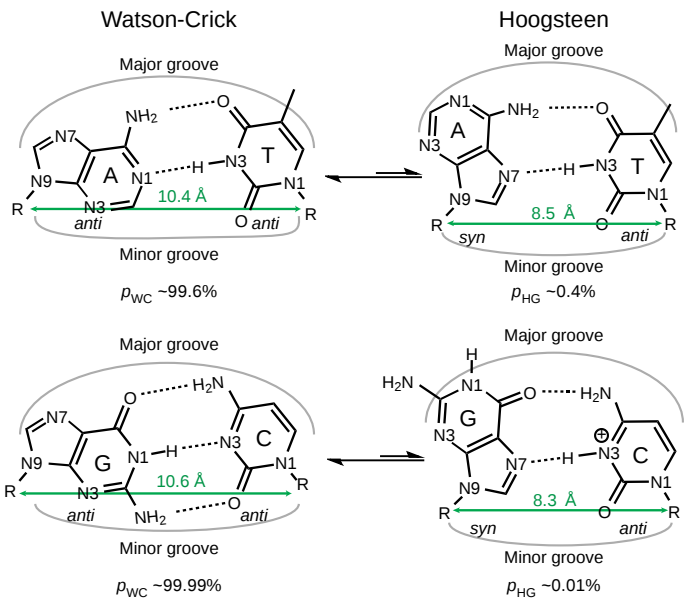


Figure 10.2.24: Xu, Y., McSally, J., Andricioaei, I. *et al*. Modulation of Figure 10.2.xxHoogsteen dynamics on DNA recognition. *Nat Commun* **9,** 1473 (2018). https://doi.org/10.1038/s41467-018-03516-1Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Hoogsteen base pairing is usually seen when DNA is distorted through interactions with bound proteins and drugs that intercalate between base pairs. Figure 10.2.25 shows an interactive iCn3D model of a Hoogsteen base pair embedded in undistorted B-DNA - MATAlpha2 homeodomain bound to DNA (1K61).

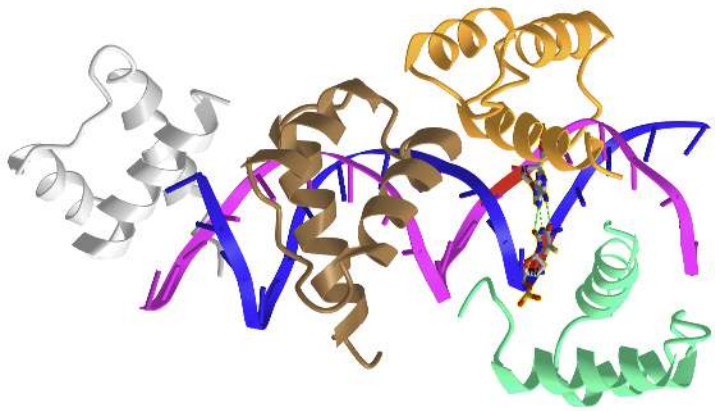


Figure 10.2.25: A Hoogsteen base pair embedded in undistorted B-DNA - MATAlpha2 Homeodomain bound to DNA (1K61). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SLLRv1m8HQXKcA

The same DNA without bound protein has no Hoogsteen base pairs. To form Hoogsteen base pairs, a rotation around the glycosidic-base bond must occur. Hoogsteen base pairs between G and C can also occur on rotation, but in addition, the N3 of cytosine is protonated, as shown in Figure 14 above.

Evidence suggests that Hoogsteen base pairing may be important in DNA replication, binding, damage, or repair. They can induce kinking of the DNA near the major groove.

There are also reverse Hoogsteen base pairing examples, as shown in Figure 10.2.26.

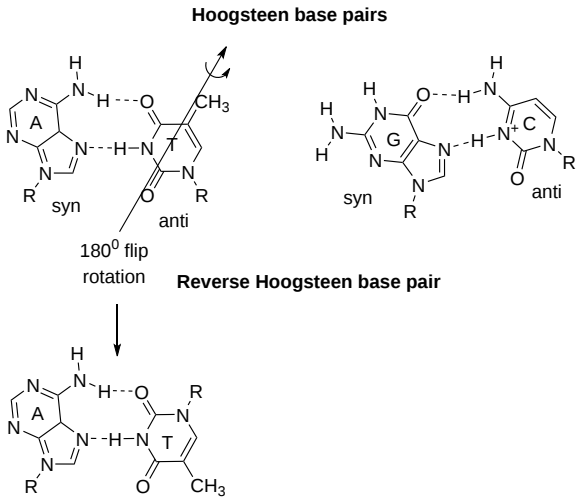


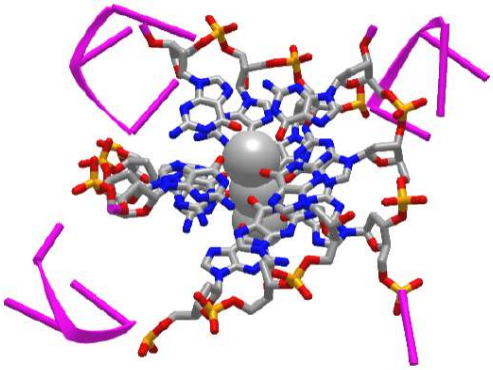
Figure **10.2.26**: The reverse Hoogsteen AT base pair

10.2.5: Additional Alternative Structures: Quadruplexes, Triple Helices, and 4-Way Junctions

Quadruplexes

These can be formed in DNA and RNA from G-rich sequences involving tetrads of guanine bases that are hydrogen-bonded. They are a bit hard to describe in words, so let's first examine one particular structure.

In human cells, telomeres (the ends of chromosomes) contain 300-8000 repeats of a simple **TTAGGG** sequence. The repetitive **TTAGGG** sequences in telomeric DNA can form quadruplexes. Figure **10.2.27** shows an interactive iCn3D model of parallel quadruplexes from human telomeric DNA (1KF1). The structure contains a single DNA strand **(**5'-AGGGTTAGGGTTAGGGTTAGGG-3') which contains four TTAGGG repeats.



NCBI iCn3D Figure **10.2.27**: parallel quadruplexes from human telomeric DNA (1KF1). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...y5joFHDgWJQsQ6

Rotate the model to see three parallel layers of quadruplexes. In each layer, **4** **noncontiguous** guanine bases interact with a K^+ ion. Hover over the guanine bases in one layer, and you will find that one layer consists of guanines 4, 10, 16, and 22, which derive from the last G in each of the repeats in the sequence of the oligomer used (5'-AGGGTTAGGGTTAGGGTTAGGG-3'). These quadruplexes certainly serve in recognition and as binding sites for telomerase proteins. The guanine-rich telomere sequences, which can form quadruplexes, may also function to stabilize chromosome ends.

A Quadruplex can be formed in 1 strand of nucleic acid (as in the above model) or from 2 or 4 separate strands. They also must have at least two stacked triads. As in the example above, single-stranded sections can form an intramolecular G-quadruplex from a $G_mX_nG_mX_oG_mX_pG_m$ sequence, where m is the number of Gs in each short segment (3 in the structure above). A G might be in a loop if a segment is longer than the others.

Triple Helices

These structures can occur in DNA (and also RNA) that contain homopurine and homopyrimidine sequences that have a mirror repeat symmetry. Hence, they can occur naturally. A mirror repeat contains a center of symmetry on a single strand. Here is an example: 5'-GCATGGTACG-3'.

They can also occur when a third single-stranded DNA (a triplex-forming oligonucleotide or TFO) binds to a double-stranded DNA. The TFOs bind through Hoogsteen base pairing in the major groove of the ds-DNA. They can bind tightly and specifically in a parallel or antiparallel fashion. Specific and locally higher concentrations of divalent cations or positively charged polyamines like spermine act to stabilize the extra negative charge density from the binding of a third polyanionic DNA strand.

An example of a triple helix system that has been studied in vitro is shown in Figure 10.2.28.

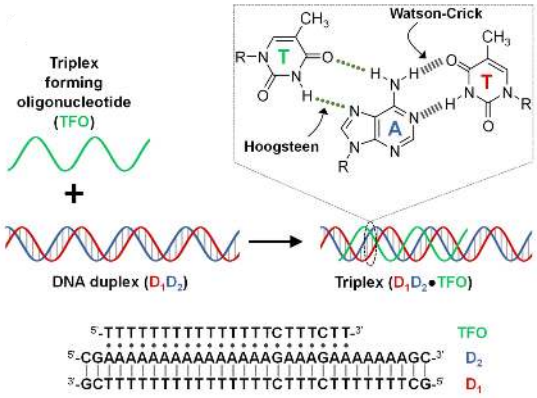


Figure 10.2.28: Intermolecular triplex formation and their oligonucleotide sequences (where "•" and "-" indicate Hoogsteen and Watson–Crick base pairings, respectively). Inset: chemical structure of a parallel T•AT triplet. Guerrini, L. and Alvarez-Puebla, R.A. Nanomaterials 2021, 11, 326. https://doi.org/10.3390/nano11020326. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

The double-stranded canonical helix (D_1D_2) consists of 31 base pairs in which strand D_1 is pyrimidine-rich and D_2 is purine-rich strand (D_2). A 22-nucleotide Triple helix-forming oligonucleotide (TFO) that is rich in pyrimidines binds the 19 AT and 2 C-GC base triplets. The TFO binds along the major groove of the D2 strand, which is purine-rich.

If the binding of the third strand in the major groove occurs at the site where RNA polymerase binds to a gene, then the third strand can inhibit gene transcription. Binding can also lead to a mutation or recombination at the site.

Figure 10.2.29 shows the base pairing of purine and pyrimidines of the third strand to the canonical AT dn GC base pairs of the original double-stranded DNA.

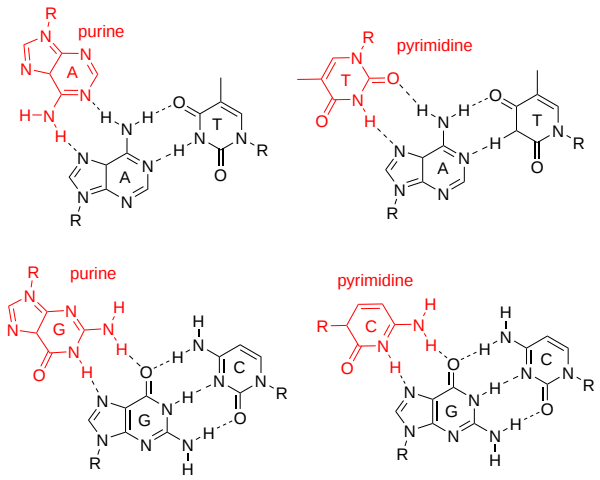


Figure 10.2.29: Base pairing in triple helix motifs. (after Jain et al. Biochimie. 2008. doi: 10.1016/j.biochi.2008.02.011

Figure 10.2.30 shows an interactive iCn3D model of a solution conformation of a parallel DNA triple helix (1BWG).

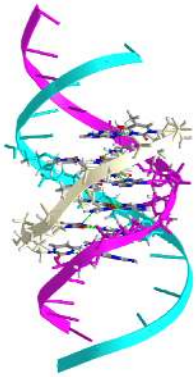


Figure 10.2.30: Solution conformation of a parallel DNA triple helix (1BWG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5JU813eNjND8E7

Triple helix formation can also occur within a single strand of DNA. The resulting structure is called **H-DNA**. An example is shown below. Note that the central blue, black, and red sequences are all mirror-image repeats (around a central nucleotide). During processes that unravel DNA (replication, transcription, repair), the self-association of individual mirror repeats can form a locally stable triple helix, as shown in Figure 10.2.31.

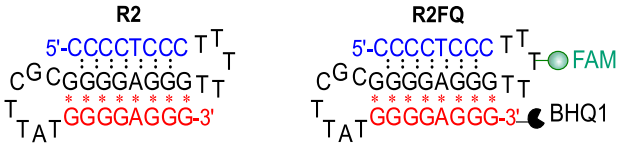


Figure 10.2.31: Schematic illustrations of (A) the H-DNA or intramolecular triplex structure used in this study; del Mundo et al. (2019) Nucleic acids research. 47. e73. 10.1093/nar/gkz237. Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/)

The * between in the G*G and A*A denote Hoogsteen hydrogen bonding (purine motifs) in this intramolecular triple helix. Reverse Hoogsteen hydrogen bonds can also occur.

Triple helices can develop when single-stranded DNA, formed during replication, transcription, or DNA repair with half of the required mirror symmetry, folds back into the adjacent major groove and base pairs, using Hoogsteen/reverse Hoogsteen bonding, which Mg^{2+} can stabilize.

Recent Updates: Four-Way Junctions

As we will see in the next section on RNA, nonhelical sections of DNA can bind small target molecules through noncovalent interactions. (RNA examples that we will see in the next chapter section include aptamers, ribozymes, and riboswitches.) One example is "lettuce" single-stranded DNA that can bind small fluorophores modeled after the intrinsic fluorophore of the green fluorescent protein. The fluorophore's fluorescence is dramatically enhanced when bound to the lettuce DNA. Figure 10.2.32 below shows the structure of extrinsic DNA fluorophores based on GFP that bind to the single-stranded "lettuce" DNA.

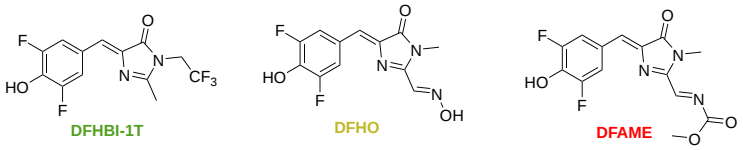


Figure 10.2.32: Structure of extrinsic DNA fluorophores based on GFP that bind to DNA. The font color of the names indicates the color of the emitted fluorescence.

Figure 10.2.33 shows an interactive iCn3D model of a solution conformation of a ssDNA:DFAME fluorophore complex (8FI0). The blue dotted lines show π-π stacking interactions, and the green dotted line a hydrogen bond.

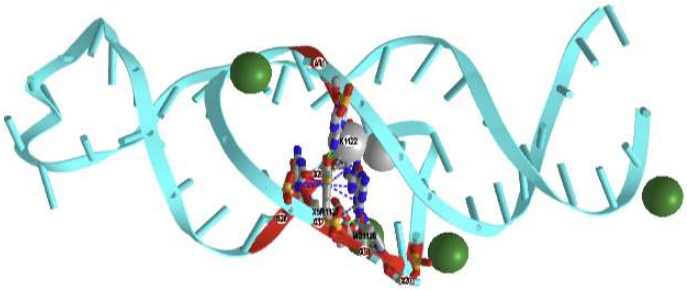


Figure 10.2.33: Solution conformation of a ss-53 mer DNA:DFAME fluorophore complex (8FI0). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...97eLqwNWmaTNC9

The DFAME ligand is shown as sticks.

Figure 10.2.34 shows a close-up of DFAME (colored spacefill) bound to the lettuce DNA.

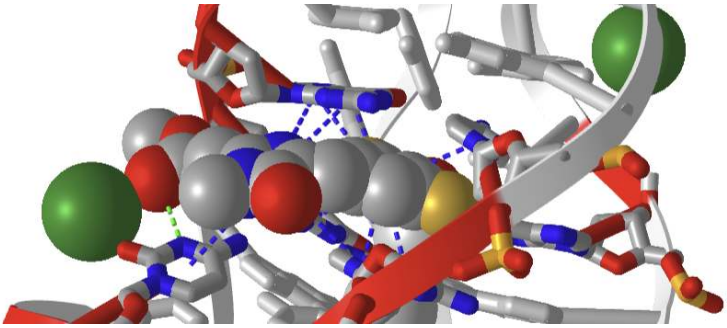


Figure 10.2.34: Pi stacking interactions (blue dotted lines) between the extrinsic fluorophore DFAME (spacefill) and ssDNA.

The DNA fold is characterized as a four-way junction (also seen in RNA, but they are more L or H-shaped). On either end are B-DNA duplexes, and the ssDNA between them forms stem-loops with odd base pairings in the stems. The overall structure is like a cloverleaf. Two coaxial stacks of nucleotides form a G-quadruplex where the fluorophore binds. Pi base stacking between diagonally packed bases and the binding of Mg^{2+} and K^{+} stabilize the structure.

10.2.6: Stability of nucleic acids

After looking at the myriad of structures showing the nearly parallel hydrogen-bonded base pairs and from ideas from most textbooks and classes you have taken, you probably think that double-stranded DNA is held together and stabilized by hydrogen bonds between the bases. It is well known that the greater the percentage of GC compared to AT, the greater the stability of the dsDNA. This translates into a higher "melting temperature (T_M)," the temperature at which the dsDNA is converted to ssDNA. There is a linear relationship between GC content and T_M. The figures above show that GC base pairs have three inter-base hydrogen bonds compared to 2 in AT base pairs. These observations support the notion that inter-base hydrogen bonds are the source of dsDNA stability.

You would be, in general, correct in this belief. Still, you'd be missing the more important contributor to ds-DNA stability, base (π) stacking, and the noncovalent interactions associated with the stacking. The main contributors to stability are hydrophobic interactions in the anhydrous hydrogen-bonded base pairs in the helix. Given that the hydrogen bond donors and acceptors that contribute to base pairing exist in the absence of competing water, the donors and acceptors are free to engage in bonding fully. The hydrogen bond interaction energy is more favorable in the stack. The stacking energy is similar for an AT-AT stack and a GC-GC stack (about -9.8 kcal/mol, 41 kJ/mol). Hence, the AT and GC base pairs contribute equally to stability. The excess stability of dsDNA enriched in GC base pairs can still be explained by the extra stabilization for an additional hydrogen bond per GC base pair

A myriad of interactions stabilizes proteins, but the folded state is marginally more stable than the ensemble of the unfolded state. Marginal stability is important as protein conformation often must be perturbed by binding and ensuing functions. The same must be true of double-stranded DNA, which must "unfold' or separate on replication, transcription, and repair. It is well known that

dsDNA structure is sensitive to hydration (see the section on A, B, and Z DNA). As we saw with proteins, small molecules like urea can also denature DNA into single strands.

DNA must be stable enough to carry genetic information but dynamic enough to allow events that require partial unfolding. Other water-soluble molecules like ethylene glycol ethers (polyethylene glycol-400) and diglyme (dimethyl ether of diethylene glycol), which are more hydrophobic than water, appear to reduce base stacking interactions while maintaining them and, at the same time, allow a longitudinal extension or breathing of the helix. This dynamic extension may be required for transitions of B-DNA to Z-DNA, for example. The extensions also allow transient "holes" to appear between base pairs, which might assist in binding intercalating agents like some transition metal complexes. The extension caused by these ethers and natural extensions would decrease base stacking but also strengthen the hydrogen bonding between bases.

Longitudinal helical extensions might be important when homologous genes recombine. In that process, the homologous DNA strand is exchanged with a paired homolog. This processing is associated with strand extension and disruption of base pairs at every third base. Recombination also must allow chain extension as it maintains base-pairing fidelity.

DNA structures get more complicated as they pack into the nucleus of a cell and form chromosomes, as shown in Figure **10.2.35**. We will study the packing of DNA in other sections.

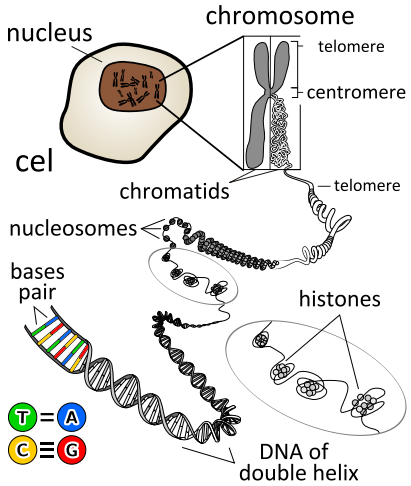


Figure **10.2.35** : Packing of DNA into the chromosome.
https://commons.wikimedia.org/wiki/F...omosome_en.svg
.
Creative Commons
Attribution 3.0 Unported

10.2.7: Summary

This chapter provides an in-depth overview of nucleic acids, one of the four fundamental classes of biological macromolecules, emphasizing their central role in storing and transmitting genetic information. It begins by contrasting DNA and RNA in terms of their chemical composition—highlighting that DNA contains deoxyribose and the bases adenine, guanine, cytosine, and thymine, while RNA uses ribose and replaces thymine with uracil. The discussion covers the nucleoside and nucleotide building blocks and explains how the sugar-phosphate backbone is constructed through the formation of phosphodiester bonds, emphasizing the directionality of nucleic acid synthesis (5′ to 3′).

The chapter then explores the structural organization of DNA and RNA. DNA's iconic double helix, stabilized by Watson–Crick base pairing and base-stacking interactions, is examined along with its major and minor grooves, which serve as critical recognition sites for proteins. In contrast, RNA, though generally single-stranded, adopts complex secondary and tertiary structures that enable diverse functions beyond mere information storage.

Additionally, alternative base pairing modes—including reverse Watson–Crick, wobble, Hoogsteen, and reverse Hoogsteen interactions—are discussed, illustrating how nucleic acids can form noncanonical structures that influence replication,

transcription, and repair. The synthesis and dynamic behavior of nucleic acids are also addressed, including the mechanism of nucleotide incorporation by polymerases and the energetic contributions from pyrophosphate hydrolysis. Finally, the chapter contextualizes these structural features within the historical framework of nucleic acid research, noting landmark discoveries and the evolution of our understanding of DNA and RNA structures.

This comprehensive review lays the foundation for further exploration into the functional roles of nucleic acids in cellular processes, highlighting both the simplicity of their repeating structures and the complexity arising from their dynamic conformations and interactions.

10.2.8: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics*. Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore*. Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

This page titled 10.2: Nucleic Acids - Structure and Function is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **8.1: Nucleic Acids - Structure and Function** by Henry Jakubowski and Patricia Flatt is licensed CC BY-SA 4.0.
- **5.7: Binding - Enzyme Linked Immunosorbant Assays (ELISAs)** by Henry Jakubowski and Patricia Flatt has no license indicated.

10.3: Nucleic Acids - RNA Structure and Function

Learning Goals (ChatGPT o1, 1/30/25)

1. **Differentiate RNA from DNA:**
   - Describe the key chemical differences (e.g., ribose vs. deoxyribose; uracil vs. thymine) and explain how these differences influence the biological functions of RNA.
2. **Analyze RNA Structural Complexity:**
   - Explain how RNA can form secondary (e.g., stem-loop, hairpin) and tertiary structures, and evaluate the role of noncanonical base pairing and chemical modifications in these processes.
3. **Understand RNA Transcription and Processing:**
   - Illustrate the process of transcription from DNA to RNA, including the formation of heteronuclear RNA and subsequent splicing events that yield mature mRNA.
4. **Classify RNA Types Based on Function and Size:**
   - Distinguish between coding RNAs (e.g., mRNA) and noncoding RNAs, and further classify noncoding RNAs into short (<200 nt) and long (>200 nt) categories with examples such as tRNA, rRNA, miRNA, snoRNA, lncRNA, and circRNA.
5. **Evaluate the Role of RNA in Protein Synthesis:**
   - Describe the functions of messenger RNA in coding for proteins and the roles of transfer RNA and ribosomal RNA in translation, including the molecular interactions that facilitate these processes.
6. **Investigate RNA–Protein Interactions:**
   - Analyze specific examples of RNA–protein complexes (e.g., ribosome components, spliceosomal snRNPs, toxin-antitoxin systems) to understand how RNA structure and binding partners influence cellular functions.
7. **Explore RNA-Mediated Gene Regulation:**
   - Explain how noncoding RNAs such as microRNAs and siRNAs regulate gene expression through mechanisms like mRNA degradation, translational inhibition, or transcriptional modulation.
8. **Apply Thermodynamic Concepts to RNA Folding:**
   - Interpret the concept of RNA's thermodynamic folding landscape and compare it to protein folding, discussing how environmental factors and protein binding can lead to dynamic structural rearrangements.
9. **Examine RNA's Role in Cellular and Disease Contexts:**
   - Discuss the significance of RNA molecules in cellular processes (e.g., splicing, ribosome function, synaptic regulation) and their emerging roles in diseases such as cancer, cardiovascular disorders, and neurological conditions.
10. **Integrate Bioinformatics and Structural Tools:**
    - Utilize available computational programs and databases (e.g., lncipedia, miRBase, CircAtlas) to predict RNA secondary structures and explore experimentally determined RNA–protein complexes via interactive 3D models.

Each goal is designed to promote critical thinking, link structure to function, and encourage practical engagement with current research tools in RNA biology.

10.3.1: RNA: Structure and Function

Ribonucleic acids are very similar in chemical structure to DNA, except they contain ribose instead of deoxyribose. They also have the pyrimidine base uracil instead of thymine, as shown in Figures 1 and 2 above. These two small changes (but mostly the first) confer on it a very different set of biological functions than DNA. This should not surprise us, and the basis of all chemistry and biochemistry is that chemical structure determines chemical and biochemical functions and activities. In the previous section, we

discussed how RNA can adopt complex tertiary structures, which require the presence of more noncanonical base pairs and chemical modification of bases. In this section, we will explore the various types of RNA structures and their functions.

The sequence of RNA is made from DNA through a process called transcription (converting the information of DNA, a nucleic acid, into RNA, another nucleic acid). RNA can form double-stranded helices, but typically, these are viral in origin. dsRNA is a pathogen-associated molecular pattern (PAMP) that binds Toll-like receptor 3 (TLR3), as seen in Chapter 5.5. If both strands of DNA are transcribed, the resulting strands can anneal to form dsRNA. In addition, a single strand of RNA can fold on itself if the 5' and 3' ends are complementary to form a stem-hairpin loop. Figure **10.3.1** shows a stem-loop from a messenger RNA (4QOZ) when it is bound to a specific RNA-binding protein (not shown).

Figure **10.3.1** A stem-loop from a messenger RNA (4QOZ) when it is bound to a specific RNA-binding protein (not shown).

Larger ssRNA can form tertiary structures with many regions of intrastrand hydrogen bonds forming secondary structures, as shown in Figure **10.3.2** for one type of RNA called a transfer RNA

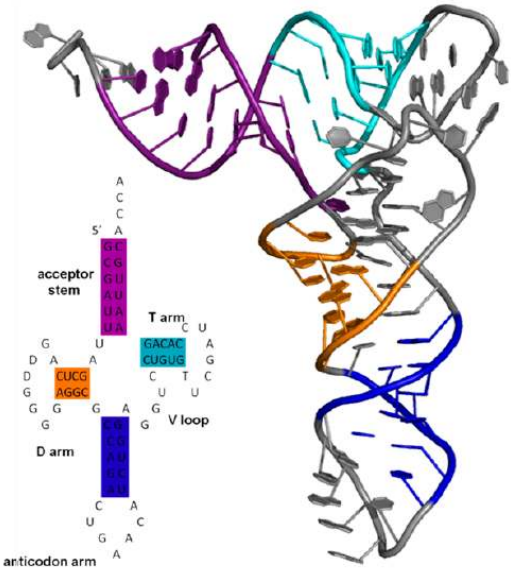


Figure **10.3.2**: Secondary and tertiary structure of tRNA showing unpaired (gray) and hydrogen-bonded secondary structure regions (color). Kyle Schneider https://commons.wikimedia.org/w/inde...curid=12309266

Figure **10.3.3** a computed model for secondary structure within a much larger single RNA molecule S11 that is part of the ribosome. The figure shows color-coded differences in accessibility when the S11 RNA is free (blue) and bound to the protein NSP2 (red), which induces structural rearrangements.

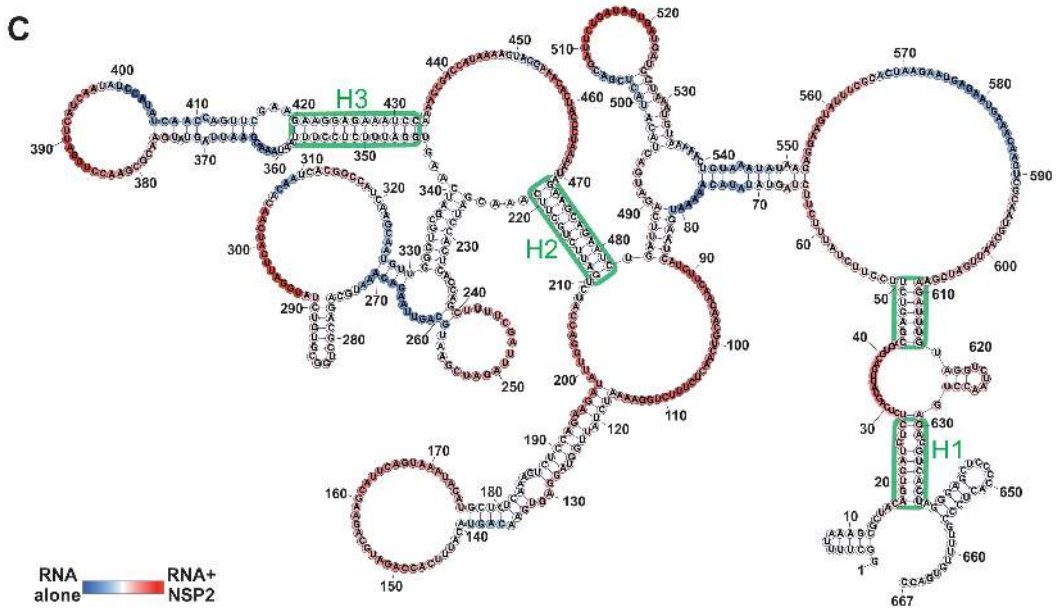


Figure **10.3.3**: Model showing secondary structure on ribosomal S11 RNA and accessibility changes on binding the protein NSP2.

You can imagine that a different set of intrachain H-bonded double-stranded regions could easily form, with the most likely determined by sequence, local environment, and protein binding partners. Each RNA molecule would have a thermodynamic folding landscape similar to that of a protein. Programs are available to determine secondary structures from RNA sequences. Each RNA molecule would have a thermodynamic folding landscape similar to that of a protein. Also, the structures are dynamic, as we can see with proteins.

Another feature that complicates RNA is that many different types of RNA are made from DNA using RNA polymerases. They are loosely divided into two types of RNA. One is **coding RNA**, which contains the sequence information that will be translated into a protein sequence. The other type is called **noncoding RNA**. These RNAs regulate many cellular processes, including transcription, which produces coding RNA.

10.3.1.1: Coding RNA

The DNA template from which the coding sequence of a translatable RNA is produced is called a **gene.** The coding RNA, which has the exact sequence that is translated into protein, is called **messenger RNA (mRNA)**. The exact sequence of RNA in mRNA that encodes a protein is derived from a longer contiguous DNA sequence in the nucleus, from which sections called **intervening sequences** or **introns** have been removed. The coding sequences of DNA, which are separated by introns, are called **exons**. When coding DNA is transcribed, a long contiguous sequence containing both exons and introns is transcribed into one long primary transcript called **heteronuclear RNA**. The introns in the heteronuclear RNA are removed in a splicing reaction catalyzed by a large complex called the **spliceosome** to form mRNA. The process is illustrated in Figure **10.3.4**. The first RNA sequence made is the heteronuclear RNA.

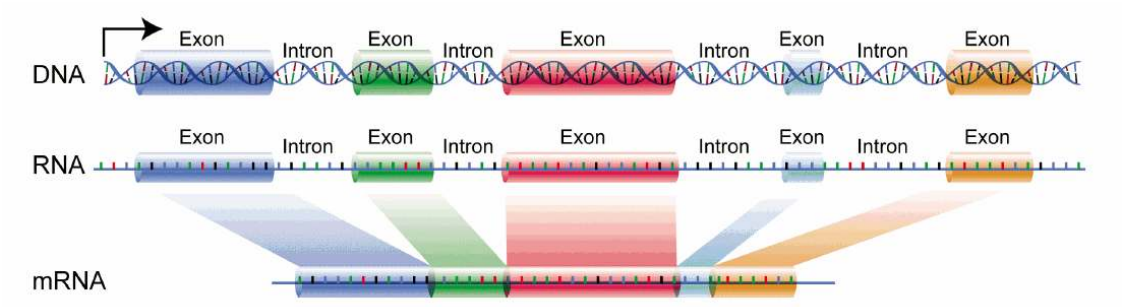


Figure **10.3.4**: https://commons.wikimedia.org/wiki/F...ns_introns.gif. This work is in the **public domain** in the United States

The long single-stranded mRNA molecule binds to ribosomes and nanomachines, which orchestrate the translation of the mRNA sequence into a protein sequence. Around 20,000 human genes produce an even larger number of mRNA that arise from differential splicing of the primary transcript.

---

mRNA molecules have a polyA sequence connected to the 3' end (see Chapter 25.2: RNA Processing for more details). PolyA binding proteins (PABP) interact with the polyA tail and are likely stabilize the RNA molecule. A PABP domain was fused to another protein called a major vault protein (MVP). The MVP (100K molecular mass) forms a large multimeric "vault" with a large internal volume that can encapsulate large molecules (A Cellular 'vaults' deployed to spy on gene activity), such as RNAs. The assembled vault may acts as a scaffold for proteins involved in signal transduction and in transport between the nucleus and cytoplasm. During assembly, the vault captures mRNA molecules produced by the cell for up to a week, enabling real-time analysis of transcription. Table **10.3.1** below shows two iCn3D models of oligomeric vaults created by the MVP

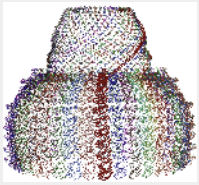
Human Vault Cage (9BW5) Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...0508a84051ead8	
Rat Half vault structure (7PKY) Click the image for a popup or use this external link:https://www.ncbi.nlm.nih.gov/Structu...23ab413f7914d9 (long load time)	

Table **10.3.1**: iCn3D models of oligomeric vaults created by the MVP. The top is the full vault and the bottom the half vault. One monomer in the half vault is shown in spacefill (brown).

Pablo Guerra *et al*. Symmetry disruption commits vault particles to disassembly. *Sci. Adv.***8**,eabj7795(2022). DOI:10.1126/sciadv.abj7795

Half-vaults (39mers) from oligomers of MVP then assemble into the full vault (78mer). In mammals, the 193-kDa vault poly(adenosine 5′-diphosphate) ribose polymerase (VPARP) protein, the 290-kDa telomerase-associated protein 1 (TEP1), or small untranslated RNAs can be found inside.

10.3.1.2: Noncoding RNA (ncRNA)

Not long ago, few thought about possible RNA transcripts from non-coding regions of the genome, except for two types of RNA required to translate mRNA. These two are ribosomal RNAs (rRNA) found in ribosomes and transfer RNAs, to which amino acids are esterified and transferred to a growing protein chain of the ribosome. Many more classes have been discovered and given names that confuse those more familiar with protein structures. One way to classify noncoding RNAs (ncRNAs) is based on size.

- short noncoding RNAs (sncRNAs) are <200 nucleotides
- long noncoding RNAs (lncRNAs)are >200 nucleotides

These function to regulate gene expression at both the transcription and post-transcriptional levels. Some have catalytic functions. Some affect chromosome structure and chemical modification.

10.3.1.2.1: Long Noncoding RNAs (lncRNAs)

There may be between 16,000 to over 100,000 human lncRNAs encoded into the genome, which adds much complexity to our understanding of the function of RNA transcripts. An online lncipedia is a database of searchable lncRNA sequences. There are many types of lncRNAs. The first we will consider is ribosomal RNA.

a. Ribosomal RNA (rRNA):

These RNAs fit the simple definition of lnRNAs (>200 nucleotides and are not protein-coding), but most would not think of them as lncRNAs since they have always been in their own category of a nonprotein-coding gene. rRNAs vary in length from between 1500 and 3000 nucleotides in bacteria and about 1800 and 5000 nucleotides in humans and are the core structure of ribosomes. These nanomachines translate bound mRNA into a protein sequence.

Figure 10.3.5 shows an interactive iCn3D model of the structure of 23S rRNA of the large ribosomal subunit from Deinococcus radiodurans (2O44) (long load time).

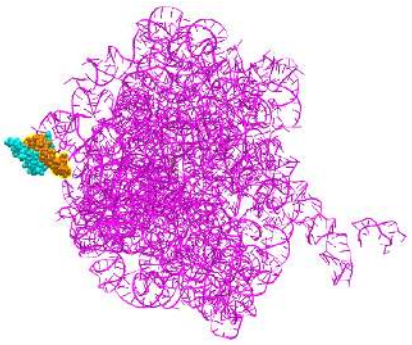


Figure 10.3.5: Structure of 23S rRNA of the large ribosomal subunit from Deinococcus radiodurans (2O44). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Tb2BVwSTe6LviPou5

The red (highlighted yellow) spacefill is the 5' start of the rRNA. The chain has a complex tertiary structure, like a protein sequence, and ends at the cyan spacefilling 3' end. It has 2880 nucleotides.

b. Other types

Let's focus on more classic examples of long noncoding RNAs (i.e., not rRNA). One way to categorize them is based on the position in the genome that encodes them. The different types include long **intergenic** noncoding RNAs (lincRNAs), intronic lncRNAs, antisense RNAs (as lncRNAs), and other variants. These are illustrated in Figure 10.3.6, where the lncRNA is pink.

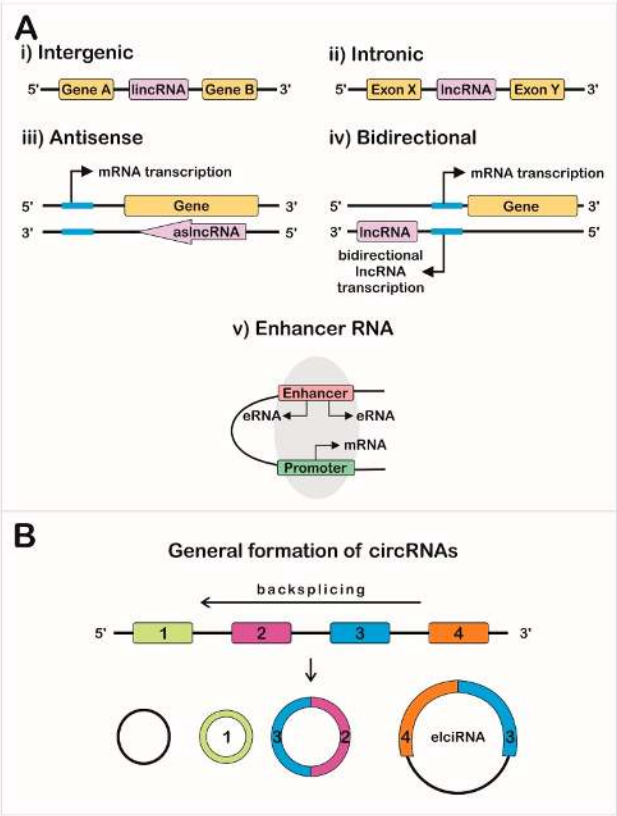


Figure \(\PageIndex{6\): https://en.Wikipedia.org/wiki/Long_non-coding_RNA. Creative Commons Attribution-Share Alike 3.0 Unported

Another variant is exon and intron-containing circRNAs (EIciRNAs), as illustrated in panel B in Figure 10.3.13. These are presumably produced from pre-mRNA for a given mRNA and appear to regulate gene expression through RNA-RNA interactions with U1 snRNA, which starts the assembly of the spliceosome on pre-mRNA when it binds to the 5′ pre-mRNA splice site.

circRNAs are found throughout the biological world and have been associated with diseases such as cancer, cardiovascular disease, and brain disorders like Alzheimer's. Their possible functions include binding to miRNAs and RNA-binding proteins, altering their activities, and binding to DNA, regulating its transcription. CircAtlas is a database of circRNAs that two techniques have experimentally validated

Now let's consider specific examples of long non-coding RNAs (lncRNAs), which are often bound to target proteins.

- mamRNA (a lnc RNA)

The lncRNA named mamRNA (**M**mi1 **a**nd **M**ei2-associated RNA) binds two proteins, Mmi1 and Mei2, in *Schizosaccharomyces pombe* that control the balance between meiosis and mitosis in yeast. (*Schizosaccharomyces pombe* is a "fission" yeast that divides by fission, not budding.) The MamRNA has two variants of length 550 and 700 nucleotides. The binding of mamRNA leads to the ubiquitinylation of Mei2 in the complex. Mmi1 R is an RNA-binding protein that binds to a modified version of adenosine methylated at N6 and is found internally in mRNA. Mei2 (meiosis protein 2) is necessary of meiosis. The binding of mamRNA leads to the ubiquitinylation of Mei2 in the complex, targeting it for proteolysis. Mei2 concentrations increase relatively, shifting yeast from mitosis to meiosis. Figure 10.3.7 shows a cartoon depicting these interactions.

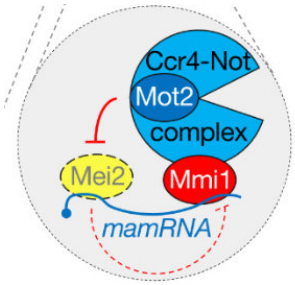


Figure 10.3.7: Binding of the lncRNA mam2 to yeast proteins Mei2 and Mmi1. Andric, V. et al. Yeast. Non-coding RNA 2021, 7, 34. https://doi.org/10.3390/ncrna7020034. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Figure 10.3.8 shows an interactive iCn3D model of the S. pombe Mei2 RRM3 protein domain bound to the Mei2 binding "domain" of mamRNA (6YYM), which in this structure is only eight nucleotides long (not the full length of this lncRNA, which is 550 and 700 nucleotides long).

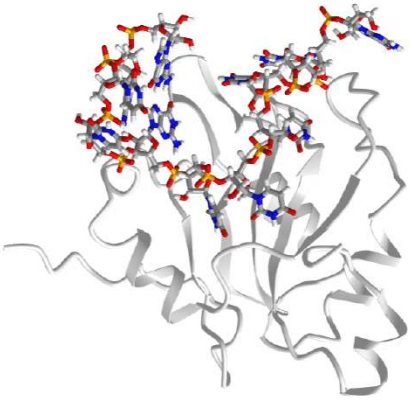


Figure 10.3.8: S. pombe Mei2 RRM3 protein domain bound to an eight nucleotide section of the yeast lncRNA mam (6YYM). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...5xmQjwAmQL9on7

- ToxI - a lncRNA inhibitor of the endonuclease ToxN

Those with a more chemistry-centric background might be surprised that viruses also "infect" bacteria. These viruses are called **bacteriophages**. It is estimated that there are over 10^{30} in nature. Some covalently incorporate into genomes, where they reside and are incorporated permanently. They are a main driver of bacterial genome evolution as they shape the bacteria's immune response and adaptation.

One very interesting example is the type III toxin-antitoxin (TA) system in E. Coli. It consists of a toxin, ToxN, which is a nuclease that cleaves internally after the second A in a AAA sequence. It acts on mRNA, but especially pre-mRNA sequences. It is inhibited by the binding of a lncRNA called ToxI (toxin inhibitor). The RNA sequence of the ToxI inhibitor has 36 "domain" repeats of a pseudoknot, one of which is sufficient to inhibit the ToxN. The ToxN endonuclease cleaves the ToxI lncRNA as it assembles the complex. It also cleaves its mRNA. Figure 10.3.9 shows an interactive iCn3D model of the protein toxin (ToxN):lncRNA (ToxI), which is a shortened version of 29 29-nucleotide section from *Pectobacterium atrosepticum* (**2xdb**).

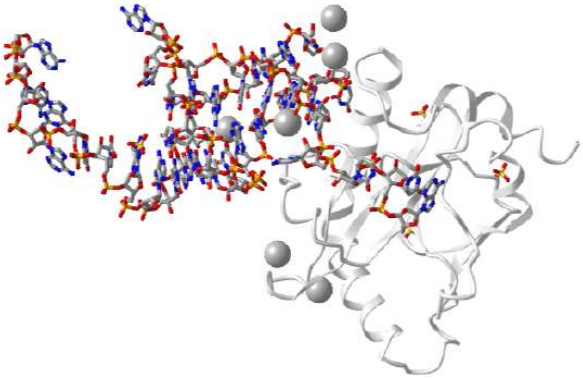


Figure 10.3.9: Protein toxin (ToxN) - 29 NT fragment of lncRNA antitoxin ToxI complex (**2xdb**). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KjNPGghiUyAXr9

As we discussed in Chapter 3.2: The Structure of Proteins- An Overview, some long noncoding RNAs have noncanonical open reading frames (ncORFs) that are translated into protein, adding an element to the complexity of RNA.

Recent Updates: January 18, 2026

New types of long noncoding RNAs are routinely discovered. Some are termed **Natural RNA-Only Assemblies**. These are often found in bacteria and viruses that infect them (bacteriophages). Some examples of iCn3D structures are shown in Table 10.3.2 below.

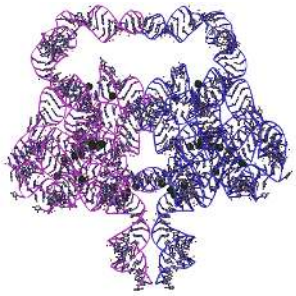
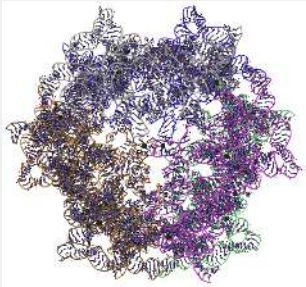
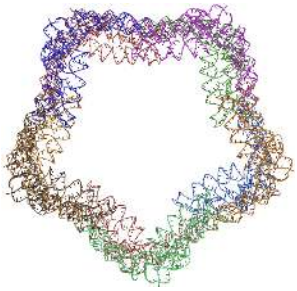
Ornate Large Extremophilic (OLE) RNAs *(pdb_00009lcr)* Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...d35637ad01dfd1	
ROOL (Rumen-Originating, Ornate, Large) RNA hexamer from bacteria found in cow stomachs (*pdb_00009j6y)* Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...0f5f943d312af6	
GOLLD (Giant, Ornate, Lake- and Lactobacillales-Derived) RNAs from bacteria 10mer (*pdb_00009l0r)* Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...9c2b7ccc40d727	

Table 10.3.2: iCn3D structures of Natural RNA-Only Assemblies

ROOL and GOLLD structures form cages with internal spaces. For more information, visit this PDP 101 page on Natural RNA-Only Assemblies by Janet Iwasa.

10.3.1.2.2: Short Noncoding RNA

Short noncoding RNAs (sncRNAs) are less than 200 nucleotides in length. By definition, this would include **transfer RNAs (tRNAs),** which bring to the ribosome amino acids covalently attached to their 3' end of the tRNA for incorporation into a growing protein chain during the translation of mRNA. As with rRNA for lncRNAs, these are really in a class of their own. Others include **small nuclear RNAs (snRNAs)** involved in splicing, **small nucleolar RNAs (snoRNAs)** involved in the modification of rRNAs, and **microRNAs (miRNAs)**, involved in the inhibition of translation and transcription, **PIWI-interacting RNAs (piRNAs)**, and endogenous **small interfering RNAs (siRNAs).** It is difficult to remember the subtle differences among these, which makes them hard to understand. We will tell their stories with a few targeted examples.

a. Transfer RNA:

Transfer RNAs act as adapter molecules between transcription and translation. They are between 76 and 90 nucleotides long and have a cloverleaf shape. An enzyme, aminoacyl-tRNA synthase, covalently attaches a select amino acid at its 3' end. Another end of the tRNA hydrogen bonds through 3 nucleotides (the anticodon) to a triplet nucleotide (the codon) on the mRNA that encodes a specific amino acid at that triplet position. Figure **10.3.10** shows an interactive iCn3D model of the structure of yeast phenylalanine tRNA (1EHZ)

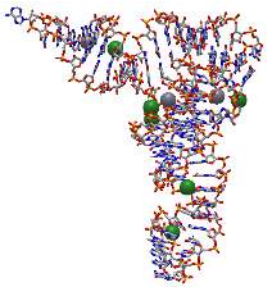


Figure **10.3.10**: Yeast phenylalanine tRNA (1EHZ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?iAWpKknDWbQm5hHL6

b. Small nuclear RNA (snRNA):

The spliceosome is a nanoparticle that catalyzes the removal of introns from pre-mRNA in eukaryotes (prokaryotes appear devoid of introns). The yeast spliceosome has a molecular weight of 1.3 million and contains five small ribonucleoproteins (RNPs) with many other associated proteins. Each of the 5 RNPs has a small nuclear RNA (U1, U2, U4, U5, and U6) enriched in uracils. U6 is highly conserved and is directly involved in catalysis. Figure **10.3.11** shows an interactive iCn3D model of the core structure of the U6 small nuclear ribonucleoprotein complex with most of the U6 RNA bound.

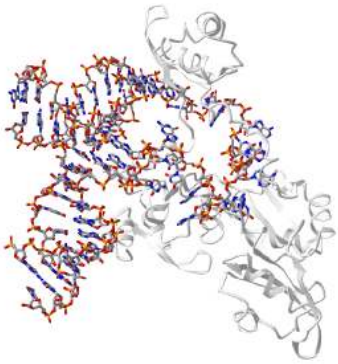


Figure **10.3.11**: Core structure of the U6 small nuclear RNA - protein (ribonucleoprotein) complex (4N0T). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...P3SLVRAKhP7W57

c. MicroRNAs (miRNAs) and small inhibitory RNAs (siRNAs)

MicroRNAs (miRNAs) control the expression of thousands of genes in plants and animals. They are single-stranded but fold on themselves to form a stem-hairpin. The miRBase is a microRNA database containing almost 40,000 miRNA sequences. miRNAs are highly conserved and are found in animals, plants, and some unicellular eukaryotes. They interact with the 3′ untranslated regions of mRNAs and inhibit or prevent their translation. Several key proteins, **RNA polymerase II, Drosha, and Dicer** are involved in the canonical pathway. The others appear to be independent of **Drosha**, which is a ribonuclease III double-stranded (ds) RNA endoribonuclease.

Dicer is a dsRNA) endoribonuclease, which cleaves long dsRNAs and short hairpin pre-microRNAs (miRNA) into fragments of either 21-23 nucleotides (short interfering RNA) or 19-25 nucleotides (microRNAs). Each has two nucleotides that are unpaired at the 3' end. These bind to the enzyme complex RISC (RNA-induced silencing complex), which then targets them to mRNA complementary to the siRNA/miRNA (RISC), causing cleavage of the mRNA and hence inhibiting translation.

Small inhibitory RNAs (siRNAs) are very similar to miRNAs (to the point that differentiating between them is somewhat arbitrary). They both engage in RNA interference (RNAi) of mRNA translation. Here are some reported differences:

- The substrate for dicer cleavage is dsRNA (that could be added exogenously) of length 30-100+ for siRNA, but the actual pre-miRNA of length 7-100 nucleotides that may contain hairpins with some mismatches (i.e., not a perfect stem and hairpin)
- The final RNA after dicer processing is double-stranded for both and 21-23 nucleotides long for siRNA and 19-25 for miRNA
- siRNAs are perfectly complementary to the target mRNA, while miRNAs, which are not necessarily perfectly complementary, typically bind to the 3' untranslated end of the mRNA
- Because of the perfect complementarity to target mRNA, siRNA interacts with only one mRN,A while miRNAs, given that they are not perfectly complementary to their target sequences, can bind different mRNAs
- Given their higher affinity binding, the siRNA leads to dicer endonuclease cleavage of the target mRNA. In contrast, inhibition of mRNA translation by miRNAs arises from binding of the miRNA to the mRNA or, if the match between the miRNA and mRNA is high enough, endonuclease cleavage of the mRNA.

Figure **10.3.12** shows canonical and several alternative pathways for their transcription and processing from the noncoding miRNA genes.

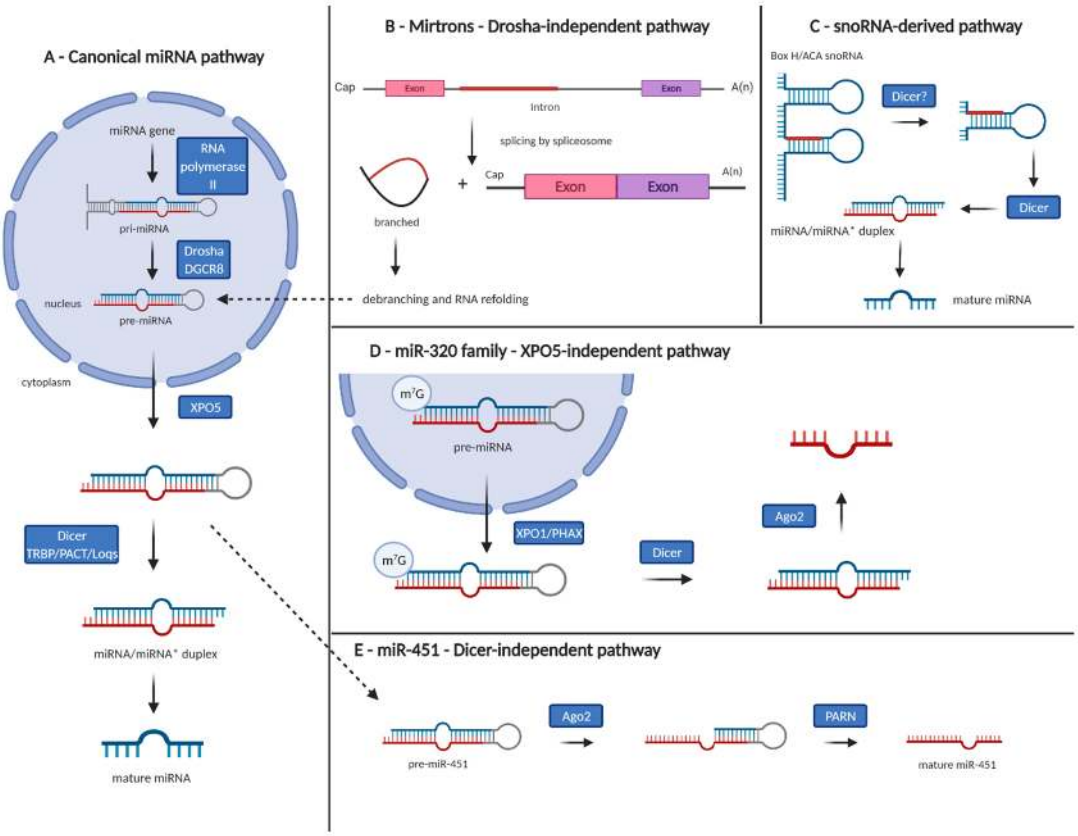


Figure **10.3.12**:

Figure 1. Canonical and non-canonical pathways of microRNA biogenesis. (A) Canonical pathway—microRNA gene is transcribed by RNA polymerase II into primary microRNA (pri-miRNA), cleaved by microprocessor complex Drosha/DGCR8, and precursor microRNA (pre-miRNA) is exported from the nucleus to the cytoplasm by Exportin 5 (XPO5) and further processed by Dicer and its partners into 18–25 nucleotide long microRNA duplex with 2-nucleotide 30 overhangs. The guide strand is subsequently bound by the Argonaute proteins 1-4 (AGO1-4) and retained in the microRNA-induced silencing complex to target mRNAs for post-transcriptional silencing. (B) Mirtrons—generated through mRNA splicing independently of the Drosha-mediated processing step. (C) Small nucleolar RNA-derived microRNAs—Drosha-independent pathway. (D) Exportin 5-independent transport of pre-miRNAs from the nucleus to the cytoplasm has been described in the case of miR-320 family. (E) Dicer-independent processing of miR-451—pre-miR-451 is directly loaded into AGO2, cleaved, and trimmed by poly(A)-specific ribonuclease PARN to produce mature miR-451. Gregorova et al. Cancers 2021, 13, 1333. https://doi.org/10.3390/cancers13061333. Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/)

siRNA that are perfect matches to specific mRNA can be easily designed and purchased for translation inhibition and through gene silencing studies. Both miRNAs and siRNAs are potentially therapeutic as it is much simpler to design a drug that targets a mRNA sequence (a 1D sequence target) than a protein active site (a 3D target). Additionally, they can be used to inhibit protein synthesis of target proteins that don't have a "druggable" active site.

The protein Argonaute is involved in miRNA and siRNA silencing of genes through their mRNAs in the RISC (RNA-induced silencing complex). RISC contains the protein argonaute 2 (AGO2) bound to a "**guide**" RNA, which is either microRNA (miRNA) or short interfering RNA (siRNA). The miRNA or siRNA directly interacts with the "target" - the mRNA.

- *Example miRNA:*

Figure **10.3.13** shows an interactive iCn3D model of human Argonaute2 Bound to a Guide (miRNA) and Target RNA (4W5O). The two RNA sequences are 5' UUCACAUUGCCCAAGUCUUU 3' and 5' CAAUGUGAAA 3'.

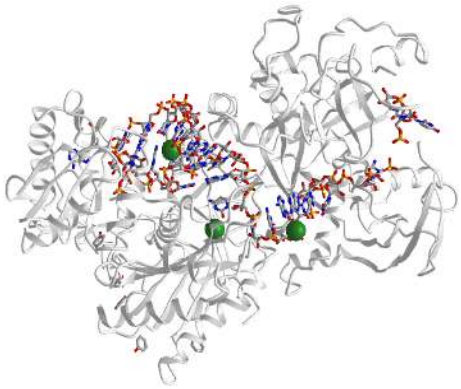


Figure 10.3.13: human Argonaute2 Bound to a Guide and Target RNA (miRNA) (4W5O). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?9fvyQmPkk4Kctei59

Recent Updates: October 9, 2024

2024 Nobel Prize in Physiology and Medicine for discovery of miRNAs

Victor Ambros and Gary Ruvkun won the Nobel Prize (October 7, 2024) for discovering miRNA and its function in the roundworm *C. elegans*, which has fewer than 1000 cells. Their basic research on miRNAs' role in protein synthesis regulation was extended to other organisms. Humans express over 1000 miRNAs that regulate translation.

Ruvkun found that two small RNAs, lin-4 (22 nucleotides) and let-7 (21 nucleotides), are required to progress from larval to adult stages in the worm. They were nonhomologous but complementary to parts of the 3'-untranslated region (UTR) of mRNAs for proteins that are downregulated in development. Let-4 is found in many animal species, including humans, and binds to the 3'-UTR of the mRNA for the protein lin-41. Ambrose discovered that a small RNA lin-4 from a non-coding gene was complementary to regions in the 3'-UTR of the mRNA for a protein lin-14.

The seed sequences of miRNAs are a conserved 7-mer at positions 2-7 from the miRNA 5'-end. This part of the miRNA must be exactly complementary to the 3'-UTR of the target mRNA, but the other bases don't have to match exactly.

The sequence of let-7 miRNA from C. elegans is 5'-UGAGGUAGUAGGUUGUAUAGU-3'. Figure 10.3.*i* below shows the aligned sequences of a let-7 miRNA and lin-41 mRNA (A) and analogs made for NMR structure determination (B,C).

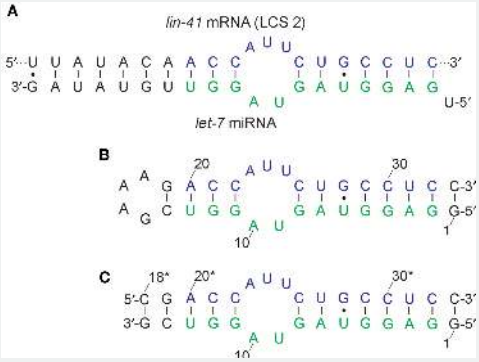


Figure 10.3.*i*: Sequences of a let-7 miRNA and lin-41 mRNA (A) and analogs.

(**A**) Schematic representation of the complex of *let-7* miRNA with 3′-UTR of the *lin-41* mRNA (LCS 2). The residues in green and blue match the sequences used in the monomolecular and dimeric constructs. (**B**) The 33-nt monomolecular RNA construct mimicking the complex. (**C**) The dimeric RNA construct. The numbering used in (B) has been preserved to facilitate comparing the two constructs. The residues originating from the *lin-41* mRNA sequence are labeled by asterisk. *Cevec, M, Thibaudeau, C and Plavec, J. Nucleic Acids Research*, Volume 36, Issue 7, 1 April 2008, Pages 2330–2337, https://doi.org/10.1093/nar/gkn088. Creative Commons CC-BY-NC license.

Two analog structures preserved most of the matching sequences from the miRNA and mRNA (green for the let-7 miRNA and blue for the lin-41 mRNA. The seed sequence of the miRNA was completely preserved. One was a single strand that formed a stem-loop structure. Figure **10.3.*j*** shows an interactive iCn3D model of the solution structure of a let-7 miRNA:lin-41 mRNA complex from C. elegans (2JXV) using the single-stranded analog.

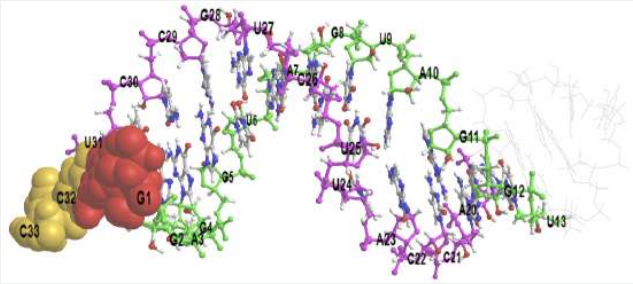


Figure **10.3.*j***: solution structure of a let-7 miRNA:lin-41 mRNA complex from C. elegans (2JXV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...xJh4paMJRiTQ47

G1 (red spacefill) is the 5'-end and C33 (yellow) is the 3'-end of the single-stranded analog. The green bases match those from the figure above. For clarity, the matching base partners of lin-41 mRNA are shown in cyan instead of blue.

The model shows two stems with an asymmetric loop containing 3 Us and 2 A. A GU wobble base is seen in the first stem between U6 and G28.

- *Example: siRNA (Small interfering RNA)*

Virus genomes are ultimately decoded into new viruses by the host's replication, transcription, and translation machinery. Host cells have evolved ways to silence mRNAs from viruses. Unfortunately, viruses, in response, evolve ways to suppress host RNA silencing. Many viral proteins are used by the host to suppress silencing. The viral p19 protein preferentially binds to host short interfering RNAs (siRNAs) rather than to microRNAs (miRNAs). A single mutation in the viral p19 proteins changes selectivity, allowing it to bind to a specific human miRNA called miR-122. This shows the subtle complexities of protein:RNA interactions. Figure **10.3.14** shows an interactive iCn3D model of the viral suppressor of RNA silencing protein and a 21-residue small interfering RNA (6BJV)

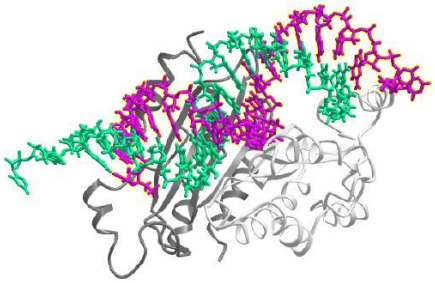


Figure **10.3.14**: Viral suppressor of RNA silencing protein p19 mutants and a double-stranded 21 nucleotide small interfering RNAs (6BJV). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VXzSgezTEPcPt9

- *Example: piRNA (a specific miRNA)*

Piwi proteins are RNA-binding proteins in plants and animals and are structurally similar to argonaute. They bind a guide RNA called piwi-interacting RNAs (piRNAs) and lead to the silencing of sequences called transposable elements that can move around the genome. piWi has endonuclease activity and can cleave mRNA. The piRNAs are just one type of miRNA. Figure **10.3.15** shows an interactive iCn3D model of Ephydatia fluviatilis (a sponge) PiwiA with a guide (piRNA) and-target RNA(7KX9)

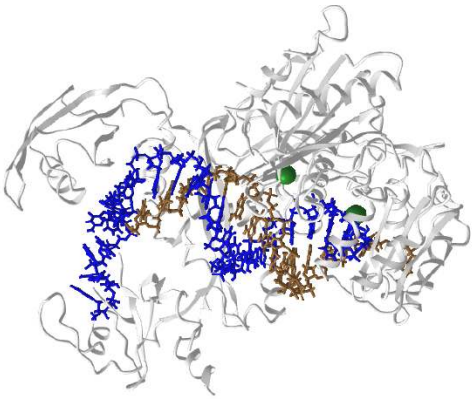


Figure 10.3.15: Ephydatia fluviatilis PiwiA-piRNA-target complex (7KX9). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?j9ikApUgsVvqsiM5A

lncRNAs and miRNAs in the brain

After the decoding of the human genome, many have been struggling to understand how the complexity of the human brain (large size, greater connectivity among neurons) arises, given that we appear to have only around 20,000 protein genes encoded by the genome (not counting small proteins of less than 100 amino acids). Long noncoding RNAs (lncRNAs) and miRNAs appear to be significant pieces of this puzzle. Their ability to regulate transcription during development may hold the key. Additional roles on these RNAs outside of transcriptional regulation are being discovered. Some are transported away from the nucleus to serve other functions in axons, dendrites, etc. For example, the lncRNA Gm38257 binds to proteins that structure the synapse (components of the spectrin/ankyrin complex) instead of simply regulating gene expression.

The repertoire of miRNA appears to be significantly increased in "intelligent" organisms such as humans and octopi. For example, a large increase (179) in miRNAs occurs in the evolutionary scale from mice (which have about 24,000 protein-encoding genes) to humans (around 20,000). miRNAs and lncRNAs may be involved (causative or correlative?) with brain disease. An example is miR-124, which is significantly elevated (3.5X) in hippocampal cells from mouse models of Alzheimer's compared to normal mice. Altered expression of the lncRNA named Gomafu, RNCR2 or MIAT) appears to affect certain psychiatric diseases.

c. small nucleolar RNA

The nucleolus is a small nuclear structure that helps assemble the ribosomal RNAs synthesized in the nucleus. They are then transported through the nuclear membrane into the cytoplasm, where they combine with proteins translated from mRNA in the cytoplasm to form complete ribosomes. As described below, rRNA is chemically modified by enzymes (much like the post-translational modification of proteins). One such modification is 2'-O-methylation in archaea and eukaryotes. A class of small nucleolar RNAs (snoRNAs) that vary from 10-21 base pairs is called C/D RNAs, and they "guide" the modification. Hence, they are also called "guide" RNAs. These snoRNAs bind to 3-4 proteins into ribonucleoproteins. Figure 10.3.16 shows an interactive iCn3D model of the box C/D ribonucleoprotein 40 nt snoRNA "guide" and a 10-nucleotide RNA target substrate. It appears that the maximal duplex RNA formed (from the guide and target) is 10 base pairs long.

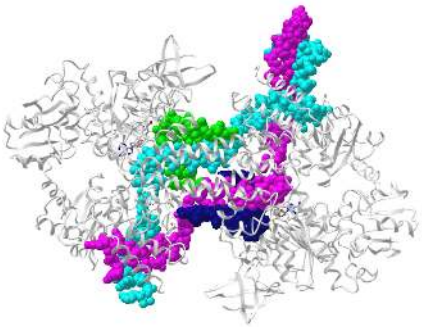


Figure 10.3.16: Box C/D ribonucleoprotein with a small nucleolar RNA (snoRNA) 12 nucleotide "guide" and 13 nucleotide target RNA for 2'-O-methylation (5GIO). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?q45tA8LEko87mYTw6

10.3.2: Summary

This chapter provides an in-depth exploration of RNA's unique chemical and biological roles, emphasizing how small differences in structure lead to diverse functions that are essential for cellular life.

Chemical Basis and Structural Diversity:
RNA differs from DNA primarily in that it contains ribose instead of deoxyribose and uracil in place of thymine. These seemingly minor differences not only alter its chemical properties but also empower RNA to adopt a wide range of structures—from simple stem-loop hairpins to complex tertiary conformations. Such structural variability is driven by noncanonical base pairing and chemical modifications, much like the dynamic folding observed in proteins.

Transcription and RNA Processing:
The chapter outlines the transcription process where RNA polymerases synthesize RNA from a DNA template. For coding RNAs, the initial heteronuclear transcript undergoes splicing to remove introns, resulting in mature messenger RNA (mRNA) that serves as the blueprint for protein synthesis. The intricacies of RNA processing are critical for ensuring that the correct genetic information is conveyed and translated.

Classification of RNA Types:
RNA molecules are broadly categorized into coding and noncoding RNAs. While mRNA carries the code for protein synthesis, noncoding RNAs (ncRNAs) serve regulatory and structural functions. The ncRNAs are further divided by size:

- **Short noncoding RNAs (<200 nucleotides):** Include tRNAs (adapters in translation), small nuclear RNAs (snRNAs) that are essential for splicing, microRNAs (miRNAs) and small interfering RNAs (siRNAs) involved in gene regulation, and other specialized molecules like piRNAs and snoRNAs.
- **Long noncoding RNAs (>200 nucleotides):** These include the well-known ribosomal RNAs (rRNAs) that form the core of the ribosome, as well as a diverse array of lncRNAs involved in transcriptional regulation, chromatin organization, and even direct protein interactions.

RNA–Protein Interactions and Functional Implications:
The chapter highlights numerous examples of RNA interacting with proteins to regulate cellular processes. From the assembly of the ribosome to RNA-based toxin-antitoxin systems in bacteria, these complexes underscore the importance of RNA's structure in mediating its interactions. Detailed structural models illustrate how RNA–protein binding can affect processes like mRNA splicing, translation, and gene silencing.

RNA in Gene Regulation and Disease:
A significant portion of the discussion is dedicated to the roles of miRNAs and lncRNAs in regulating gene expression. These molecules control the translation and stability of mRNAs and are implicated in various diseases, including cancer, cardiovascular disorders, and neurodegenerative conditions. Notably, recent advances—highlighted by the awarding of the 2024 Nobel Prize in Physiology and Medicine—underscore the importance of miRNAs in developmental processes and disease pathogenesis.

Integration of Computational and Experimental Tools:
The chapter also emphasizes modern techniques used to predict RNA secondary structures and analyze RNA–protein complexes.

Databases such as miRBase and lncipedia, along with interactive 3D structural models, allow biochemists to visualize the dynamic nature of RNA and its myriad interactions within the cell.

In summary, this chapter equips students with a comprehensive understanding of RNA's multifaceted roles—from its chemical foundation and diverse structural forms to its critical involvement in gene regulation and disease. Through the integration of molecular details and modern analytical tools, junior and senior biochemistry majors gain valuable insights into how RNA not only translates genetic information but also orchestrates a complex network of regulatory processes within the cell.

10.3.3: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics*. Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore*. Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

This page titled 10.3: Nucleic Acids - RNA Structure and Function is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **8.2: Nucleic Acids - RNA Structure and Function** by Henry Jakubowski and Patricia Flatt has no license indicated.

10.4: Nucleic Acids - Comparison of DNA and RNA

Learning Goals (ChatGPT o3-mini, 2/1/25)

1. **Explain Chemical Modifications:**
   - Describe how intentional chemical modifications—such as methylation and hydroxymethylation—alter the structure and function of DNA and RNA, including their roles in regulating transcription and translation.
2. **Distinguish Epigenetic Versus Epitranscriptomic Changes:**
   - Compare how DNA methylation (and histone modifications) contribute to epigenetic regulation, while analogous modifications in RNA contribute to the emerging field of epitranscriptomics.
3. **Analyze Mutation Mechanisms:**
   - Explain how spontaneous chemical reactions, such as the hydrolytic deamination of cytosine, lead to point mutations and discuss the role of repair enzymes (e.g., uracil-DNA glycosylases) in maintaining genomic integrity.
4. **Evaluate the Impact of Chemical Agents:**
   - Assess how external mutagens (like nitrous acid and alkylating agents) cause point mutations and structural rearrangements in DNA, and predict their potential consequences for genomic stability.
5. **Compare Uracil and Thymine Roles:**
   - Discuss the chemical rationale behind using uracil in RNA and thymine in DNA, emphasizing the stabilizing effect of the methyl group in thymine to prevent erroneous repair of deaminated cytosine.
6. **Examine Backbone Linkage Chemistry:**
   - Evaluate why DNA and RNA employ phosphodiester bonds instead of alternative linkages (e.g., carboxylic acid esters or amides) and how the chemical properties of these bonds contribute to the stability and overall structure of nucleic acids.
7. **Understand Sugar Structure and Flexibility:**
   - Describe how the presence or absence of a 2'-OH group in ribose versus deoxyribose influences molecular stability, backbone flexibility, and the ability to form extended double-helical structures.
8. **Interpret Sugar Puckering and Helical Conformations:**
   - Analyze the effects of ribose puckering (C3'-endo vs. C2'-endo) on the formation of A-RNA versus B-DNA helices, and explain how these conformational differences affect nucleic acid-protein interactions.
9. **Differentiate Base Pairing Modes:**
   - Contrast Watson-Crick and Hoogsteen base pairing in DNA, discuss how their dynamic equilibrium is influenced by structural modifications, and understand the implications for protein recognition and damaged DNA repair.
10. **Integrate Structure-Function Relationships:**
    - Synthesize the comparative structural features of DNA, RNA, and proteins to explain how these differences have been evolutionarily selected to optimize genetic information storage, replication fidelity, and catalytic activity.

These goals are intended to guide your study of how subtle chemical differences in nucleic acids impact their structure, stability, and biological functions, thereby deepening your understanding of molecular biology at a fundamental level.

Now that we understand the structures of DNA and the structures and various functions of RNA, we can more fully explore how their chemical similarities and differences contribute to different functions.

10.4.1: Chemical modifications of DNA and RNA

Post-translational modifications of proteins alter their structural/functional properties. Likewise, intentional chemical modifications of nucleic acid bases alter both their structures and potentially their transcriptional and translational status. Figure **10.4.1** shows common modifications of bases in DNA.

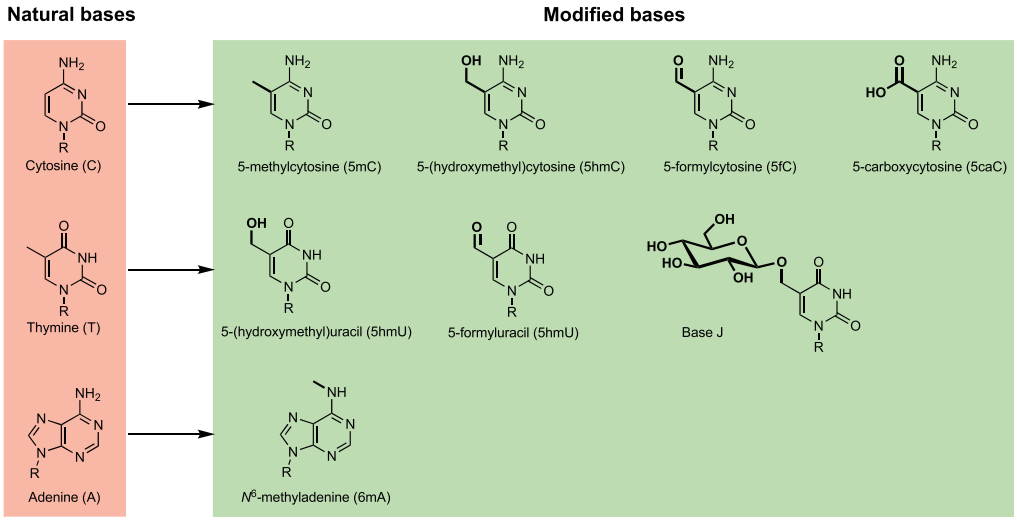


Figure **10.4.1**: Common modifications of bases in DNA. Matthew K.Bilyard et al. Current Opinion in Chemical Biology. Volume 57, August 2020, Pages 1-7. https://doi.org/10.1016/j.cbpa.2020.01.014. Under a Creative Commons license

Likewise, RNA is chemically modified. Figure **10.4.2** shows common modifications of bases in RNA. Methylation and subsequent hydroxylation to hydroxymethyl are common to both DNA and RNA. Methylation of DNA often represses the transcription of the DNA into RNA. Hence, it has huge potential to alter gene transcription. Such changes to the DNA are called **epigenetic** modifications. These changes can also be passed down to future generations and affect a cell's phenotype. Histone proteins involved in DNA packing into nucleosomes can also be methylated and acetylated, altering the interaction of the DNA with the nucleosome core and further packing, again affecting transcription.

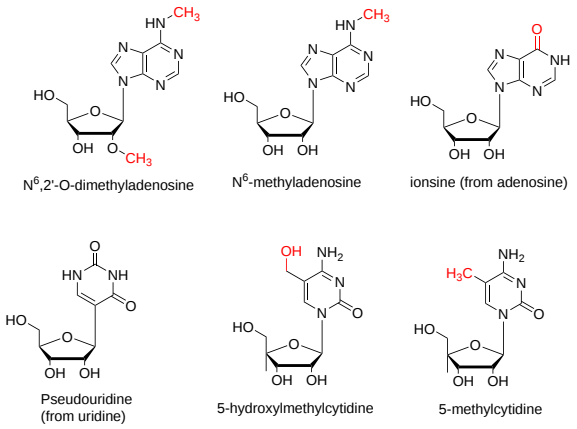


Figure **10.4.2**: Common modifications of bases in RNA

Chemical modification to RNA can also change the reading of the genome. The **epitranscriptome** refers to the collective chemical modifications to RNA, and its understanding is part of a new field, **epitranscriptomics**.

10.4.2: Mutations

A mutation can arise from the chemical modification of bases. Uracil in RNA is a demethylated form of thymine in DNA. In RNA, AU base pairs replace AT base pairs. Why is there a need for uracil in RNA? The question could be rephrased as to why there is a need for thymine, with its extra methyl group, in DNA. It's useful to think about the consequences of replacing a single H in a molecule with a $-CH_3$. Take HOH (i.e., water) as an example. Our bodies are over 60% water. We drink liters of water with a concentration of 55 M each day. Yet if we drink 0.07 L of methanol (CH_3OH), half of us would die! Let's probe some consequences

of the U (no $-CH_3$) and T (with $-CH_3$) changes in DNA. It can get confusing, but remember that the normal base pairs in DNA are AT, but AU base pairs also form (they are the norm in RNA). The $-CH_3$ substituent on thymine does not affect its base pairing.

a. Spontaneous deamination of cytosine in DNA

Why are we now discussing cytosine in DNA? One reason is that the most common mutation in DNA is a C to T replacement. One way that happens is through the spontaneous hydrolytic deamination of cytosine in DNA to uracil, which we have presumed to be found only in RNA. The mechanism for this deamination and subsequent conversion of a G**C** to an A**T** base pair is shown in Figure **10.4.3**. The inset box shows a simplified mechanism for spontaneous deamination.

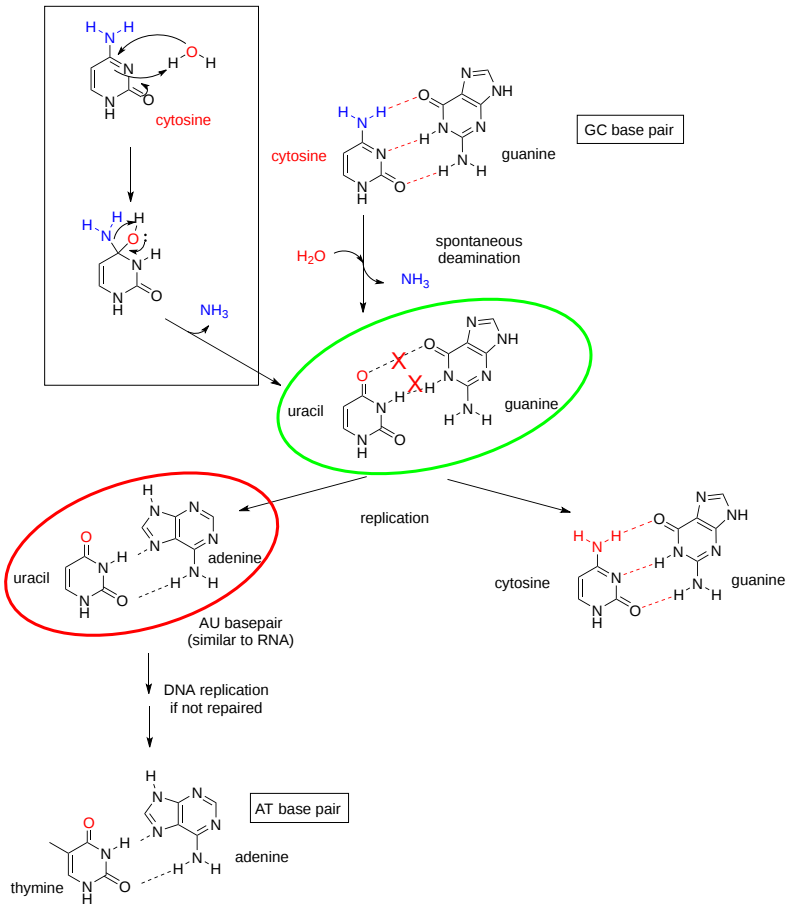


Figure **10.4.3**: GC to AT base pair mutation on spontaneous hydrolytic deamination of cytosine in DNA.

Hence, a possible consequence of the deamination reaction is a GC to AT base pair mutation if the uracil in DNA is not removed before DNA replication. Fortunately, the enzyme uracil-DNA glycosylases can remove any uracils found in DNA, leaving an abasic site, which can be fixed with DNA repair enzymes.

We can now ask why T and not U in DNA. Pretend you are a DNA repair enzyme and see a UA base pair in DNA. How can you tell if the UA base pair is correct and intended to be there or if it should be a CG base pair that underwent deamination? The most common uracil-DNA glycosylases remove the uracil whether it is across from guanine, the correct base but which can not hydrogen bond with uracil (in the **green oval** in Figure **10.4.3**), or if it is across from adenine, the wrong base (in **red oval**), which is present after a round of replication. Evolution has addressed this problem by adding a methyl group to uracil to form thymine and using that base, which forms a base pair with adenine. Now, no decision on which base across from a uracil (guanine if the uracil arose from deamination) or across from a "uracil-like" thymine (adenine) is correct.

b. Other mutations

Since we are considering chemical modifications to DNA and mutations, giving a more expanded background on them is appropriate. In addition to mutations caused by spontaneous hydrolytic cytosine deamination, mutations can also arise by adding a wrong base during DNA replication, by chemical damage caused by radiation or chemical modifying agents. How many mistakes

in replication are made? You would be ecstatic if you received a 99% on an examination. That's not good enough for DNA replication. In Cell Biology by the Numbers, they calculate it this way. Assume the replication /repair is so good that it takes 10^8 replications to make a mistake (an error rate of 10^{-8}/BP). Assume also that there are 3 x 10^9 base pairs in the human genome. This leads to a mutation rate of 10-100 mutations/genome/generation or about 0.1-1 mutations/genome/replication. Not bad!

Figure **10.4.4** shows how common point mutations might randomly arise.

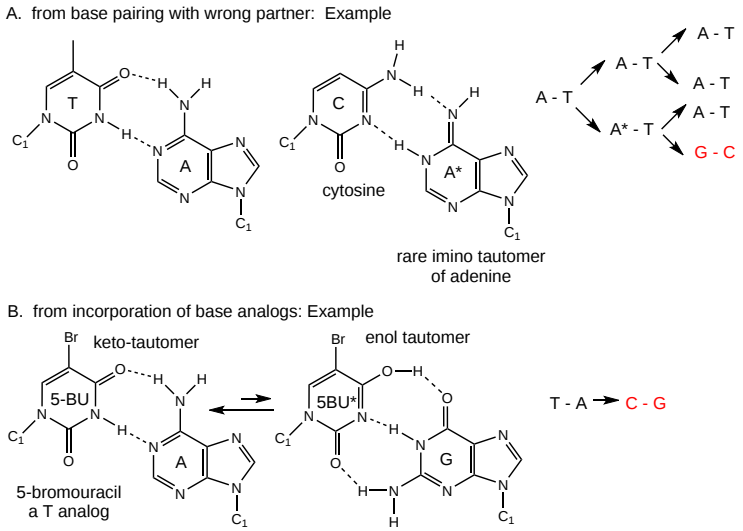


Figure **10.4.4**: How common point mutations might arise randomly

Chemical agents can also cause point mutations. Figure **10.4.5** shows point mutations arising from oxidative deaminations (not hydrolytic) by nitrous acid/nitrosamines and from alkylating agents.

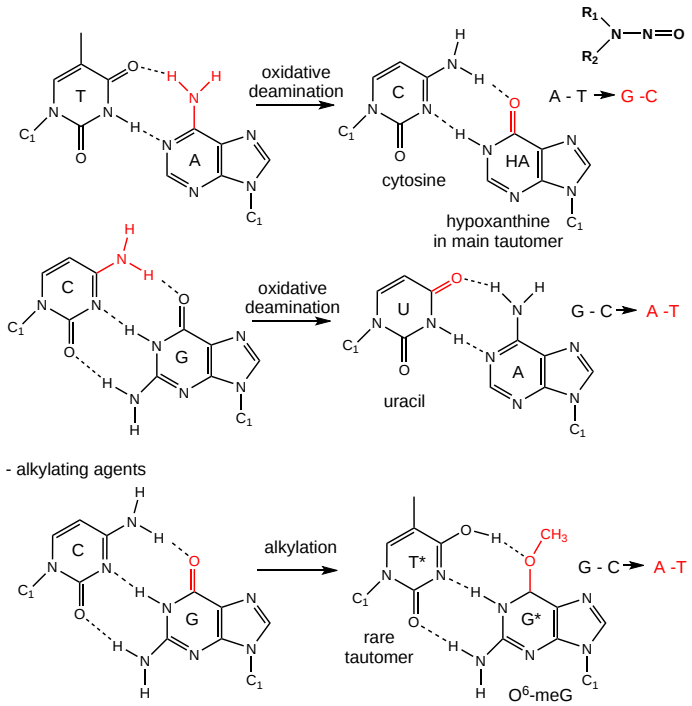


Figure **10.4.5**: Nitrous acid/nitrosamines and alkylating agent point mutations

Figure **10.4.6** shows a variety of alkylating agents with mutagenic potential.

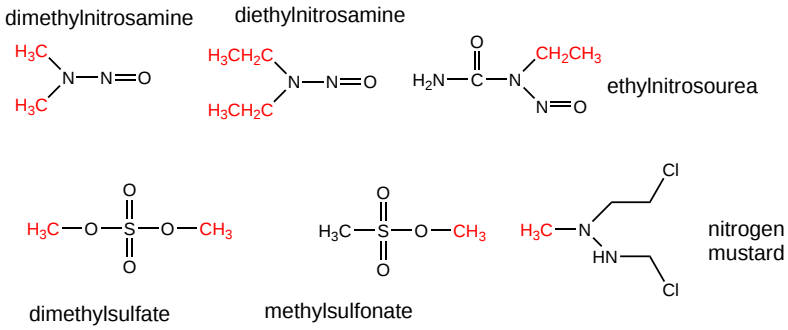


Figure 10.4.6: Alkylating agents with mutagenic potential.

Finally, large-scale changes in chromosome structure can also occur, as shown in Figure 10.4.7, usually with profound consequences.

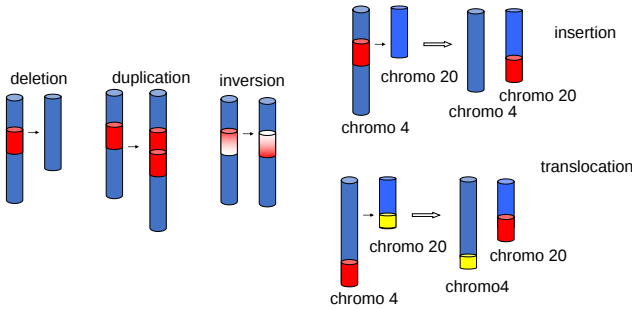


Figure 10.4.7: Large scale structural rearrangements in DNA

10.4.3: Why DNA and RNA - A Chemical Perspective

Asking a "why" question (like above) in the sciences is inappropriate, as teleological questions are more philosophical or religious. Yet we will in this section, in part, to be in the company of Alexander Rich, who wrote a very cool article entitled "Why RNA and DNA have different Structures".

Given that RNA expresses catalytic activities and can carry genetic information (some viruses have ds and ss RNA as their genome), it has been suggested that early life might have been based on RNA. DNA would evolve later as a more secure carrier of genetic information. Inspecting the chemical properties of DNA, RNA, and proteins shows them to have attributes needed for their expressed function. Let's examine each for structural features that might be important for function.

a. Why does DNA lack a 2' OH group (found in RNA), which has been replaced with hydrogen? This required the evolutionary creation of a new enzyme, ribonucleotide reductase, to catalyze the replacement of the OH in a ribonucleotide monomer to form the deoxyribonucleotide form. One possible explanation is offered in the figure below. DNA, the main carrier of genetic information, must be an extremely stable molecule. An OH present on C'2 could act as a nucleophile and attack the proximal P in the phosphodiester bond, leading to a nucleophilic substitution reaction and potential cleavage of the bond. RNA, an intermediary molecule whose concentration (at least as mRNA) should rise and fall based on the need for a potential transcript, should be more labile to such hydrolysis. Figure 10.4.8 shows a possible reaction diagram for the internal cleavage of RNA. (The reaction would probably proceed with no actual intermediate, but just a transition state.

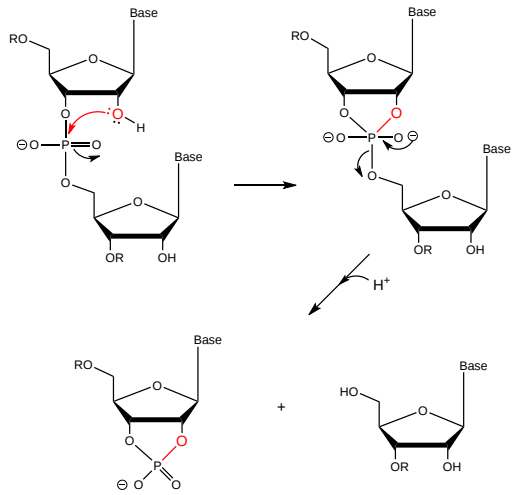


Figure 10.4.8: Internal cleave of RNA using the C'2-OH as an intramolecular nucleophile

b. Why do both DNA and RNA contain a phosphodiester link between adjacent monomers instead of more "traditional" links such as carboxylic acid esters, amides, or anhydrides? One possible explanation is given below. Nucleophilic attack on the sp^3 hybridized P in a phosphodiester is much more difficult than for a more open sp^2 hybridized carboxylic acid derivative. In addition, the negative charge on the O in the phosphodiester link would decrease the likelihood of a nucleophilic attack. The negative charges on both strands in ds-DNA probably help keep the strands separated, allowing the traditional base pairing and double-stranded helical structure to be observed. The cleavage of the phosphodiester link in DNA and a hypothetical ester link is shown in Figure 10.4.9. Again, the reaction of the phosphodiester shows a pentavalent intermediate, but most likely, the reaction proceeds directly from the transition state.

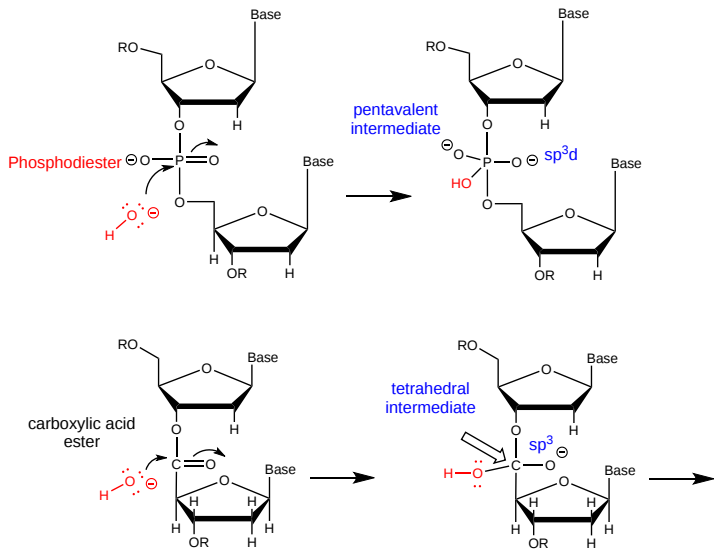


Figure 10.4.9: cleavage of the phosphodiester link in DNA and a hypothetical ester link

c. Why is DNA found as a repetitive double-stranded helix, but RNA is usually found as a single-stranded molecule that can form complicated tertiary structures with some dsRNA motifs?

Another reason for the absence of the 2' OH in DNA is that it allows the deoxyribose ring in DNA to pucker just the right way to allow extended ds-DNA helices (B type). The pucker in deoxyribose and ribose can be visualized by visualizing a single plane in the sugar ring defined by the ring atoms C1', O, and C4'. If a ring atom points in the same direction as the C4'-C5' bond, the ring atom is defined as **endo**. If it points in the opposite direction, it is defined as **exo**. In the most common form of double-stranded DNA, B-DNA, the iconic extended double helix, C2' is in the endo form. It can also adopt the C3' endo form, forming another less common helix, a more open ds-A helix. In contrast, steric interference prevents ribose in RNA from adopting the 2'endo

conformation. It allows only the 3'endo form, precluding the occurrences of extended ds-B-RNA helices but allowing more open, A-type helices.

Figure **10.4.10** shows another comparison between the A-RNA and B-DNA double helices and the C'3 and C'2 endo forms of the ribose

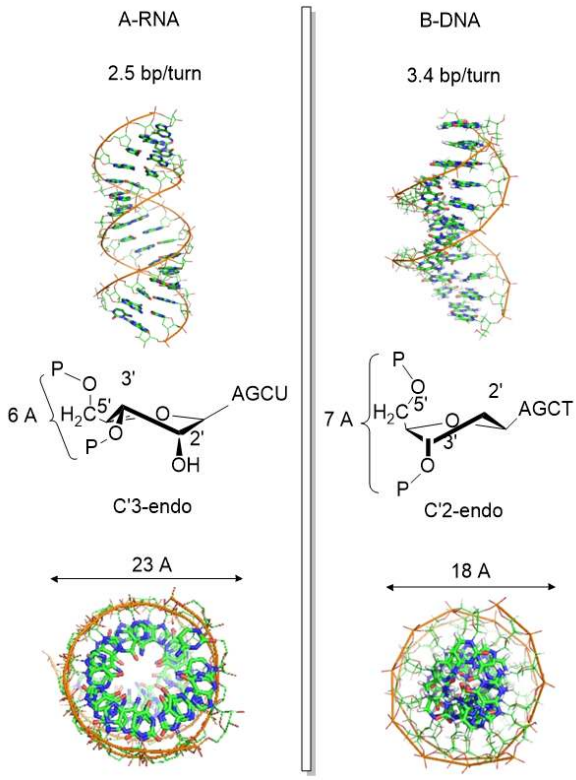
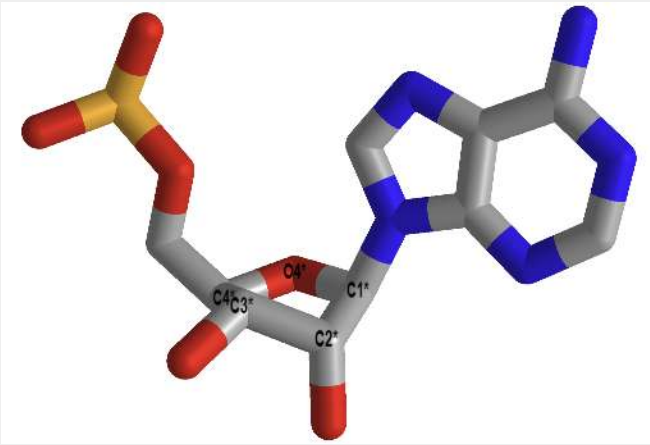
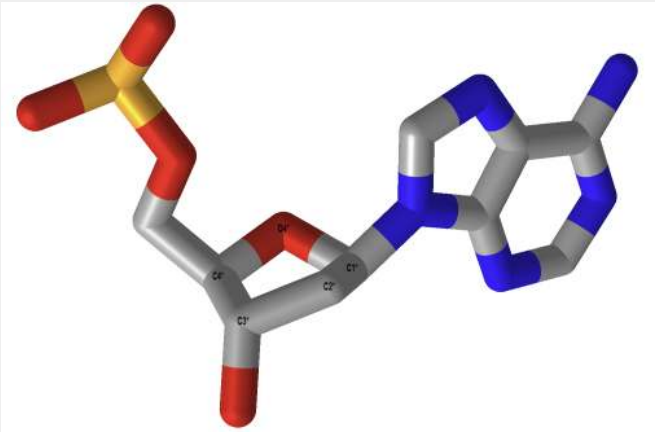


Figure **10.4.10**: after Zhou et al. Nature Structural and Molecular Biology. doi:10.1038/nsmb.3270

Figure **10.4.5** shows interactive iCn3D models of the pentoses in a strand of A-RNA (413D), double-stranded, left, and B-DNA (1BNA), double-stranded, right.

C'3-endo ribose, A-RNA (413D, double stranded)	C'2 endo ribose, B-DNA (1BNA, double stranded)
	
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KPueqrBADczh26	Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BEn5nqsCQG2JH6

d. <u>What about the molecular dynamics of A-RNA and B-DNA?</u>

The information above suggests that the sugar ring of DNA is conformationally more flexible than the ribose ring of RNA. This can be inferred from the observation that dsDNA can adopt B and A forms, which requires a switch from the 2' endo in the B form to the 3'endo form in the A form. The smaller H on the 2'C would offer less steric interference with such flexibility. The rigidity in ribose is associated with a smaller 5'O to 3'O distance in RNA, leading to a compression of the nucleotides into a helix with a smaller number of base pairs/turn.

The increased flexibility in DNA allows rotation around the C1'-N glycosidic bond connecting the deoxyribose and base in DNA, allowing different orientations of AT and GC base pairs with each other. The normal "anti" orientation allows "Watson-Crick" (WC) base pairing between AT and GC base pairs, while the altered rotation allows "Hoogsteen" (Hoog) base pairs. Figure **10.4.11** shows the different orientations for an AT base pair.

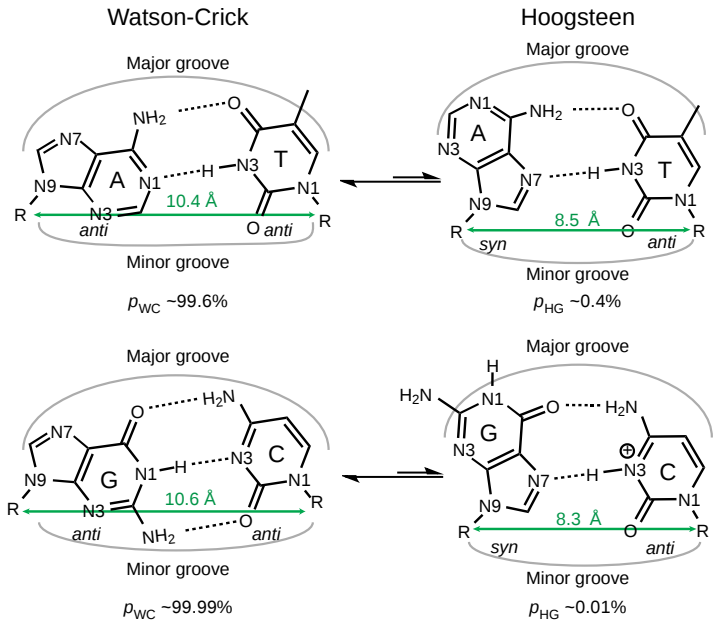


Figure **10.4.11**: Xu, Y., McSally, J., Andricioaei, I. *et al*. Modulation of Hoogsteen dynamics on DNA recognition. *Nat Commun* **9,** 1473 (2018). https://doi.org/10.1038/s41467-018-03516-1Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The Watson-Crick (WC) and Hoogsteen (HG) base pairs in B-DNA are in a dynamic equilibrium, with the equilibrium greatly favoring the WC form, as indicated by the arrows in the figure above. In a DNA:protein complex, the WC ↔ HG equilibrium can favor the WG form for AT and GC^+ forms (in the latter, the C is protonated) when those base pairs are also involved in protein recognition. They can also occur more frequently in damaged DNA. In contrast, molecular dynamics studies show that the HG base pairs A-U and GC^+ are strongly disfavored in ds A-RNA.

One type of DNA damage is methylation on N1-adenosine and N1-guanosine. This modification prevents normal Watson-Crick base pairing, but for DNA, these modified bases can still engage in Hoogsteen base pairing, preserving the overall structure of dsDNA and its ability to carry genetic information stably. This same methylation occurs normally in post-transcriptionally modified RNA. Hence, N1 adenosine and N1 guanosine methylation prevent any base pairing in the modified RNA. These properties make DNA a better carrier of molecular information and offer another way to regulate RNA's structural and functional properties.

Hoogsteen base pairs can be found in distorted dsDNA structures (caused by protein:DNA interactions) and normal B-DNA. Figure **10.4.12** shows a Hoogsteen base pair between dA7 and dT37 in the MAT α 2 homeodomain:DNA complex (pdb 1K61). Note that the dA base in the Hoogsteen base pair is rotated syn (with respect to the deoxyribose ring) instead of the usual anti, allowing the Hoogsteen base pair.

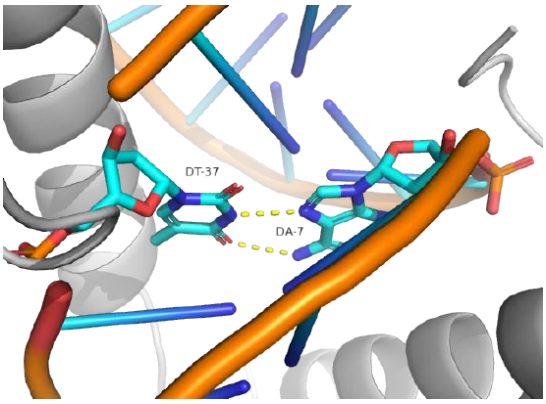


Figure 10.4.12: Hoogsteen base pair between dA7 and dT37 in the MAT α 2 homeodomain:DNA complex (pdb 1K61)

10.4.4: A Structural Comparison

Now, let's review the structures adopted by the three major macromolecules: DNA, RNA, and proteins. DNA predominantly adopts the classic ds-BDNA structure, although it is wound around nucleosomes and "supercoiled" in eukaryotic cells since it must be packed into the nucleus. Prokaryotic DNA is typically packed into a more amorphous nuclear region, the nucleoid, through interactions with other proteins that also facilitate supercoiling. It is, in effect, a dynamic molecular condensate.

The extended ds-BDNA helical form arises partly from the significant electrostatic repulsions of two strands of this polyanion (even in counter-ions). Given its high charge density, it is unsurprising that it forms complexes with positive proteins and does not adopt complex tertiary structures. RNA, conversely, can not form long B-type double-stranded helices (due to steric constraints of the 2'OH and the resulting 3'endo ribose pucker). Rather, it can adopt complex tertiary conformations (albeit with significant counter-ion binding to stabilize the structure) and, in doing so, can form regions of secondary structure (ds-A RNA) in the form of stem/hairpin forms. Proteins, with their combination of polar charged, polar uncharged, and nonpolar side chains, offer little electrostatic hindrance in adopting secondary and tertiary structures. RNA and proteins can adopt tertiary structures with potential binding and catalytic sites, making them ideal catalysts for chemical reactions. Given its four-nucleotide alphabet, RNA can also carry genetic information, making it an ideal candidate for the first evolved macromolecules enabling the development of life. Proteins with abundant organic functionalities eventually supplanted RNA as a better choice for life's catalyst. DNA, with its greater stability, would supplant RNA as the choice for the primary carrier of genetic information (Figure 10.4.13):

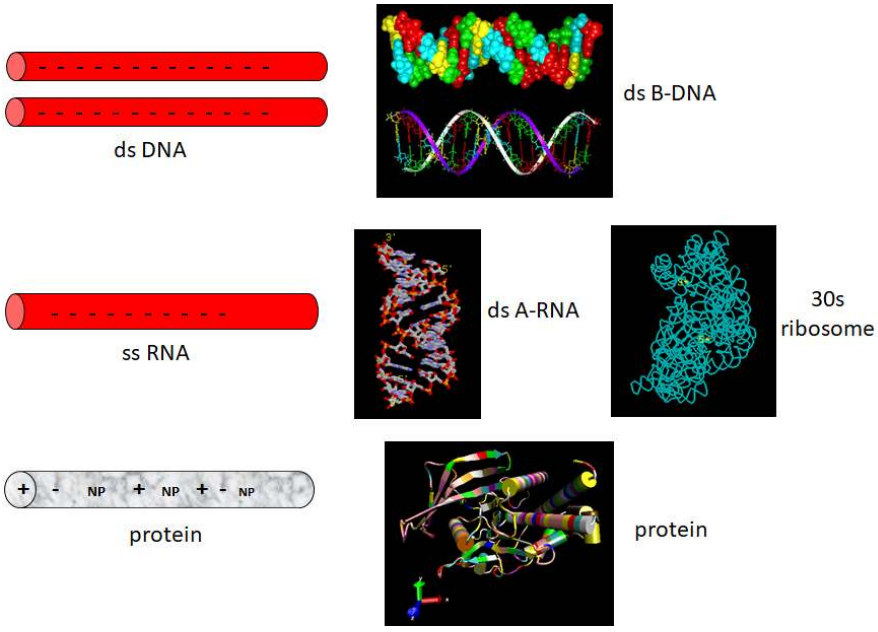


Figure 10.4.13: Summary comparison of DNA, RNA, and Protein structures

A final note on the simplicity of the dsDNA structure. A mutation causing a single base pair change in DNA does **not** change the iconic ds-stranded DNA structure. If it did, DNA would not be a reliable molecule to store and read out the genetic blueprint. In contrast, a single mutation in the DNA leading to a single amino acid substitution may lead to a protein with altered structure and function. This could be deleterious or even fatal to the organism. On the other hand, the new protein structure might have new functionalities that allow adaptation to new environments or allow new types of reactions. Evolution would favor the latter.

10.4.5: Summary

This chapter examines how subtle chemical variations and modifications in nucleic acids—DNA and RNA—lead to distinct structural and functional outcomes that are essential for life. Designed for junior and senior biochemistry majors, the chapter integrates concepts from chemical modifications and mutation mechanisms to explain why DNA and RNA serve different roles in the cell.

Chemical Modifications and Regulation:
Both DNA and RNA undergo intentional chemical modifications that can alter their physical structure and influence gene expression. In DNA, common modifications such as methylation (and subsequent hydroxymethylation) are central to epigenetic regulation. These modifications can repress transcription and are heritable, impacting cell phenotype over generations. RNA modifications, collectively referred to as the epitranscriptome, similarly affect RNA stability, processing, and translation, highlighting the dynamic regulatory potential of these molecules.

Mutation Mechanisms and DNA Repair:
The chapter discusses how chemical modifications can also lead to mutations. For example, the spontaneous hydrolytic deamination of cytosine to uracil in DNA represents a major source of point mutations, potentially converting GC to AT base pairs if not corrected. DNA repair enzymes, such as uracil-DNA glycosylases, play a critical role in identifying and excising these aberrant bases. Additionally, errors during DNA replication and damage induced by chemical agents (e.g., nitrous acid, alkylating agents) contribute to mutations ranging from single base changes to large-scale chromosomal rearrangements.

Structural Basis for Functional Differences:
A significant portion of the chapter is devoted to understanding why DNA and RNA, despite their chemical similarities, adopt very different structures and fulfill distinct roles:

- **Sugar Chemistry and Backbone Stability:**
  DNA's deoxyribose lacks a 2'-OH group, making it more chemically stable and less prone to self-cleavage compared to RNA, which contains ribose with a reactive 2'-OH. This difference is crucial for DNA's role as a long-term, reliable repository of genetic information, whereas RNA's inherent lability suits its function in transient information transfer and catalysis.
- **Phosphodiester Linkages:**
  Both nucleic acids use phosphodiester bonds to link nucleotides. The inherent chemical resistance of these bonds to nucleophilic attack—enhanced by their negative charge—contributes to the overall stability of the DNA double helix and the more dynamic secondary and tertiary structures found in RNA.
- **Sugar Puckering and Helical Forms:**
  The conformational flexibility of the sugar ring is a key determinant of nucleic acid structure. In DNA, the deoxyribose can adopt different puckers (C2'-endo for B-DNA, C3'-endo for A-DNA), which influences the overall helix geometry. In contrast, the ribose in RNA is more restricted, favoring the C3'-endo conformation, which precludes extended double-stranded helices and promotes the formation of complex tertiary structures.
- **Base Pairing Dynamics:**
  The chapter also explores the equilibrium between Watson-Crick and Hoogsteen base pairing in DNA. Although Watson-Crick pairing predominates in stable B-DNA, dynamic shifts to Hoogsteen pairings can occur—particularly in protein-DNA complexes or damaged regions—affecting recognition and repair mechanisms. In RNA, such alternative pairing is less common due to structural constraints imposed by the ribose.

Comparative Structural Overview:
In a broader context, the chapter compares the structural attributes of DNA, RNA, and proteins. DNA's robust double-helical structure, which is maintained despite mutations, underscores its reliability as a genetic archive. RNA's versatility in folding into complex structures enables it to perform catalytic and regulatory roles, while proteins, with their diverse side chains, are well-

suited to function as dynamic catalysts and structural components. This comparison reinforces the evolutionary specialization of these macromolecules for their respective biological roles.

In summary, the chapter provides a comprehensive exploration of how chemical modifications and intrinsic structural features govern the stability, function, and evolutionary utility of DNA and RNA. Understanding these principles is fundamental for appreciating the molecular basis of gene regulation, mutation, and the overall maintenance of genetic integrity in biological systems.

10.4.6: References

Börner, R., Kowerko, D., Miserachs, H.G., Shaffer, M., and Sigel, R.K.O. (2016) Metal ion induced heterogeneity in RNA folding studied by smFRET. *Coordination Chemistry Reviews* 327 DOI: 10.1016/j.ccr.2016.06.002 Available at: https://www.researchgate.net/publication/303846502_Metal_ion_induced_heterogeneity_in_RNA_folding_studied_by_smFRET

Hardison, R. (2019) B-Form, A-Form, and Z-Form of DNA. Chapter in: R. Hardison's *Working with Molecular Genetics*. Published by LibreTexts. Available at: https://bio.libretexts.org/Bookshelves/Genetics/Book%3A_Working_with_Molecular_Genetics_(Hardison)/Unit_I%3A_Genes%2C_Nucleic_Acids%2C_Genomes_and_Chromosomes/2%3A_Structures_of_Nucleic_Acids/2.5%3A_B-Form%2C_A-Form%2C_and_Z-Form_of_DNA

Lenglet, G., David-Cordonnier, M-H., (2010) DNA-destabilizing agents as an alternative approach for targeting DNA: Mechanisms of action and cellular consequences. *Journal of Nucleic Acids* 2010, Article ID: 290935, DOI: 10.4061/2010/290935 Available at: https://www.hindawi.com/journals/jna/2010/290935/

Mechanobiology Institute (2018) What are chromosomes and chromosome territories? *Produced by the National University of Singapore*. Available at: https://www.mechanobio.info/genome-regulation/what-are-chromosomes-and-chromosome-territories/

National Human Genome Research Institute (2019) The Human Genome Project. *National Institutes of Health*. Available at: https://www.genome.gov/human-genome-project

Wikipedia contributors. (2019, July 8). DNA. In *Wikipedia, The Free Encyclopedia*. Retrieved 02:41, July 22, 2019, from https://en.Wikipedia.org/w/index.php?title=DNA&oldid=905364161

Wikipedia contributors. (2019, July 22). Chromosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:18, July 23, 2019, from en.Wikipedia.org/w/index.php?title=Chromosome&oldid=907355235

Wikilectures. Prokaryotic Chromosomes (2017) In MediaWiki, Available at: https://www.wikilectures.eu/w/Prokaryotic_Chromosomes

Wikipedia contributors. (2019, May 15). DNA supercoil. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:40, July 25, 2019, from en.Wikipedia.org/w/index.php?title=DNA_supercoil&oldid=897160342

Wikipedia contributors. (2019, July 23). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 16:19, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Histone&oldid=907472227

Wikipedia contributors. (2019, July 17). Nucleosome. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:17, July 26, 2019, from en.Wikipedia.org/w/index.php?title=Nucleosome&oldid=906654745

Wikipedia contributors. (2019, July 26). Human genome. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:12, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Human_genome&oldid=908031878

Wikipedia contributors. (2019, July 19). Gene structure. In *Wikipedia, The Free Encyclopedia*. Retrieved 06:16, July 27, 2019, from en.Wikipedia.org/w/index.php?title=Gene_structure&oldid=906938498

This page titled 10.4: Nucleic Acids - Comparison of DNA and RNA is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **8.3: Nucleic Acids - Comparison of DNA and RNA** by Henry Jakubowski and Patricia Flatt is licensed CC BY-SA 4.0.

10.5: Enzymes for Genetic modifications

Learning Goals (ChatGPT o3-mini, 2/1/25)

1. **Describe the Mechanism of Restriction Endonucleases:**
   - Explain how Type II restriction enzymes recognize specific palindromic sequences in double-stranded DNA and catalyze double-strand breaks, generating predictable fragment sizes.
2. **Understand the Restriction-Modification System:**
   - Illustrate how bacteria use restriction enzymes together with DNA-methyltransferases to differentiate between self and non-self DNA, thereby protecting their genomes from viral infection.
3. **Analyze the Catalytic Mechanism of Type II Restriction Enzymes:**
   - Describe the role of divalent metal ions (e.g., Mg^{2+}) in activating water for nucleophilic attack on the phosphodiester bond and discuss the stereochemical evidence (inversion of configuration) that supports a direct hydrolytic mechanism.
4. **Interpret Enzyme Specificity and Nomenclature:**
   - Discuss how restriction enzymes are named based on their organism of origin and how isoschizomers and neoschizomers differ in their cleavage patterns and evolutionary origins.
5. **Compare and Contrast Gene Editing Technologies:**
   - Compare traditional restriction enzyme-based cloning methods with modern CRISPR-Cas9 gene editing, emphasizing their mechanisms, applications, and limitations.
6. **Explain the Structure and Function of CRISPR-Cas9:**
   - Describe the architecture of Cas9, including its two catalytic nuclease domains (HNH and RuvC) and the roles of the guide RNA (crRNA and tracrRNA, or the synthetic sgRNA) in conferring target specificity.
7. **Understand the Allosteric Regulation of Cas9:**
   - Analyze how the binding of guide RNA triggers conformational changes in Cas9 that prime the enzyme for DNA binding and cleavage, and explain the role of the protospacer adjacent motif (PAM) in target recognition and cleavage specificity.
8. **Evaluate the Applications and Challenges of CRISPR Gene Editing:**
   - Assess how CRISPR-Cas9 is used in disease diagnosis, therapy, and genetic modifications, and discuss strategies for minimizing off-target effects and ensuring efficient delivery into target cells.
9. **Integrate Concepts in Protein Structure and Enzyme Mechanism:**
   - Relate the structural features of both restriction enzymes and Cas9 to their catalytic mechanisms, emphasizing how protein conformation, cofactor binding, and specific amino acid residues contribute to their function.

These goals are intended to guide your study and critical analysis of gene editing systems, helping you connect fundamental enzyme mechanisms with their revolutionary applications in biotechnology and medicine.

It is difficult to read newspapers and newsmagazines without encountering the CRISPR-Cas9 gene editing system that has the potential to make gene editing routine in disease diagnosis, treatment, and cure, as well as in genetic modification of organisms to improve their quality and quantity for food and natural product production. In this chapter section, we will explore the mechanism of restriction enzymes that made gene cloning possible as well as the CRISPR-Cas gene editing system.

10.5.1: Restriction Endonucleases

A ***restriction enzyme***, ***restriction endonuclease***, or ***restrictase*** is an enzyme that cleaves DNA into fragments at or near specific recognition sites within molecules known as restriction sites. Restriction enzymes are one class of the broader endonuclease group

of enzymes. Restriction enzymes are commonly classified into five types, which differ in their structure and whether they cut their DNA substrate at their recognition site or if the recognition and cleavage sites are separate. To cut DNA, all restriction enzymes make two incisions, once through each sugar-phosphate backbone (i.e., each strand) of the DNA double helix. Here, we will focus on the Type II restriction enzymes routinely used in molecular biology and biotechnology applications.

As with other restriction enzymes, Type II Restriction Enzymes occur exclusively in unicellular microbial life forms—mainly bacteria and archaea (prokaryotes)—and are thought to protect these cells from viruses and other infectious DNA molecules. Inside a prokaryote, the restriction enzymes selectively cut up *foreign* DNA in a process called ***restriction digestion*;** meanwhile, host DNA is protected by a modification enzyme (a methyltransferase) that modifies the prokaryotic DNA and blocks cleavage. Together, these two processes form the ***restriction-modification system***.

The first Type II Restriction Enzyme discovered was HindII from the bacterium *Haemophilus influenzae* Rd. The event was described by Hamilton Smith (Figure 7.23) in his Nobel lecture, delivered on 8 December 1978:

‘"In one such experiment we happened to use labeled DNA from phage P22, a bacterial virus I had worked with for several years before coming to Hopkins. To our surprise, we could not recover the foreign DNA from the cells. With Meselson’s recent report in our minds, we immediately suspected that it might be undergoing restriction, and our experience with viscometry told us that this would be a good assay for such an activity. The following day, two viscometers were set up, one containing P22 DNA and the other Haemophilus *DNA. Cell extract was added to each and we began quickly taking measurements. As the experiment progressed, we became increasingly excited as the viscosity of the* Haemophilus *DNA held steady while the P22 DNA viscosity fell. We were confident that we had discovered a new and highly active restriction enzyme. Furthermore, it appeared to require only* Mg^{2+} *as a cofactor, suggesting that it would prove to be a simpler enzyme than that from* E. coli *K or B.*

After several false starts and many tedious hours with our laborious, but sensitive viscometer assay, Wilcox and I succeeded in obtaining a purified preparation of the restriction enzyme. We next used sucrose gradient centrifugation to show that the purified enzyme selectively degraded duplex, but not single-stranded, P22 DNA to fragments averaging around 100 bp in length, while Haemophilus *DNA present in the same reaction mixture was untouched. No free nucleotides were released during the reaction, nor could we detect any nicks in the DNA products. Thus, the enzyme was clearly an endonuclease that produced double-strand breaks and was specific for foreign DNA. Since the final (limit) digestion products of foreign DNA remained large, it seemed to us that cleavage must be site-specific. This proved to be case and we were able to demonstrate it directly by sequencing the termini of the cleavage fragments.’"*

Figure (\PageIndex{35}\): Hamilton Smith and Daniel Nathans at the Nobel Prize press conference, 12 October 1978 (reproduced with permission from Susie Fitzhugh). Original Repository: Alan Mason Chesney Medical Archives, Daniel Nathans Collection. Image from: Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527.

Restriction enzymes are named according to the taxonomy of the organism in which they were discovered. The first letter of the enzyme refers to the genus of the organism, and the second and third to the species. This is followed by letters and/or numbers identifying the isolate. Roman numerals are used to specify different enzymes from the same organism. For example, the enzyme ‘HindIII’ was discovered in *Haemophilus influenzae*, serotype d, and is distinct from the HindI and HindII endonucleases also present within this bacterium. The DNA-methyltransferases (MTases) accompanying restriction enzymes are named similarly, and

given the prefix ‘M.’. When there is more than one MTase, they are prefixed ‘M1.’, ‘M2.’, etc, if they are separate proteins or ‘M1∼M2.’ when joined.

Restriction Enzymes that recognize the same DNA sequence, regardless of where they cut, are termed ***‘isoschizomers’*** (iso = equal; skhizo = split). ***Isoschizomers*** that cut the same sequence at different positions are called ***‘neoschizomers’*** (neo = new). ***Isoschizomers*** that cut at the same position are frequently, but not always, evolutionarily drifted versions of the same enzyme (e.g. BamHI and OkrAI). ***Neoschizomers,*** on the other hand, are often evolutionarily unrelated enzymes (e.g., EcoRII and MvaI).

Type II Restriction Enzymes are a conglomeration of many different proteins that, by definition, have the common ability to cleave duplex DNA at a fixed position within, or close to, their recognition sequence. This cleavage generates reproducible DNA fragments and predictable gel electrophoresis patterns, properties that have made these enzymes invaluable reagents for laboratory DNA manipulation and investigation. Almost all Type II Restriction Enzymes require divalent cations, usually Mg^{2+}, as essential components of their catalytic sites. Ca^{2+}, on the other hand, often acts as an inhibitor of Type II Restriction Enzymes.

The recognition sequences of Type II Restriction Enzymes are ***palindromic***, with two possible types of palindromic sequences. The ***mirror-like palindrome*** is similar to those found in ordinary text, in which a sequence reads the same forward and backward on a single strand of DNA, as in GTAATG. The ***inverted repeat palindrome*** is also a sequence that reads the same forward and backward, but the forward and backward sequences are found in complementary DNA strands (i.e., of double-stranded DNA), as in GTATAC (GTATAC being complementary to CATATG). ***Inverted repeat palindromes*** are more common and have greater biological importance than ***mirror-like palindromes.*** The position of cleavage within the palindromic sequence can vary depending on the enzyme and can produce either single-stranded overhanging sequences (sticky ends) or blunt-ended DNA products. Table 10.5.8 below shows examples of staggered and blunt end cuts by restriction enzymes.

EcoR1	
Sma1	

Table 10.5.8: Staggered and blunt end cut sequences by EcoR1 and Sma1

The host can use methylation to protect its own genome from cleavage. For example, the methylation of the EcoRI recognition sequence by the M.EcoRI methyltransferase (MTase), changes the sequence from GAATTC to GAm6ATTC (m6A = N6-methyladenine). This modification completely protects the sequence from EcoRI cleavage.

Type II Restriction Enzymes initially bind non-specifically with the DNA and proceed to slide down the DNA scanning for recognition sequences as shown in Figure (\PageIndex{36}\):. Upon binding to the correct palindromic sequence, the enzyme associates with the metal cofactor and mediates catalytic cleavage of the DNA using the mechanism of ***strain distortion*** and ***catalysis by approximation***.

Figure (\PageIndex{36}\): DNA Recognition and Cleavage by Type II Restriction Endonucleases. (A) Pictorial view of an EcoRV dimer scanning nonspecifically along the DNA until a specific binding site is recognized. This causes coupling with the metal cofactor and strain distortion of the DNA. Hydrolysis of the phosphodiester bond is mediated, and the DNA cleavage products are released from the enzyme. (B) shows a space-filling model of EcoRV DNA recognition and cleavage. Figure (A) from Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527. and Figure (B) from Thomas Splettstoesser

One of the most important questions regarding the catalytic mechanism of a hydrolase is whether hydrolysis involves a covalent intermediate, as is typical for the proteases described previously. This can be decided by analyzing the stereochemical course of the reaction. This was done first for EcoRI and later for EcoRV. Both enzymes were found to cleave the phosphodiester bond with inversion of the chiral center at the phosphorus, which argues against the formation of a covalent enzyme–DNA intermediate. Thus, it is proposed that cleavage involves the direct nucleophilic attack of the substrate by a water molecule, as shown in Figure (\PageIndex{37}\) below.

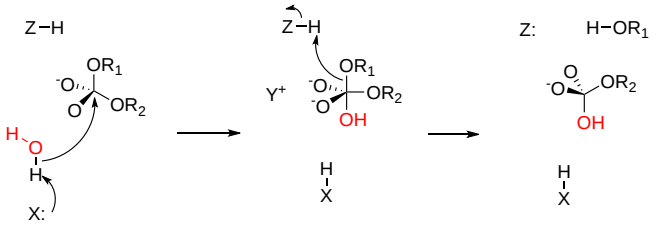


Figure (\PageIndex{37}\): A General Mechanism for DNA Cleavage by EcoRI and EcoRV. An activated water molecule attacks the phosphorus in line with the phosphodiester bond to be cleaved, which proceeds with an inversion of configuration.

X, Y, and Z are a general base, a Lewis acid, and a general acid, respectively. Figure adapted from: Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527.

Type II restriction enzymes typically form a homodimer when binding with DNA, as shown in the crystal structure of ***Bgl*II** in Figure 7.26B. ***Bgl*II** catalyzes phosphodiester bond cleavage at the DNA backbone through a phosphoryl transfer to water. Studies on the mechanism of restriction enzymes have revealed several general features that seem true in almost all cases. However, the actual mechanism for each enzyme is most likely some variation of this general mechanism (Figure 7.25). This mechanism requires

a base to generate the hydroxide ion from water, acting as the nucleophile and attacking the phosphorus in the phosphodiester bond. Also required is a Lewis acid to stabilize the extra negative charge of the pentacoordinate transition state phosphorus, as well as a general acid or metal ion that stabilizes the leaving group (3'-O^-). Two divalent metal cofactors are required in some Type II Restriction Enzymes (such as in EcoRV and BamHI). In contrast, other enzymes only require one divalent metal cofactor (such as in EcoRI and BglII).

Structural studies of endonucleases have revealed a similar architecture for the active site with the residues following the weak consensus sequence Glu/Asp-$(X)_{9\text{-}20}$-Glu/Asp/Ser-X-Lys/Glu. ***Bgl*II's** active site is similar to other endonucleases', following the Asp-$(X)_9$-Glu-X-Gln sequence. In its active site, a divalent metal cation, most likely Mg^{2+}, interacts with Asp-84, Val-94, a phosphoryl oxygen, and three water molecules. One of these water molecules can act as a nucleophile because of its proximity to the scissile phosphoryl group (Figure 7.26A). The nucleophilic water molecule is positioned for attack onto the phosphoryl group by a hydrogen bond with the side chain amide oxygen of Gln-95 and its contact with the metal cation. Interaction with the metal cation effectively lowers its pK_a, promoting the water's nucleophilicity as shown in Panel A of Figure (\PageIndex{38}\) below (from Pingoud, A., Wilson, G.G., and Wende, W. (2014) Nuc Acids Res 42(12):7489-7527). During hydrolysis, the divalent cation can stabilize the 3'-O^- leaving group and coordinate proton abstraction from one of the coordinated water molecules

Figure (\PageIndex{38}\): Proposed Reaction Mechanism for the Type II Restriction Endonuclease, *Bgl*II. (A) Schematic diagram of the catalytic mechanism demonstrating the utility of Mg2+ ions and polar amino acid residues within the active site to activate and position a water molecule for nucleophilic attack on the phosphodiester bond of the DNA substrate. (B) Crystal structure of the BglII dimer with double-stranded DNA and (C) Coordination of the Mg2+ cofactor within the active site of the BglII enzyme.

Figures from G Williams

10.5.2: CRISPR-Cas 9

The **CRISPR** (clustered regularly interspaced short palindromic repeats) operon was initially discovered as part of the adaptive immune system of bacteria and archaea, which must defend themselves against viruses (bacteriophages) and unwanted plasmids transferred from both bacteria. It would be ideal for bacteria to recognize previous exposure to viruses and their nucleic acids as the basis of their immunological memory system. Given the tendency of viral DNA to integrate into the host genome (which allows later transcription and translation of the viral genes in the process of new virus production), immunological memory could be based on that integrated viral DNA. Without going into detail, viral DNA can be integrated between two direct repeats in the bacterial genome. DNA from different viruses from previous exposures is also incorporated in the same fashion. One site of integration is the CRISPR operon. The DNA of the CRISPR operon contains both protein-coding and noncoding regions, which are transcribed and processed to form at least three RNA molecules, as shown in Figure (\PageIndex{24}\) below.

- a coding Cas 9 mRNA this is translated to produce the **Cas** 9 (**C**RISPR-**as**sociated protein);
- a noncoding cr-RNA (CRISPR RNA)

- a noncoding tracr-RNA (trans-activating CRISPR RNA)

Figure (\PageIndex{24}\): DNA of the CRISPR operon

The two mature noncoding RNAs eventually associate to form a binary complex. When using CRISPR-Cas 9 in eukaryotic gene editing applications, the two noncoding RNAs are covalently combined into one large **s**ynthetic **g**uide RNA (**sgRNA**), described later in this section. The Cas 9 protein is an endonuclease that cleaves both strands of bound target dsDNA in a blunt-end fashion at specific sequences. This occurs after the DNA binds to two arginines (1333 and 1335) in Cas9 through a short (3-5+ bases) recognition **protospacer adjacent motif** (**PAM**) located three base pairs from the cleavage site. The DNA must also bind in a complementary and specific fashion to the **protein-bound** noncoding cr+tracr-RNAs (or a single sg-RNA molecule for gene editing applications). The binding and cleavage of target DNA would render the DNA from an invading bacteriophage inactive.

Basic research into the bacterial CRISPR system has led to revolutionary and explosive eukaryotic applications of this gene editing system. The hope is that CRISPR technology will give us a precise and incredibly cheap way to do gene therapy in diseased cells and organisms. Given its role in transforming our ability to edit the genome and potentially cure genetically based diseases, we will explain its mechanism.

We have discussed the structure and function of many proteins. Protein enzymes are key to life as they catalyze almost all biological reactions. Most key enzymes are regulated. The activity of Cas 9 must be carefully controlled. Think of the consequences if the enzyme cleaves promiscuously at off-site targets! This section will help you understand several critical features of this enzyme:

1. How does the enzyme find its correct target site, a 20-nucleotide DNA sequence, and a proximal PAM site, among all the possible alternative sites? Think of how many PAM sequences must be in the host DNA genome!
2. How can the enzyme be "turned" on when it finds its target site and remains off when free, but more importantly, when bound off-site?

First, we will discuss the apo-form of the enzyme without bound substrate and RNA.

Apo- and Holo-Cas 9

This section will focus on the Type II-A Cas9 from *Streptococcus pyogenes* (SpyCas9 or SpCas9). Cas 9 is an endonuclease that cleaves both strands of DNA 3 base pairs from a DNA motif, NCC/NGG, called PAM. It has two distinct lobes. The nuclease lobe (NUC), amino acids 1-56 and 718-1368, has two different nuclease domains for the two cleavages. The recognition or receptor lobe (REC), amino acids 94-717, interacts with the RNA molecules. There is also an arginine-rich bridge helix (57-93).

The enzyme has two catalytic nuclease domains:

- HNH-like nuclease domain cleaves the "target" DNA strand, which is complementary to the RNA that confers specificity to the enzyme. The key catalytic residues are His 840 and Asn 854. It also contains a Mg ion;
- Ruv-like domain cleaves the complementary "non-target" strand with key active site residues Asp 10, Glu 762, Asp 986, and His 983. It also contains a bound Mn ion. The two lobes are separated by two linkers, amino acids 712-717, and an arginine-rich bridge (basic helix - BH), amino acids 628-658.

The overall structure of the apoenzyme (without bound RNA and DNA,pdb id 4cmp) is shown in Figure (\PageIndex{25}\) below, which shows the NUC domain (light blue) with the two catalytic domains (HNH and Ruv), the REC domain (orange) and the BH helix (red).

Figure (\PageIndex{25}\): Apoenyzme Cas9 (without bound RNA and DNA (4cmp)

A close up view showing the two catalytic sites is shown in Figure (\PageIndex{26}\) below.

Figure (\PageIndex{26}\): Two catalytic sites in Cas9

Figure 10.5.27 shows an interactive iCn3D model of Streptococcus pyogenes Cas9 in complex with guide RNA and target DNA (4OO8) (long load time). The Cas9 enzyme is shown as a gray transparent surface with an underlying cartoon rendering. The DNA is shown as colored sticks. The RNA is shown as a cyan cartoon.

Figure 10.5.27: **Streptococcus pyogenes Cas9 in complex with guide RNA and target DNA (4OO8)**. (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RjzJBFVt5qRjS7 (long load time)

A comparison of the crystal structure of the apo-Cas 9 and the ternary Cas 9: sgRNA:DNA target strand complex shows a significant conformational change on binding nucleic acids. The structure of the holoenzyme (ternary complex) is shown in Figure (\PageIndex{28}\) below.

Figure (\PageIndex{28}\): Structure of the holoenzyme Cas9 (ternary complex)with bound guide RNA and DNA

The extent of the conformation change between apo- and holo-Cas 9 enzymes can be seen by examining the distance between D435 and E 944/945 in Figure (\PageIndex{29}\) below. The importance of this change will be described later.

Figure (\PageIndex{29}\): Distance between D435 and E 944/945 going from the apo-Cas9 (left) to the holo-Cas 9 (right) enzyme

Figure (\PageIndex{30}\) below shows the pathway from the transcription of the relevant CRISPR genes (coding and noncoding) to the assembly of the ternary complex and the blunt end cut of the target DNA strand three nucleotides from the PAM sequence.

Figure (\PageIndex{30}\): Pathway from the transcription of the relevant CRISPR genes (coding and noncoding) to the assembly of the ternary complex and the blunt-end cut of the target DNA strand three nucleotides from the PAM sequence

Figure (\PageIndex{31}\) below shows an expanded view of the ternary complex.

Figure (\PageIndex{31}\): Expanded view of the ternary complex of Cas9 with guide RNA and DNA

Mechanism of DNA binding and cleavage

The above figures do not speak to the mechanism of the binding processes that form the ternary complex. Kinetic and structural studies have been conducted to elucidate the mechanism of binding and cleavage and address the following questions:

- which binds first, the RNA or DNA?
- What are the consequences of the profound conformational changes on the formation of the ternary complex?

The **specificity** of target DNA binding depends both on enzyme:PAM DNA and enzyme:sgRNA (or tracr- and crRNA) interactions. It should seem improbable that the trinucleotide PAM DNA sequence (NGG in *S. pyogenes*), which interacts with a pair of arginines (R 1333, R 1335) through H-bonding, as shown in the images above, and other local sites in Cas 9 would provide the sole or even the majority of the binding interactions. Figure (\PageIndex{32}\) below shows the Args:PAM interaction (pdb code 4un3)

Figure (\PageIndex{32}\): Args:PAM interaction in holo-Cas9 (4un3)

Hence, it is most likely that RNA binds first. Indeed, it does with the tracRNA implicated in the recruitment of Cas and the crRNA providing specificity for target DNA binding. The resulting Cas9:RNA binary complex could then search the relevant DNA genome. That would include the DNA of the bacteriophage in viral infection or eukaryotic DNA if the CRISPR DNA operon with the genes for Cas 9 and an sgRNA was transfected into the eukaryotic cell. After RNA binding, the enzyme changes conformation and allows loose DNA binding through Cas 9: PAM interactions.

Studies have shown that the apo form can also bind DNA, but it does so loosely and indiscriminately. It dissociates quickly, and binding is affected by generic polyanions such as the glycosaminoglycan heparin, which indicates its nonspecific nature. Once bound, both off-target and target DNAs would then be surveyed. If a target DNA contained a PAM sequence, the complex would undergo another conformational change to position the HNH and Ruv nuclease catalytic residues and locally unwind the duplex DNA to make the blunt-end cuts.

Cas 9 binding to the PAM site would promote better interaction of the unwound DNA and the bound RNA. If no PAM were present, no catalytically effective Cas 9:target DNA would form. This prevents off-site cleavage. These allosteric changes and controls are vital to the function of the endonuclease. Here are some findings that support this proposed mechanism:

- the conformation of apo Cas 9 is catalytically inactive;
- on binding RNA to form a binary complex, Cas 9 undergoes a dramatic conformational change, mostly in the REC lobe. However, on binding DNA in a nonspecific fashion, the conformational changes are much smaller. This suggests that most changes in conformation occur before DNA binding. In a way, RNA acts as an allosteric activator of the enzyme (as well as the major source of binding specificity to target DNA). Conformational changes can be determined directly by comparison of crystal structures or spectral techniques such as fluorescence resonance energy transfer (FRET) between two different attached fluorophores.
- Cas 9: RNA interactions lead to ordering of the region of the RNA that interacts with the DNA PAM sequence and adjacent deoxynucleotides (a "seed sequence"), allowing the Cas 9:RNA complex to scan and interact with potential DNA targets with PAM sequences;
- Once a PAM site is found, conformational changes lead to unwinding of the dsDNA, which allows heteroduplex formation between the crRNA and the target DNA strand;
- since Cas 9 recognizes a variety of DNA target sequences (but of course, only a specific PAM sequence), the binding of the target sequence depends on the geometry, not the sequence, of the target DNA;
- since binding of off-target DNA to the Cas 9:RNA complex occurs but with very infrequent cleavage, binding and cleavage are very distinct steps;
- on specific DNA binding, the HNH catalytic site moves near the sessile DNA bond site. Crystal structures show that the active site His is not sufficiently close to facilitate cleavage, suggesting that binding a second metal ion (see below) may be necessary. Molecular dynamics studies show that the HNH domain is "remarkably plastic."

Figure (\PageIndex{33}\) below shows an animation that illustrates the relative conformational changes going from the apo Cas 9 to the binary Cas 9:sgRNA complex to the ternary Cas 9: sgRNA: target DNA complex. The NUC catalytic domain is shown in light blue, the REC (receptor or RNA binding domain) in orange, sgRNA in red, and the target DNA in green. Note again that on binding RNA to form a binary complex, Cas 9 undergoes a dramatic conformational change, mostly in the REC lobe. The pdb protein sequences shown were aligned using pdbEfold.

Figure (\PageIndex{33}\): Animation illustrating conformational changes going from the apo Cas 9 to the binary Cas 9:sgRNA complex to the ternary Cas 9: sgRNA: target DNA complex

A potential abbreviated catalytic mechanism for the Ruv nuclease domain is shown in Figure (\PageIndex{34}\) below. The red arrows indicate the second set of electron movements. His 983 acts as a general base to abstract a proton from the water, making it a more potent nucleophile. An intermediate trigonal bipyramidal phospho-intermediate is formed, and the preceding transition state is stabilized by the proximal Mg^{2+} ion (an example of electrostatic or metal ion catalysis). The magnesium is positioned through its interaction with negatively charged carboxyl groups of Asp 10, Glu 762, and Asp 986.

Figure (\PageIndex{33}\): Abbreviated catalytic mechanism for the Ruv nuclease domain of Cas9

A second metal ion might be recruited to the Ruv site to facilitate the cleavage of the DNA further. The HNH catalytic site has a structure (beta-beta-alpha) and a conserved His in common with a class of nucleases that require one metal ion. In contrast, the Ruv catalytic site does not have this common secondary structural motif. It has a critical histidine, which is a common feature found in endonucleases that use two metal ions.

CRISPR and Eukaryotic Gene Editing

How could blunt-end cutting of both DNA strands by Cas 9 lead to the holy grail of specific eukaryotic gene editing with no off-site effects? Cutting the DNA genome seems like a bad idea. It is potentially so bad that many DNA repair mechanisms have

evolved to fix the cut. These include homologous recombination. If corrective DNA is supplied, as well as the components of the CRISPR system, a cell could effectively add the corrective DNA after the double-stranded cut and repair a deleterious mutation. Consult a molecular biology textbook for more insight into homologous recombination.

Mutations in the PAM sequence prevent Cas9 nuclease activity. Hence, the NGG PAM sequence is vital for the above interactions and activities. This would seem to limit the utility of CRISPR-Cas 9 in eukaryotic gene editing until one realizes that the GG dinucleotide has a 5.2% frequency of occurrence in the human genome, corresponding to over 160 million occurrences. Even then, it might not occur in a desired gene target. Cas 9 nuclease from other bacteria extends the range of activity of the CRISPR/Cas system as they interact with other PAM sequences (NNAGAA and NGGNG for *S. thermophilus* and NGGNG for *N. meningitidis*). Likewise, mutations in the *S. pyogenes* PAM (NGG) have been made as well. A D1135E mutation retains but increases the specificity for the normal NGG PAM site. D1135V, R1335Q, and T1337R mutations alter the optimal PAM recognition site to NGAN or NGNG.

CRISPR editing can be easily used to knock out specific genes. In addition, if cells are transfected with a plasmid with many target sequences, the system can edit multiple genes in one experiment. This would be very useful in studies of diseases linked to multiple genes. Since the cost of CRISPR reagents (plasmids, RNAs) is so inexpensive, and the specificity of editing is so high, the great excitement about CRISPR use for gene editing in human disease and for modification of plant and fungal genomes is warranted.

Other systems have been developed to bind to and cleave a target DNA sequence. They typically contain a protein that binds to a specific DNA target and an associated endonuclease that cleaves within the target DNA site. Typical prokaryotic restriction enzymes bind to and cut at a specific nucleotide sequence (for example, EcoR1 cleaves at G/AATTC palindromic sequences) to form sticky ends. The protein itself binds to this DNA recognition site. Other examples are based on the structure of known transcription factors. Libraries of genetically engineered proteins with Zn finger DNA binding domains (designed for specific DNA target sequences) fused to endonucleases have been created. Other examples are proteins called TALENs (transcription activator-like effector nucleases). These are fusion proteins containing a TAL effector DNA-binding domain and a nuclease. In each case, a 3D-folded protein is the specific target DNA recognition molecule. Think how much easier it is to make a 1D-DNA recognition element, a simple linear RNA sequence, which would adopt the correct 3D structure on binding its complementary target.

One major problem in using CRISPR for gene editing must be solved: how to get the CRISPR components into the correct cells in an organism. In effect, it's the same problem faced by small drug designers, only the components are much larger. *Ex vivo* applications, when diseased cells are removed from the body, repaired by CRISPR, and then reinjected, are likely to have more success. In these cases, electroporation would allow the uptake of Cas 9 and the sgRNA. I*n vivo* therapy has included adeno-associated viruses in which Cas9 and sgRNA genes could be encapsulated. This technique, used for other gene delivery systems, can be tolerated immunologically. However, this system allows for continual gene expression, which is undesirable for gene editing. After an initial "fix" of a mutant gene, continued expression of the CRISPR-Cas 9 genes would increase the chances for off-target cutting. A more recent approach is to deliver the mRNA in artificial lipid nanoparticles that can be taken into cells. Once free and translated into protein and sgRNA inside the cell, gene editing can occur before the RNA and protein are degraded.

10.5.3: Summary

Chapter Summary: Restriction Enzymes and CRISPR-Cas9 Gene Editing

This chapter explores two revolutionary tools that have transformed molecular biology and biotechnology: the classical restriction endonucleases and the modern CRISPR-Cas9 gene editing system.

It begins by examining the fundamental principles of restriction enzymes, particularly the widely used Type II enzymes. These enzymes recognize specific palindromic DNA sequences and cleave both strands at precise sites, generating reproducible fragment patterns essential for gene cloning and analysis. The chapter details the restriction-modification system in prokaryotes, where bacteria and archaea protect their own genomes via DNA methylation while using these enzymes to degrade foreign DNA. Emphasis is placed on the catalytic mechanism—where divalent metal ions, usually Mg^{2+}, activate a water molecule to directly hydrolyze the phosphodiester bond with inversion of configuration—and on how enzyme nomenclature and classifications (including isoschizomers and neoschizomers) reflect evolutionary relationships and cleavage patterns.

The discussion then transitions to the CRISPR-Cas9 system, which has rapidly gained attention due to its unprecedented precision and versatility in gene editing. Originating as part of the adaptive immune system in bacteria, CRISPR-Cas9 utilizes a guide RNA

(either as separate crRNA and tracrRNA or as a fused single guide RNA) to direct the Cas9 endonuclease to a target DNA sequence adjacent to a protospacer adjacent motif (PAM). Detailed structural insights reveal that Cas9 consists of distinct lobes—the nuclease (NUC) lobe housing the HNH and RuvC catalytic domains, and the recognition (REC) lobe responsible for RNA binding. Binding of the guide RNA triggers significant conformational changes that "activate" Cas9, enabling it to specifically bind and cleave target DNA with blunt-end cuts.

Furthermore, the chapter outlines the critical factors that govern Cas9 specificity and activity, including the allosteric changes induced upon PAM binding and the kinetic distinction between nonspecific DNA association and target recognition. It also addresses practical considerations in applying CRISPR-Cas9 for eukaryotic gene editing, such as delivery strategies (e.g., viral vectors, lipid nanoparticles) and the importance of minimizing off-target effects.

In summary, this chapter integrates the biochemical principles underlying restriction enzymes and CRISPR-Cas9, illustrating how understanding enzyme mechanisms, structure-function relationships, and regulatory features has not only advanced basic molecular biology but also paved the way for transformative applications in disease diagnosis, treatment, and genetic engineering.

10.5.3.0.1: 7.4 References:

1. Wikipedia contributors. (2020, April 21). Nucleophile. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:39, April 26, 2020, from en.Wikipedia.org/w/index.php?title=Nucleophile&oldid=952368939
2. Oregon Institute of Technology (2019) Organic Chemistry II (Lund). In Libretexts. Retrieved 10:58 am, April 27, 2020 from: https://chem.libretexts.org/Courses/Oregon_Institute_of_Technology/OIT%3A_CHE_332_--_Organic_Chemistry_II_(Lund)
3. Wikipedia contributors. (2020, April 12). Bond cleavage. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:15, April 27, 2020, from en.Wikipedia.org/w/index.php?title=Bond_cleavage&oldid=950494652
4. Wikipedia contributors. (2020, February 24). Arrow pushing. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:25, April 27, 2020, from en.Wikipedia.org/w/index.php?title=Arrow_pushing&oldid=942438883
5. Wikipedia contributors. (2020, April 16). Acid dissociation constant. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:48, April 27, 2020, from en.Wikipedia.org/w/index.phptitle=Acid_dissociation_constant&oldid=951313744
6. Farmer, S., Reusch, W., Alexander, E., and Rahim, A. (2016) Organic Chemistry. Libretexts. Available at: https://chem.libretexts.org/Core/Organic_Chemistry
7. Ball, et al. (2016) MAP: The Basics of GOB Chemistry. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Introductory_Chemistry_Textbook_Maps/Map%3A_The_Basics_of_GOB_Chemistry_(Ball_et_al.)/14%3A_Organic_Compounds_of_Oxygen/14.10%3A_Properties_of_Aldehydes_and_Ketones
8. McMurray (2017) MAP: Organic Chemistry. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_(McMurry)
9. Soderburg (2015) Map: Organic Chemistry with a Biological Emphasis. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_With_a_Biological_Emphasis_(Soderberg)
10. Ophardt, C. (2013) *Biological Chemistry.* Libretexts. Available at:https://chem.libretexts.org/Core/Biological_Chemistry/Proteins/Case_Studies%3A_Proteins/Permanent_Hair_Wave
11. Soderberg, T. (2016) *Organic Chemistry with a Biological Emphasis*. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Organic_Chemistry_Textbook_Maps/Map%3A_Organic_Chemistry_with_a_Biological_Emphasis_(Soderberg)
12. Ball, et al. (2016) MAP: The Basics of General, Organic, and Biological Chemistry. Libretexts. Available at:https://chem.libretexts.org/Textbook_Maps/Introductory_Chemistry_Textbook_Maps/Map%3A_The_Basics_of_GOB_Chemistry_(Ball_et_al.)
13. Clark, J. (2017) Organic Chemistry. Libretexts. Available at:https://chem.libretexts.org/Core/Organic_Chemistry/Amides/Reactivity_of_Amides/Polyamides
14. Wikipedia contributors. (2018, December 28). Metabolism. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:28, December 29, 2018, fromen.Wikipedia.org/w/index.php?title=Metabolism&oldid=875751739
15. Ball, Hill, and Scott. (2012) Enzyme Activity, section 18.7 from the book Introduction to Chemistry: General, Organic and Biological (v1.0) retrieved on Dec 31, 2018 fromhttps://2012books.lardbucket.org/books/introduction-to-chemistry-general-organic-and-biological/s21-07-enzyme-activity.html

16. Wikipedia contributors. (2018, November 29). Mechanism of action. In *Wikipedia, The Free Encyclopedia*. Retrieved 05:00, January 1, 2019, fromen.Wikipedia.org/w/index.php?title=Mechanism_of_action&oldid=871201209
17. Mótyán, J.A., Tóth, F., and Tőzsér, J. (2013) Research Applications of Proteolytic Enzymes in Molecular Biology. *Biomolecules 3*(4), 923-942; https://doi.org/10.3390/biom3040923
18. Wikipedia contributors. (2020, April 11). Adenylate kinase. In *Wikipedia, The Free Encyclopedia*. Retrieved 19:28, May 4, 2020, from en.Wikipedia.org/w/index.php?title=Adenylate_kinase&oldid=950311736
19. Ahern, K., Rajagopal, I., and Tan, T. (2019) Biochemistry Free and Easy. Available at Oregon State University (http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy) and Libretexts (https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Free_For_All_(Ahern%2C_Rajagopal%2C_and_Tan)/04%3A_Catalysis/4.03%3A_Mechanisms_of_Catalysis)
20. Wikipedia contributors. (2020, April 16). Serine protease. In *Wikipedia, The Free Encyclopedia*. Retrieved 14:32, May 6, 2020, from en.Wikipedia.org/w/index.php?title=Serine_protease&oldid=951309456
21. Wikipedia contributors. (2020, April 16). Restriction enzyme. In *Wikipedia, The Free Encyclopedia*. Retrieved 15:12, May 16, 2020, from en.Wikipedia.org/w/index.php?title=Restriction_enzyme&oldid=951351229
22. Pingoud, A., Wilson, G.G., and Wende, W. (2014) Type II restriction endonucleases - a historical perspective and more. *Nuc Acids Res* 42(12)7489-7527. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4081073/pdf/gku447.pdf
23. Wikipedia contributors. (2019, July 25). BglII. In *Wikipedia, The Free Encyclopedia*. Retrieved 20:48, May 16, 2020, from en.Wikipedia.org/w/index.php?title=BglII&oldid=907885716
24. De la Peña, M, GarcÍa-Robles, I., and Cervera, A. (2017) The Hammerhead Ribozyme: A Long History for a Short RNA. *Molecules* 22(1):78. Retrieved from: https://www.mdpi.com/1420-3049/22/1/78/htm
25. Jakubowski, H. (2019) Biochemistry Online. *Libretexts*. Available at: https://bio.libretexts.org/Bookshelves/Biochemistry/Book%3A_Biochemistry_Online_(Jakubowski)

This page titled 10.5: Enzymes for Genetic modifications is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **8.6: Enzymes for Genetic modifications** by Henry Jakubowski and Patricia Flatt has no license indicated.

10.6: DNA Isolation, Sequencing, and Synthesis

Learning Goals

1. **Differentiate Genomic DNA and cDNA:**
   - Explain the differences between genomic DNA (gDNA) and complementary DNA (cDNA), including gene structure (introns vs. exons) and their implications for gene expression and cloning strategies.
2. **Trace the Flow of Genetic Information:**
   - Describe the central dogma in eukaryotic cells by outlining how genomic DNA is transcribed into pre-mRNA, processed into mature mRNA, and how cDNA is synthesized from mRNA via reverse transcription.
3. **Contrast Prokaryotic and Eukaryotic Gene Organization:**
   - Compare the organization of genes in eukaryotes (with introns, UTRs, promoters, and regulatory elements) versus prokaryotes (often organized as operons and polycistronic transcripts).
4. **Master DNA Extraction and Purification Techniques:**
   - Identify and explain various methods for DNA extraction and purification, such as phenol/chloroform extraction and silica gel chromatography, and understand the chemical basis for DNA adsorption and elution.
5. **Understand DNA Sequencing Methodologies:**
   - Outline the evolution of DNA sequencing from early methods (e.g., Maxam-Gilbert) to Sanger sequencing (chain-termination) and its automation, and discuss the principles and advantages of next-generation sequencing (NGS) platforms including Illumina, Roche 454, Ion Torrent, nanopore, and SMRT sequencing.
6. **Evaluate the Impact of Sequencing Technologies:**
   - Assess how advances in sequencing have reduced costs, increased throughput, and improved accuracy, thereby revolutionizing genomics, diagnostics, and evolutionary studies.
7. **Apply Principles of Polymerase Chain Reaction (PCR):**
   - Explain the fundamental steps of PCR (denaturation, annealing, extension), the role of primers and thermostable DNA polymerases, and understand the exponential amplification of target DNA, including applications in diagnostics and research.
8. **Differentiate RT-PCR from qPCR:**
   - Distinguish between reverse-transcription PCR (RT-PCR), which converts RNA into cDNA for amplification, and real-time (quantitative) PCR (qPCR), which allows for quantitative analysis of gene expression.
9. **Explore DNA Synthesis and Gene Optimization:**
   - Describe chemical DNA synthesis methods, including oligonucleotide synthesis via phosphoramidite chemistry, and understand how gene synthesis and codon optimization are used to improve protein expression in heterologous systems.
10. **Connect Synthetic Biology and Vaccine Development:**
    - Discuss how advances in gene synthesis have enabled the creation of synthetic genes, complete synthetic plasmids, and DNA vaccines, and relate these technologies to modern applications such as synthetic influenza vaccines and RNA-based COVID-19 vaccines.
11. **Understand RNA Vaccine Design and Modification:**
    - Explain the rationale for modifying RNA (e.g., replacing uracil with pseudouracil) to enhance stability and reduce immunogenicity, and describe how lipid nanoparticle encapsulation facilitates efficient delivery of RNA vaccines.
12. **Analyze the Biosynthesis of RNA Modifications:**
    - Detail the post-transcriptional modification of uridine to pseudouridine by pseudouridine synthases and its impact on RNA structure, function, and vaccine efficacy.

These learning goals aim to provide a comprehensive framework for understanding how genomic and complementary DNA techniques underlie modern molecular biology, diagnostics, and therapeutic applications.

10.6.1: Genomic and complementary DNA

The ability to sequence an organism's DNA has revolutionized our understanding of biology and evolution. DNA can be isolated directly from living, dead, and even extinct species, and the "primary" sequence of A, G, C, and T bases in the molecule is determined. We can read (sequence), write (synthesize), and edit (mutate) DNA at will. Before we explore how to purify, sequence (read), and synthesize (write) DNA, it's important to differentiate two different types of DNAs, genomic and **complementary DNA** (cDNA), which is made by reverse transcribing messenger RNA into DNA. Since mRNA has no nucleotides encoded by introns, c-DNA provides just the coding sequences for protein.

Genomic deoxyribonucleic acid (gDNA) is chromosomal DNA, which does not include the extra-chromosomal DNA found in the mitochondria of eukaryotes or plasmids in bacteria (plasmids will be discussed in more detail in section 5.3 during the discussion of gene cloning and expression). Most organisms have the same genomic DNA in every cell (one exception is the genomic DNA for antibodies in B cells and T cells receptors in T-cells that have been altered as the cells become more terminally differentiated. It is also important to remember that only certain genes are active (expressed) in each cell. The subset of expressed genes is specific to a given differentiated cell and allows for the expression of specific cell functions. Liver cells, for example, don't express the gene for the protein opsin, which is expressed in retinal cells and is required for vision.

The genome of an organism (encoded by the genomic DNA) is the (biological) heredity information passed from one generation of organism to the next. That genome is transcribed to produce various RNAs, which are necessary for the function of the organism. RNA polymerase II transcribes precursor mRNA (pre-mRNA) in the nucleus. pre-mRNA is then processed by splicing to remove introns, leaving the exons in the mature messenger RNA (mRNA). Additional processing includes the addition of a 5' cap and a poly(A) tail to the pre-mRNA. The mature mRNA may then be transported to the cytosol and translated by the ribosome into a protein. Other types of RNA include ribosomal RNA (rRNA) and transfer RNA (tRNA). These types are transcribed by RNA polymerase II and RNA polymerase III, respectively, and are essential for protein synthesis. However, 5s rRNA is the only rRNA transcribed by RNA Polymerase III. ***cDNA is derived from mRNA, containing only exons but no introns.***

Figure **10.6.1** shows the flow of information stored in eukaryotic DNA and its eventual expression in mRNA.

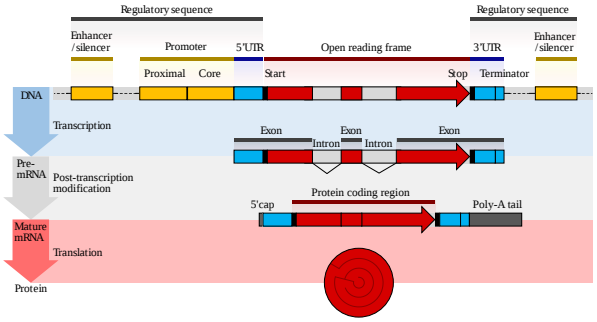


Figure **10.6.1**: Eukaryotic gene structure. Shafee T, Lowe R (2017). *WikiJournal of Medicine* **4** (1). DOI:10.15347/wjm/2017.002. ISSN 20024436.Wikimedia Commons

Red indicates coding exons, which are separated by gray introns. At the beginning and end of a gene sequence (encoded by three exons in the figure below) are 5' and 3' **untranslated regions (UTRs),** which are also transcribed and are represented in the mature mRNA. Also, there are potential regulatory sequences which are not transcribed (yellow) on both sides of the transcribable part of the gene. The 5'-end **promoter** is where transcription factors and RNA polymerase assemble before transcription starts. In addition, there are regulatory **enhancers** and **silencers** more distal from the gene sequences. An open reading frame (ORF) is a region of the DNA that can be decoded into a mRNA and doesn't have a stop signal (codon) in it that would prematurely terminate transcription.

In contrast, **complementary DNA** (**cDNA**) is synthesized from a single-stranded RNA (e.g., messenger RNA (mRNA) or microRNA) template in a reaction catalyzed by the enzyme reverse transcriptase. Reverse transcriptase is an enzyme found in retroviruses such as HIV with RNA as their core genetic material. Upon entering the host cell, the RNA is reverse-transcribed to produce a copy of the cDNA that can then be integrated into the host's genomic DNA. In biotechnology, reverse transcriptase is often used to create cDNA from the mRNA expressed in specific cells or tissues. In this way, the eukaryotic genes can be cloned without any introns housed in the structure. This is especially useful if the goal is to express the protein in a prokaryotic (bacterial) host. Recall that bacterial DNA contains no intron sequences within its chromosomal DNA. Thus, if you are using a prokaryotic system to express eukaryotic proteins, you must use cDNA, as the prokaryotic system will not be able to remove intron sequences following gene transcription. The term *cDNA* is also used, typically in a bioinformatics context, to refer to an mRNA transcript's sequence expressed as DNA bases (GCAT) rather than RNA bases (GCAU).

The gene organization of prokaryotes is different in that they don't have introns. In addition, some genes for a common pathway, for example, are continuous in the DNA. These stretches of DNA are called operons. Transcription from an operon produces a polycistronic RNA transcript. The words cis and are used in chemistry to describe R groups on the same size (cis) or opposite sides (trans) of a double bond. In DNA, cis-elements are in a single DNA section, while trans-elements usually refer to proteins (away from the gene) binding to the DNA. Hence, the term polycistronic is used for bacterial operons (with multiple genes sequentially arranged in the DNA sequence). Figure **10.6.2** shows the organization of prokaryotic gene structure.

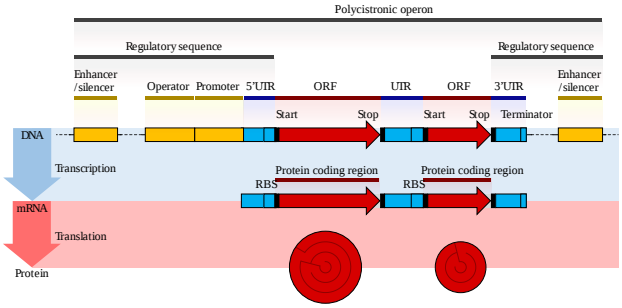


Figure **10.6.2**: Prokaryotic gene structure. Shafee T, Lowe R (2017). "Eukaryotic and prokaryotic gene structure". *WikiJournal of Medicine* **4** (1). DOI:10.15347/wjm/2017.002. ISSN 20024436.Wikimedia Commons

10.6.2: DNA Extraction/Purification

The first DNA isolation was done in 1869 by Friedrich Miescher. Now, purification kits are available from multiple manufacturers.

DNA can be isolated from whole tissue or cell cultures. Let's consider just DNA extraction from cells grown in the lab. Cells are collected by centrifugation and then treated with detergents like sodium dodecyl sulfate to lyse the cell membranes. Proteases and DNAase-free RNAase can be added to digest proteins and RNA.

Methods involving phenol/chloroform extractions:

In older methods, a mixture of phenol and chloroform or phenol/chloroform/isoamyl alcohol is used to extract DNA from the solution. Students who have performed liquid/liquid extractions in chemistry lab courses should recognize that this will form a biphasic mixture with water. Nonpolar substances like lipids and cellular debris will partition into the nonpolar phase (chloroform/phenol) or in the interface between them (as suspended insoluble material). Chloroform is very dense as it contains a chlorine atom). Phenol is somewhat soluble in water (8 g/100 g water) but very soluble in chloroform. In mixing during the extraction, the dissolved phenol alters the properties of water sufficiently enough to push the delicate equilibrium of proteins from the native to the denatured states, which aggregate and precipitate. On settling, DNA will remain in the aqueous phase. The use of chloroform/phenol in DNA extractions has a potential problem. Phenol (hydroxylated benzene) can lose one electron from the oxygen atom, forming a free radical that can be stabilized by resonance with pi electrons in the aromatic ring. Free radicals can damage DNA, so most new purification methods do **not** use phenol/chloroform extractions.

Most methods involve precipitating the extracted DNA at some point in the purification process using cold ice-cold ethanol or isopropanol. DNA is to a first approximation a long polyanion so it would be very difficult to purify "naked" DNA from solution since the extensive negative charges on the DNA would prevent aggregate and precipitate formation. This is not a problem if the ionic strength of the medium is sufficiently high so that positively charged counter-ions can shield the negative charges from each other, allowing precipitation.

Methods involving adsorption chromatography using silica gel: Nucleic acids bind or adsorb to a solid phase (silica or other) depending on the pH and the salt concentration of the buffer solution. Small spin columns are employed if small amounts are needed (such as for isolating recombinant plasmid from bacteria). This method relies on the fact that nucleic acid will bind to the solid phase of silica gel under certain conditions and then be released when those conditions are altered. These features are illustrated in Figure **10.6.3**.

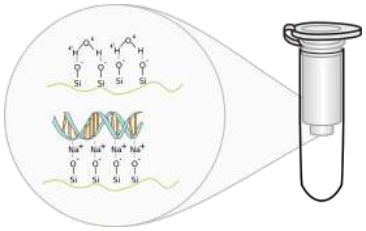


Figure 10.6.3: DNA-Silica gel interactions. *Image by Squidonius*

The DNA solution is applied to a small spin column containing silica gel and placed in a mini-centrifuge. On spinning, the nucleic acid will bind to the silica gel membrane as the solution passes through. After multiple "spin" washes to remove non-specific cellular components from the column, the DNA is eluted with a low salt elution buffer (or simply water). Unlike RNA, which degrades very quickly, DNA is quite stable and can be stored for long periods at -20°C.

Many student readers have used silica gel chromatography or silica gel thin-layer chromatography to separate and analyze organic mixtures. These techniques are usually performed in a mixture of organic solvents (e.g., hexane/ethyl acetate). Using silica gel to purify DNA from an aqueous solution might seem strange, so we will briefly explore how DNA binds.

In silica gel, each silicon atom is tetrahedrally bonded (sp^3 hybridization) to four oxygen atoms, with each oxygen atom covalently linked to two silicon atoms. At the particle's surface, the oxygen atoms are capped with H atoms, so the entire surface contains a "sea" of OHs, and therefore, hydrogen bond donors and acceptors. At lower pHs, some could ionize to form O^- ions. If salt concentrations are high enough, the percentage of O^- ions increases since they are stabilized by the cations in the salt (shifting the equilibrium to the ionized state).

DNA, a long negatively charged anion, can bind to the silica surface using two types of noncovalent interaction. It can form hydrogen bonds with the silica gel surface OHs. In addition, it can interact with the surface through ion-ion interactions mediated by bridging cations (like Na^+ from the high-concentration salt solution), as illustrated in the figure above. The binding solutions used in the spin column adsorption steps have high concentrations of chaotropic salts that disrupt water structure and hydrogen bonding. The salts also denature proteins and, in effect, dehydrate the DNA. Some chaotropic salts include sodium iodide, sodium perchlorate, guanidinium thiocyanate, and guanidinium chloride. The sodium or guanidinium acts as bridging cations, allowing the adsorption of the negative DNA to negative charges on the silica gel surface. Sodium acetate and Tris-HCl are included to buffer pH from 6-7. Now, it becomes easy to understand how pure water or low salt concentration solutions elute the bound DNA after extensive column washing since pure water or low salt solutions would strip the bound intermediary cations from the silica column.

After isolation, the DNA is dissolved in a slightly alkaline buffer, usually in a Tris-EDTA buffer or in ultra-pure water. EDTA binds divalent cations like Ca^{2+}, which activate nucleases. Modifications to these standard techniques are often done if the tissue being used is difficult to break down, if contaminants persist in the lysis solution that inhibit further reactions, or if the sample is extremely minimal, as is often the case in forensic investigations. In addition, different commercial kits will be tailored to isolate larger genomic DNA or smaller plasmid DNA.

The purity of a DNA preparation is usually determined by measuring the absorbance of the solution at 230, 260 (peak absorbance nucleic acids), and 280 nm (peak absorbance proteins), often using an instrument that requires a tiny droplet of solution. Figure (\PageIndex{4}\) below shows the relative absorbance spectra of proteins and nucleic acids.

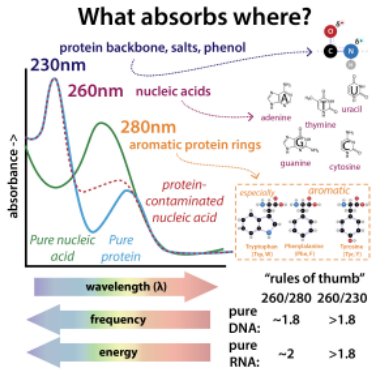


Figure (\PageIndex{4}\) relative absorbance spectra of proteins and nucleic acids. Brianna Bibel. The Bumbling Chemist https://thebumblingbiochemist.com/36...purity-ratios/

The A_{260}/A_{280} gives a measure of protein contamination, with a value around 1.8, indicating "pure" DNA. The A_{260}/A_{230} gives information on protein and other solution contamination. A pure solution of DNA with an A_{260} = 1.0 has a concentration of 50 µg/mL = 50 ng/µL.

When the temperature of a dsDNA solution is increased, the A_{260} increases in a certain range by about 37%. This is called the hyperchromic effect which occurs when the bases in DNA unstack on denaturation of dsDNA to single-stranded DNA as the intrastrand hydrogen bonds break. Figure (\PageIndex{5}\) below shows a graph of A_{260} vs temperature. The midpoint of the cooperative change in the absorbance signifies that the population of DNA molecules is 50% denatured.

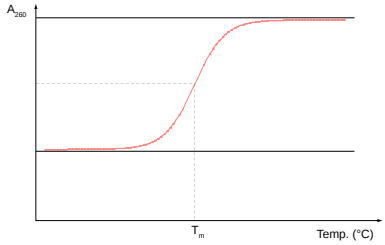


Figure (\PageIndex{5}\): Thermal denaturation curve of dsDNA. Fdardel - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/inde...?curid=5949679

Given the cooperative nature of unfolding, it is less likely that the population consists of individual molecules that are 50% denatured (i.e. a single molecule being 50% double-stranded and 50% single-stranded). Accordingly, a lower concentration (33 µg/mL) of fully single-stranded DNA gives an A_{260} = 1.

10.6.2.1: DNA Sequencing Techniques

DNA sequencing determines the order of nucleotides in a DNA sequence. It includes any method or technology used to determine the order of the four bases: adenine, guanine, cytosine, and thymine. The advent of rapid DNA sequencing methods has greatly accelerated biological and medical research and discovery.

Knowledge of DNA sequences has become indispensable for basic biological research and in numerous applied fields such as medical diagnosis, biotechnology, forensic biology, virology, and biological systematics. Comparing healthy and mutated DNA sequences can diagnose different diseases, including various cancers, characterize antibody repertoire, and can be used to guide patient treatment. Having a quick way to sequence DNA allows for faster and more individualized medical care to be administered and for more organisms to be identified and cataloged.

The rapid speed of sequencing attained with modern DNA sequencing technology has been instrumental in sequencing complete DNA sequences, or genomes, of numerous types and species of life, including the human genome and other complete DNA sequences of many animal, plant, and microbial species.

The first DNA sequences were obtained in the early 1970s by academic researchers using laborious methods based on two-dimensional chromatography. Following the development of fluorescence-based sequencing methods with a DNA sequencer, DNA sequencing has become easier and faster.

The canonical structure of DNA has four bases: thymine (T), adenine (A), cytosine (C), and guanine (G). DNA sequencing is the determination of the physical order of these bases in a molecule of DNA. However, epigenetic processes often modify DNA bases to control gene expression. Thus, more modified bases may be present in a DNA molecule than the standard four bases. For example, in some viruses (specifically, bacteriophages), cytosine may be replaced by hydroxymethyl- or hydroxymethylglucose cytosine. In eukaryotic DNA, variant bases with methyl groups or phosphosulfate may be found as shown in Figure (\PageIndex{6}\) below. Depending on the sequencing technique, a particular modification, e.g., the 5mC (5 -methylcytosine) common in humans, may or may not be detected.

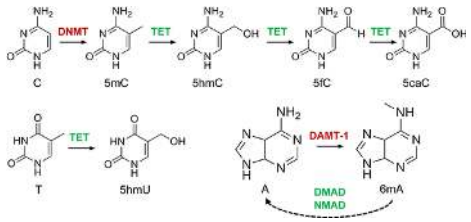


Figure (\PageIndex{6}\): **DNA modifications with epigenetic regulatory functions and their interdependencies.** Cytosine (C) is methylated to 5-methylcytosine (5mC) by DNA methyltransferases (DNMT) and then further oxidized to 5hmC, 5fC, and 5caC by Tet dioxygenases. 5-Hydroxyuracil (5hmU) is produced by Tet-catalysed oxidation of thymine (T). N6-methyladenine (6mA) is likely catalyzed by DNA N6 adenine methyltransferases (DAMT-1 in *C. elegans*), even though the biochemical activity of these enzymes remains to be characterized. The Tet-like ALKB enzymes NMAD (N6-methyl adenine demethylase 1) and DMAD (DNA 6mA demethylase) are involved in 6mA demethylation in *C. elegans* and *Drosophila*, respectively, possibly by using a conserved dioxygenase mechanism. *Image by Breiling and Lyko (2015) Epigenetics and Chromatin 8:24*

10.6.2.1.1: Early DNA sequencing methods

The first method for determining DNA sequences involved a location-specific primer extension strategy established by Ray Wu at Cornell University in 1970. DNA polymerase catalysis and specific nucleotide labeling, which figure prominently in current sequencing schemes, were used to sequence the cohesive ends of lambda phage DNA. Between 1970 and 1973, Wu, R Padmanabhan, and colleagues demonstrated that this method can be employed to determine any DNA sequence using synthetic location-specific primers. Frederick Sanger then adopted this primer-extension strategy to develop more rapid DNA sequencing methods at the MRC Centre, Cambridge, UK. He published a method for "DNA sequencing with chain-terminating inhibitors" in 1977. Walter Gilbert and Allan Maxam at Harvard also developed sequencing methods, including one for "DNA sequencing by chemical degradation." In 1973, Gilbert and Maxam reported the sequence of 24 base pairs using a method known as wandering-spot analysis. Sequence advancements were aided by the concurrent development of recombinant DNA technology, allowing DNA samples to be isolated from sources other than viruses.

Maxam-Gilbert sequencing requires radioactive labeling at one 5' end of the DNA and purification of the DNA fragment to be sequenced. Chemical treatment then generates breaks at a small proportion of one or two of the four nucleotide bases in each of the four reactions (G, A+G, C, C+T). The concentration of the modifying chemicals is controlled to introduce, on average, one modification per DNA molecule. Thus a series of labeled fragments is generated from the radiolabeled end to the first "cut" site in each molecule. The fragments in the four reactions are electrophoresed side by side in denaturing acrylamide gels for size separation. To visualize the fragments, the gel is exposed to X-ray film for autoradiography, yielding a series of dark bands, each corresponding to a radiolabeled DNA fragment, from which the sequence may be inferred.

The technical aspects of Maxam-Gilbert sequencing caused it to go out of favor once the Sanger sequencing method had been well established, as described below.

10.6.2.1.2: Sanger Sequencing Method

The chain-termination method Frederick Sanger and coworkers developed in 1977 soon became the method of choice, owing to its relative ease and reliability. When invented, the chain-terminator method used fewer toxic chemicals and lower amounts of radioactivity than the Maxam-Gilbert method. Because of its comparative ease, the Sanger method was soon automated and used in the first generation of DNA sequencers.

The classical chain-termination method requires a single-stranded DNA template, a DNA primer, a DNA polymerase, normal deoxynucleotide triphosphates (dNTPs), and modified di-deoxynucleotide triphosphates (ddNTPs), the latter of which terminate DNA strand elongation. These chain-terminating nucleotides lack a 3'-OH group required to form a phosphodiester bond between two nucleotides, causing DNA polymerase to cease the extension of DNA when a modified ddNTP is incorporated. The ddNTPs may be radioactively or fluorescently labeled for detection in automated sequencing machines.

The DNA sample is divided into four separate sequencing reactions, containing all four standard deoxynucleotides (dATP, dGTP, dCTP, and dTTP) and the DNA polymerase. To each reaction is added only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP), while the other added nucleotides are ordinary ones as shown in (\PageIndex{7}\) below.

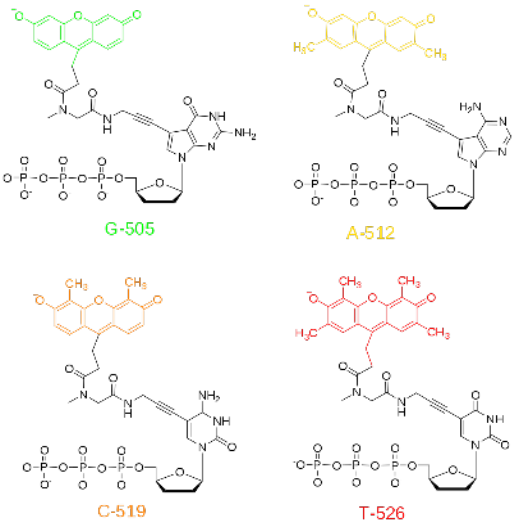


Figure (\PageIndex{7}\): **Fluorescent ddNTPs for Sanger Sequencing.** Dideoxynucleotides are utilized for sequencing as they cannot be extended further once incorporated into the nascent DNA. *Image by Fibonachi*

The dideoxynucleotide concentration should be approximately 100-fold lower than that of the corresponding deoxynucleotide (e.g., 0.005mM ddTTP : 0.5mM dTTP) to allow enough fragments to be produced while still transcribing the complete sequence. Four separate reactions are needed in this process to test all four ddNTPs. This is illustrated in Figure (\PageIndex{8}\) below.

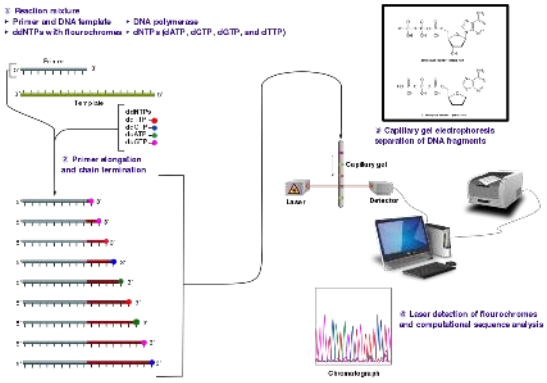


Figure (\PageIndex{8}\): **The Sanger (chain-termination) method for DNA sequencing.** (1) A primer is annealed to a sequence, (2) Reagents are added to the primer and template, including DNA polymerase, dNTPs, and a small amount of all four dideoxynucleotides (ddNTPs) labeled with fluorophores. During primer elongation, the random insertion of a ddNTP instead of a dNTP terminates the synthesis of the chain because DNA polymerase cannot react with the missing hydroxyl. This produces all possible lengths of chains. (3) The products are separated on a single-lane capillary gel, where the resulting bands are read by an imaging system. (4) This produces several hundred thousand nucleotides a day, data that require storage and subsequent computational analysis. *Image by Estevezj*

Following rounds of template DNA extension from the bound primer, the resulting DNA fragments are heat-denatured and separated by size using gel electrophoresis. This technique was frequently performed using a denaturing polyacrylamide-urea gel with each of the four reactions run in one of four individual lanes (lanes A, T, G, C). The DNA bands may then be visualized by autoradiography or UV light, and the DNA sequence can be directly read off the X-ray film or gel image, as shown in Figure (\PageIndex{9}\) below.

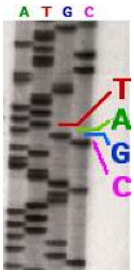


Figure (\PageIndex{9}\): **Traditional Sanger Sequencing Gel.** Sequence visualized by autoradiography. Each lane contains a single reaction that has all four regular nucleotides and a small amount of one of the dideoxynucleotides (ddNTPs). Over time, the ddNTPs will be incorporated at each position containing that specific nucleotide. The gel can then be read from the bottom to the top, as the smallest fragments (those fragments terminated the closest to the primer at the 5'-end) will run the farthest distance in the gel. The sequence of this fragment is 5'-TACGAGATATATGGCGTTAATACGATATATTGGAACTTCTATTGC-3'. *Image by John Schmidt*

Automation of the Sanger sequencing method was made possible when the shift from radioactively tagged nucleotides to fluorescently tagged nucleotides was made. In automated sequencing, capillary gel electrophoresis is performed rather than separating the samples using gel electrophoresis. The output from capillary electrophoresis is fluorescent peak trace chromatograms, as shown in **Figure** (\PageIndex{10}\) below. Automated DNA-sequencing instruments (DNA sequencers) can sequence up to 384 DNA samples in a single batch. Batch runs may occur up to 24 times a day, greatly enhancing the speed with which samples may be sequenced and analyzed. Common challenges of DNA sequencing with the Sanger method include poor quality in the first 15-40 bases of the sequence due to primer binding and deteriorating quality of sequencing traces after 400-500 bases.

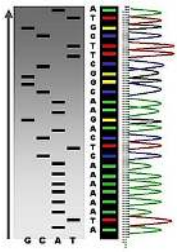


Figure (\PageIndex{10}\): **Side by Side Comparison of Gel Electrophoresis and Capillary Electrophoresis.** *Lefthand Diagram* shows the traditional autoradiogram of Sanger sequencing samples. The *Righthand Diagram* shows the same reactions using fluorescently tagged ddNTPs separated by capillary electrophoresis. The chromatogram output is shown on the far right. *Image by Abizar*

Sanger sequencing is the method that prevailed from the 1980s until ~2005. Over that period, great advances were made in the technique, such as fluorescent labeling, capillary electrophoresis, and general automation. These developments allowed much more efficient sequencing, leading to lower costs. In mass production form, the Sanger method is the technology that produced the first human genome in 2001, ushering in the age of genomics.

10.6.2.1.3: Microfluidic Sanger Sequencing

Microfluidic Sanger sequencing is a ***lab-on-a-chip application*** for DNA sequencing, in which the Sanger sequencing steps (thermal cycling, sample purification, and capillary electrophoresis) are integrated on a wafer-scale chip using nanoliter-scale sample volumes (Figure (\PageIndex{11}\)). This technology generates long and accurate sequence reads while obviating many of the significant shortcomings of the conventional Sanger method (e.g., high consumption of expensive reagents, reliance on expensive equipment, personnel-intensive manipulations, etc.) by integrating and automating the Sanger sequencing steps.

Figure (\PageIndex{11}\): **Lab-On-A-Chip Technologies.** Example of a microfluidic lab-on-a-chip device sitting on a polystyrene dish. Stainless steel needles inserted into the device serve as access points for fluids into small channels within the device, which are about the size of a human hair. *Image by National Institute of Standards and Technology (NIST)*

10.6.2.1.4: Next-generation sequencing (NGS)

Next-generation sequencing (NGS), or high-throughput sequencing, is the catch-all term used to describe several modern sequencing technologies. These technologies allow for the sequencing of DNA and RNA much more quickly and cheaply than the previously used Sanger sequencing and, as such, revolutionized the study of genomics and molecular biology. We present information from specific companies that have developed these new technologies without endorsements. Such technologies include:

Illumina Sequencing - In NGS, vast numbers of short reads are sequenced in a single stroke using the lab-on-a-chip technology described above. To do this, the input sample must be cleaved into short sections. In Illumina sequencing, 100-150 bp reads are used. Somewhat longer fragments are ligated to generic adaptors and annealed to a slide using the adaptors. PCR is carried out to amplify each read, creating a spot with many copies of the same read. They are then separated into single-stranded DNA to be sequenced, as shown in Figure (\PageIndex{12}\) below.

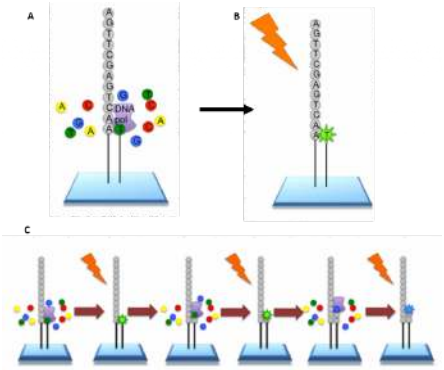


Figure (\PageIndex{12}\): **Procedure for Illumina Sequencing.** (A) The slide with PCR-amplified fragments of DNA is flooded with nucleotides and DNA polymerase. These nucleotides are fluorescently labeled, with each color corresponding to a specific base. The reactions also have a terminator present, so that only one base is added at a time. (B) An image is taken of the slide. In each reaction location, there will be a fluorescent signal indicating that a specific base has been added. (C) The data is recorded and the slide is then prepared for the next cycle. In preparation, the terminators are removed, which will allow the next base to be added, and the fluorescent signal is cleaved, preventing the fluorescent signal from contaminating the next image. The process is repeated, adding one nucleotide at a time (G, A, T, or C) and imaging in between. All of the sequence reads will be the same length as single bases are added at each cycle. *Image modified from EMBL-EBI*

Roche 454-Sequencing is similar to the Illumina process but can sequence much longer reads. Like Illumina, it does this by sequencing multiple reads simultaneously by reading optical signals as bases are added. As in Illumina, the DNA or RNA is fragmented into shorter reads, in this case, up to 1kb (1,000bp). Generic adaptors are added to the ends, and these are annealed to beads, one DNA fragment per bead. The fragments are then amplified by PCR using adaptor-specific primers. Each bead is then placed in a single well of a slide, with each well containing a single bead, covered in many PCR copies of a single sequence. The wells also contain DNA polymerase and sequencing buffers (Figure (\PageIndex{13}\).

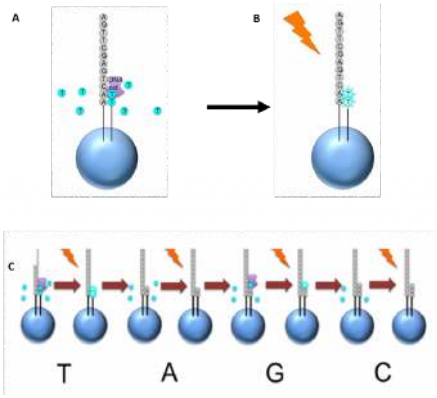


Figure (\PageIndex{13}\): **Procedure for Roche 454 Sequencing.** (A) Once the PCR product is attached to the bead, the slide is flooded with one of the four NTP species. Where this nucleotide is next in the sequence, it is added to the sequence read. If that single base repeats, then more will be added. So if we flood with Guanine bases, and the next in the sequence is G, one G will be added, however, if the next part of the sequence is GGGG, then four Gs will be added. (B) The addition of each nucleotide releases a light signal. These locations of signals are detected and used to determine which beads the nucleotides are added to. (C) The NTP mix is washed away. The next NTP mix is now added, and the process is repeated, cycling through the four NTPs. All of the sequence reads from 454 sequencings will be different lengths, because different numbers of bases will be added with each cycle. *Image modified by EMBL-EBI*

Newer technologies, such as the ***Ion Torrent Technology,*** detect sequence data using electrical signals on a semiconductor chip rather than optically reading dye-labeled nucleotides. This is possible as the addition of a dNTP to the DNA polymer causes the release of an H^+ ion (Figure (\PageIndex{14}\)). As in other kinds of NGS, the input DNA or RNA is fragmented, this time ~200bp. Adaptors are added, and one molecule is placed on a bead. The molecules are amplified on the bead by emulsion PCR. Each bead is placed into a single well of a slide. This is illustrated in Figure (\PageIndex{14}\) below.

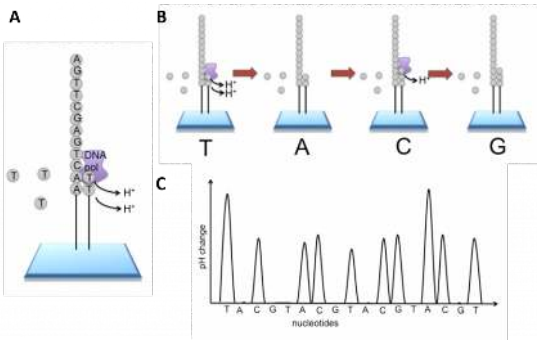


Figure (\PageIndex{14}\): **Ion Torrent Sequencing Technology.** (A) Similar to 454 sequencings, the slide is flooded with a single species of dNTP, along with buffers and polymerase. The pH is monitored in each of the wells after the addition of the specific dNTP. The pH will decrease when the dNTP is incorporated into the polymer, causing the release of a proton (H^+). The changes in pH allow the determination of different dNTP species. (C) The pH change, if any, is used to determine how many bases (if any) were added with each cycle. *Image modified from EMBL-EBI*

Nanopore Sequencing

In this technique, flow cells are constructed that contain nanopores in a nonlipid membrane that is resistant to current flow. When a voltage is applied across a membrane separating two salt solutions, a sensor chip can detect the current through each nanopore channel. When a larger molecule moves through the pore, a disruption in basal current occurs. Computer algorithms have been developed to detect base-specific changes in the current as the base (even chemically modified ones) moves through the membrane. The sequence is then decoded in real time.

Single-stranded DNA or RNA can be driven through the pore by a transmembrane potential in a process similar to electrophoresis. The enzyme DNA helicase, a motor protein, can be attached to the outer part of the pore protein. This enzyme binds to single-stranded DNA and moves along it in a process that requires ATP. If the helicase is attached to the pore protein, the single DNA would move, allowing control of its movement through the pore.

The nanopores are made of real membrane proteins (which we discuss in Chapter 11). One example is α-hemolysin, a heptamer with an inner pore diameter of 1 nm. When embedded in real cells, it can allow the flow of K^+ (diameter around 250 pm = 0.25 nm) and other ions across the cell membrane, changing the osmotic balance and lysing the cell. The pore size of the proteins used in nanopore sequencing allows single-stranded DNA to flow through it. An interactive iCn3D model of protein used for nanopore sequencing, Curli transport lipoprotein CsgG (4uv3), is shown in the membrane bilayer in Figure **10.6.15**.

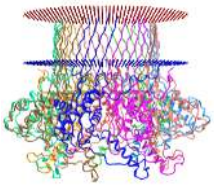


Figure **10.6.15**: Curli transport lipoprotein CsgG (4uv3) for DNA nanopore sequencing. (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LC5K3vSjYjtHN7

Nanopore technologies have enabled the production of small hand-held DNA sequencing devices that can be plugged into a USB drive on a laptop computer and utilized in the field under real-time collection conditions. Future modifications might replace protein pores with synthetic solid-state nanopores. For instance, the membrane might be made of graphene with pores of a specific size.

Figure **10.6.16** shows an animation of a single-stranded DNA moving through a protein pore (blue) assisted by a motor protein (magenta)

Figure 10.6.16: Single-stranded DNA moving through a protein pore (blue) assisted by a motor protein (magenta) Oxford Nanopore Tech. https://nanoporetech.com/how-it-works

Recent Updates: 12/6/2025

DNA sequencing using Expandomers

One problem with nanopore sequencing is the signal-to-noise ratio as each base transits the pore. The distance between bases in a strand is only 3.4 Å, so transit of each base through the pore is very quick, crowding the signal, causing lower resolution (i.e., lower signal-to-noise ratio). How to solve the problem? One way would be to increase the distance between the bases. This might at first seem ludicrous, since it would require a significant change in the structure of the single-stranded DNA passing through the pore and a new synthetic method to achieve it. Yet this feat has been accomplished by Mark Kokoris et al., using remarkable chemistry and genetic engineering.

They used click chemistry (See Chapter 7.5.6 Click Chemistry and Bioorthogonal Reactions) to attach a modified dNTP containing alkynes to a larger polymer to form an **expandomer**. Deoxynucleotides have an average molecular weight of 328 and a -1 charge for the monophosphate. In contrast, the expandomer has an average molecular weight of 16K and a charge of -38 (similar in size and charge to a deoxynucleotide of 35 base pairs. Figure 10.6.xx below shows (left) the mechanism of the copper-catalyzed azide-alkyne cycloaddition (CuAAC) Click Chemistry and (right) a click chemistry-modifiable dCTP

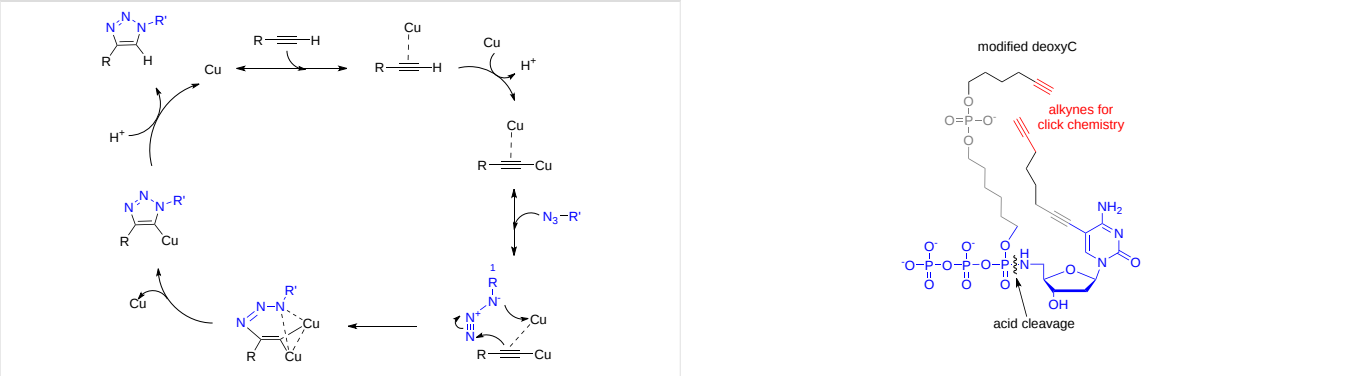


Figure 10.6.xx: (left) Mechanism of the Copper-Catalyzed Azide-Alkyne Cycloaddition (CuAAC) Click Chemistry and (right) click chemistry-modifiable dCTP. Left figure mechanism from Organic Chemistry Portal. https://www.organic-chemistry.org/na...chemistry.shtm.

A cartoon representation of a typical expandomer (for dCTP) is shown in Figure 10.6.xx below.

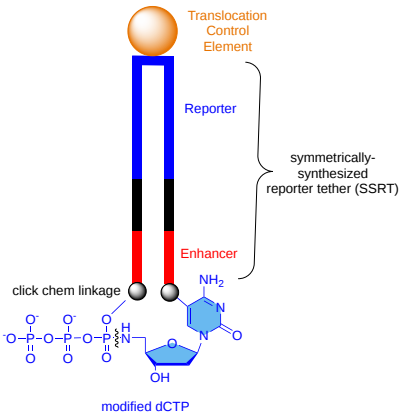


Figure 10.6.xx: A cartoon representation of a typical expandomer (for dCTP)

A chemical structure of a typical expandomer for X-dCTP is shown in Figure 10.6.xx below.

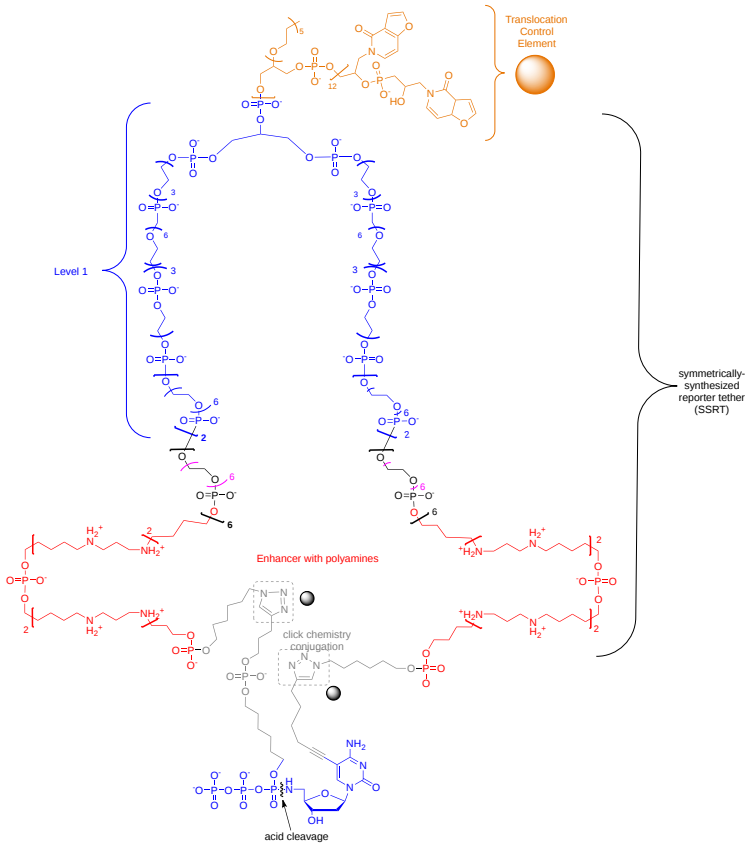


Figure **10.6.xx**: Chemical structure of an expandomer for dCTP. After Sequencing by Expansion (SBX) – a novel, high-throughput single-molecule sequencing technology, https://www.biorxiv.org/content/10.1....639056v2.full.

The covalent link between the P and N in the X-dNTP is broken by acid catalysis, allowing the macrocyclic, symmetrically synthesized report tether to be linearized and expanded, thereby allowing each expandomer to be threaded through the nanopore. Each expandomer base in the synthesized sequence is separated by one tether length.

The expandomer is made from the target DNA strand to be sequenced, which hybridizes to a solid phase **extension oligomer (EO)** from which the expandomer will be synthesized (in a 5' to 3' direction). Figure **10.6.xx** shows the extension oligo (EO) connected through a concentrator-leader-spacer to the solid support

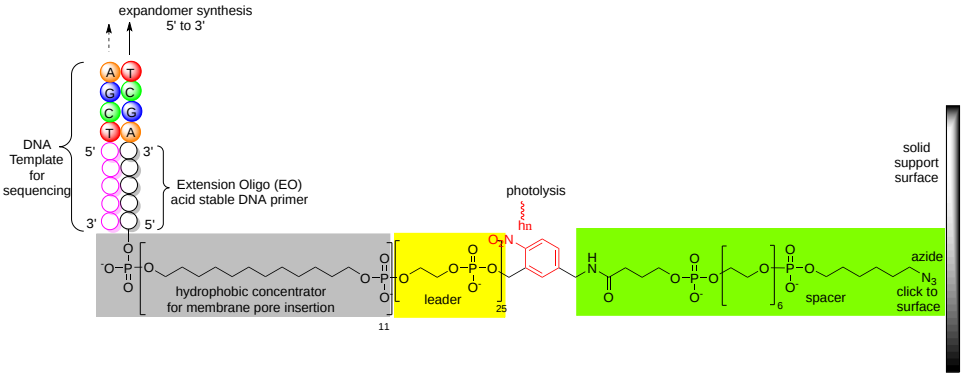


Figure **10.6.xx**: Extension oligo (EO) connected through a concentrator-leader-spacer to the solid support

The leader carries a high negative charge, which helps concentrate the structure on the positive potential side of the membrane pore. The concentrator is highly hydrophobic, facilitating interaction with the lipid bilayer. Both features dramatically increase the diffusion of the EO structure to the nanopore. After the entire complementary strand is synthesized, the expandomer can be cleaved from the solid support by photolysis.

The genetically engineered **expandomer synthase** (XP synthase), a mutated form (36 amino acid substitutions) of Dpo4 polymerase, a DNA polymerase IV from Saccharolobus solfataricus (a hyperthermophile), is used. It has a more open active site that can accommodate even damaged bases, such as thymine dimers. It reads just short stretches of DNA, so it is distributive, not processive (reading long lengths before the polymerase dissociates from the chain), and has a very high error rate. Hence, it can accommodate the large expanded version of each incoming deoxynucleotide in its active site.

Let's look in more detail at the components of the expandomer.

Translocation Control Element (TCE)

The structure in Figure xx above shows that the TCE has a more hydrophobic polyethylene glycol chain (-O-CH_2-CH_2-O-) and a polyphosphate polyanionic chain. It is likely that at low voltage, the hydrophobic chain folds over and closes the pore, whereas at higher transpore voltage, it moves through the pore. With repeated high-voltage pulses, the expandomer ratchets through the chain, one

expandomer base section per pulse.

Base Reporter

The TCE likely positions one arm of the **Base Reporter** in the pore barrel, thereby inhibiting ion flow through the pore to an extent determined by the specific base in the XNTP and the exact structure of the reporter used for each base.

Enhancer

The polyamine repeat of the enhancer confers a significant net positive charge, which promotes interaction between the expandomer and XP synthase. Mutations in the XP synthase increase the negative charge density at the open binding site of the synthase.

Here is a link to an animation showing DNA sequencing using expandomers. The method and the video are from Roche. Expandomer sequencing was recently used to sequence an entire human genome in less than 4 hours!

Single Molecule, Real-Time (SMRT) sequencing

In this technique, either RNA or DNA is converted to dsDNA. Deoxynucleotide "adapters" are added to connect the 5'-end of strand 1 to the 3'-end of strand 2 and another adapter to connect the 3'-end of strand 1 to the 5'-end of strand 2, resulting in a "circular" ss-DNA molecule. This single molecule is then drawn into a nanophotonic nanowell made in a thin metal film deposited on glass. The dimensions of each well allow only a single circular ss-DNA molecule. Hundreds of different circular ss-DNA molecules entering individual wells are shown in Figure **10.6.17**. The blue stretches represent the adapters.

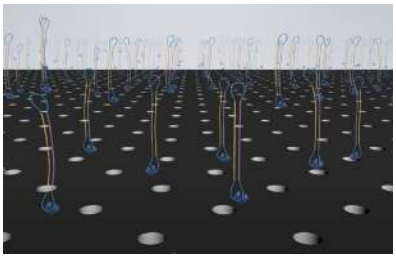


Figure **10.6.17**. Nanowells for SMRT sequencing. https://www.youtube.com/embed/_lD8JyAbwEo

The wells are approximately 100 nm in diameter. DNA polymerase and dNTPs can be added to the nanowalls, which contain a single molecule of immobilized circular ssDNA. The DNA is immobilized by its biotinylated or attachment to magnetic beads which interact with streptavidin-coated wells. When confined to the wells, the apparent concentration of reactants for the polymerization can be quite high, allowing robust DNA polymerase activity.

DNA sequencing using real-time fluorescence monitoring can be done in a massively parallel fashion. The fluorophore is connected to the terminal phosphate of the dNTP. When the DNA polymerase makes a phosphodiester bond, the fluorophore is released as a leaving group, resulting in a natural, unmodified DNA that continues to grow. The YouTube video below shows the entire process of a single molecule, real-time sequencing.

The four main advantages of Next Generation Sequencing (NGS) over classical sequencing are described below.

Sample size

NGS is cheaper, quicker, needs significantly less DNA, and is more accurate and reliable than Sanger sequencing. Let us look at this more closely. For Sanger sequencing, a large amount of template DNA is needed for each read. Several strands of template DNA are needed for each base being sequenced (i.e., for a 100bp sequence, you'd need many hundreds of copies; for a 1000 bp sequence, you'd need many thousands of copies), as a strand that terminates on each base is needed to construct a full sequence. In NGS, a sequence can be obtained from a single strand. In both kinds of sequencing, multiple staggered copies are taken for contig construction and sequence validation.

Speed

NGS is quicker than Sanger sequencing in two ways. Firstly, the chemical reaction may be combined with the signal detection in some versions of NGS, whereas in Sanger sequencing, these are two separate processes. Secondly and more significantly, only one read (maximum ~1kb) can be taken at a time in Sanger sequencing. In contrast, NGS is massively parallel, allowing 300Gb of DNA to be read on a single run on a single chip.

Cost

The reduced time, people power, and NGS reagents mean the costs are much lower. The first human genome sequence cost in the region of $2.7 billion in 2003. Using modern Sanger sequencing methods, aided by data from the known sequence, a full human genome still cost $300,000 in 2006. Sequencing a human genome with NGS today costs roughly $1,000.

Accuracy

Repeats are intrinsic to NGS, as each read is amplified before sequencing, and because it relies on many short overlapping reads, each section of DNA or RNA is sequenced multiple times. Also, because it is much quicker and cheaper, it is possible to repeat more than Sanger sequencing. More repeats mean greater coverage, which leads to a more accurate and reliable sequence, even if

individual reads are less accurate for NGS.

The nanopore and single molecule real-time (SMRT) sequencing were recently employed to complete the sequence for the full human genome (2022). Previous genomic sequences were missing regions with highly repetitive sequences at the centromeres and telomeres. The "Telomere-to-Telomere (T2T) Consortium performed the analysis, adding 200 megabases of new sequence information missing from the previous best sequence.

10.6.2.2: DNA Synthesis Techniques

DNA synthesis is the natural or artificial creation of deoxyribonucleic acid (DNA) molecules. The term DNA synthesis can refer to ***DNA replication*** (which will be covered in more detail in Chapter XX), ***polymerase chain reaction (PCR),*** or ***gene synthesis*** (physically creating artificial gene sequences).

10.6.2.2.1: Polymerase Chain Reaction (PCR)

The Polymerase chain reaction (PCR) is widely employed in the basic and biomedical sciences. PCR is a laboratory technique utilized to amplify specific segments of DNA for a wide range of laboratory and/or clinical applications. Building on the work of Panet and Khorana's successful amplification of DNA *in-vitro*, Kary Mullis and coworkers developed PCR in the early 1980s, having been met with a Nobel prize only a decade later. Allowing for more than the billion-fold amplification of specific target regions, it has become instrumental in many applications, including cloning genes, the diagnosis of infectious diseases, and screening prenatal infants for deleterious genetic abnormalities.

Fundamentals

The main components of PCR are a template, primers, free nucleotide bases, and the DNA polymerase enzyme. The ***DNA template*** contains the specific region you wish to amplify, such as the DNA extracted from a piece of hair. ***Primers***, or oligonucleotides, are short strands of single-stranded DNA complementary to the 3' end of each target region. A forward and a reverse primer are required, one for each complementary strand of DNA. ***DNA polymerase*** is the enzyme that carries out DNA replication. Thermostable analogs of DNA polymerase I, such as Taq polymerase, which was originally found in a bacterium that grows in hot springs, are a common choice due to their resistance to the heating and cooling cycles necessary for PCR.

PCR takes advantage of the complementary base pairing, double-stranded nature, and melting temperature of DNA molecules. This process involves cycling through 3 sequential rounds of temperature-dependent reactions: DNA melting (denaturation), annealing, and enzyme-driven DNA replication (elongation). ***Denaturation*** begins by heating the reaction to about 95 ^{o}C, disrupting the hydrogen bonds that hold the two strands of template DNA together. Next, the reaction is reduced to around 50 to 65 ^{o}C, depending on the physicochemical variables of the primers, enabling the ***annealing*** of complementary base pairs. The primers, which are added to the solution in excess, bind to the beginning of the 3' end of each template strand and prevent re-hybridization of the template strand with itself. Lastly, enzyme-driven DNA replication, or ***elongation***, begins by setting the reaction temperature to the amount that optimizes DNA polymerase's activity, which is around 75 to 80 ^{o}C. At this point, DNA polymerase, which needs double-stranded DNA to begin replication, synthesizes a new DNA strand by assembling free nucleotides in solution in the 3' to 5' direction to produce two full sets of complementary strands. The newly synthesized DNA is now identical to the template strand and will be used as such in the progressive PCR cycles. The steps in PCR are animated in the video below.

Figure (\PageIndex{17}\) below illustrates the step involved in PCR amplification of target DNA.

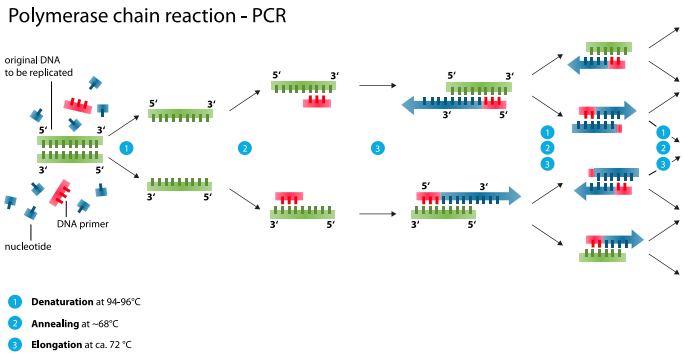


Figure (\PageIndex{17}\): Steps involved in PCR amplification of target DNA.

Figure (\PageIndex{18}\) shows a video animation of a PCR reaction.

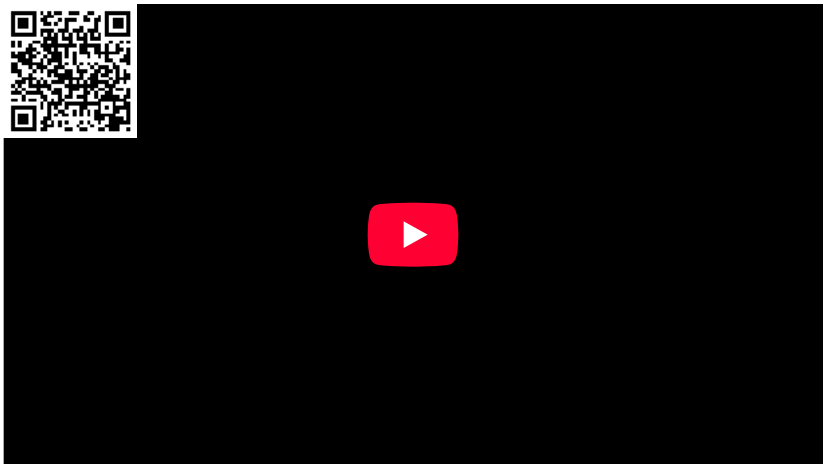


Figure (\PageIndex{18}\): Video animation of a PCR reaction.

Given that previously synthesized DNA strands serve as templates, DNA amplification using PCR increases exponentially, where the copies of DNA double at the end of each replication step. The exponential replication of the target DNA eventually plateaus around 30 to 40 cycles mainly due to reagent limitation, but can also be due to inhibitors of the polymerase reaction found in the sample, self-annealing of the accumulating product, and accumulation of pyrophosphate molecules.

10.6.2.2.2: Real-Time PCR

At its advent, PCR technology was limited to qualitative and semi-quantitative analysis due to limitations on the ability to quantify nucleic acids. At that time, the DNA product was separated by size via agarose gel electrophoresis to verify if the target gene was amplified successfully. Ethidium bromide, a molecule that fluoresces when bound to dsDNA, could give a rough estimate of DNA amount by roughly comparing the brightness of separated bands, but was not sensitive enough for rigorous quantitative analysis.

Improvements in fluorophore development and instrumentation led to thermocyclers that no longer required measurement of only end-product DNA. This process, known as ***real-time PCR,*** or ***quantitative PCR (qPCR),*** has allowed for detecting dsDNA during amplification. qPCR thermocyclers can excite fluorophores at specific wavelengths, detect their emission with a photodetector, and record the values. The sensitive collection of numerical values during amplification has strongly enhanced quantitative analytical power.

Two main types of fluorophores are used in qPCR: those that bind specifically to a given target sequence and those that do not. The sensitivity of fluorophores has been an important aspect of qPCR development. One of the most effective and widely used non-specific markers, SYBR Green, after binding to the minor groove of dsDNA, exhibits a 1000-fold increase in fluorescence compared to being free in solution (Video 5.1). However, if even more specificity is desired, a sequence-specific oligonucleotide, or hybridization probe, can be added, which binds to the target gene at some point in front of the primer (after the 3' end). These hybridization probes contain a reporter molecule at the 5' end and a quencher molecule at the 3' end. The quencher molecule effectively inhibits the reporter from fluorescing while the probe is intact. However, upon contact with DNA polymerase I, the hybridization probe is cleaved, allowing for the dye's fluorescence (Video 5.1).

10.6.2.2.3: Reverse-Transcription PCR

Since its advent, PCR technology has been creatively expanded, and ***reverse-transcription PCR (RT-PCR)*** is one of the most important advances. Real-time PCR is frequently confused with reverse-transcription PCR but is a separate technique. In RT-PCR, the DNA amplified is derived from mRNA using reverse-transcriptase enzymes to produce a cDNA copy of the gene. Using primer sequences for genes of interest, traditional PCR methods can be used with the cDNA to study the expression of genes qualitatively. Currently, reverse-transcription PCR is commonly used with real-time PCR, which allows one to quantitatively measure the relative change in gene expression across different samples.

Figure (\PageIndex{19}\) shows a video animation video showing the use of reverse transcription Polymerase Chain Reaction (RT-PCR) in COVID-19 testing.

Figure (\PageIndex{19}\): Video animation of the Reverse Transcription Polymerase Chain Reaction (RT-PCR)

Issues of Concern

One disadvantage of PCR technology is that it is extremely sensitive. Trace amounts of RNA or DNA contamination in the sample can produce extremely misleading results. Another disadvantage is that the primers designed for PCR require sequence data and, therefore, can only be used to identify the presence or absence of a known pathogen or gene. Another limitation is that sometimes the primers used for PCR can anneal non-specifically to similar sequences, but not identical, to the target gene.

Another potential issue of using PCR is the possibility of primer dimer (PD) formation. PD is a potential by-product and consists of primer molecules that have hybridized with each other due to the strings of complementary bases in the primers. The DNA polymerase amplifies the PD, leading to competition for PCR reagents that could be used to amplify the target sequences.

Clinical Significance

PCR amplification is an indispensable tool with various applications within medicine. Often, it is used to test for the presence of specific alleles, such as in the case of prospective parents screening for genetic carriers. Still, it can also be used to diagnose the presence of disease directly and for mutations in the developing embryo. For example, the first time PCR was used in this way was for the diagnosis of sickle cell anemia through the detection of a single gene mutation.

Additionally, PCR has greatly revolutionized the diagnostic potential for infectious diseases, as it can rapidly determine the identity of microbes that were traditionally unable to be cultured or that required weeks for growth. Pathogens routinely detected using PCR include Mycobacterium tuberculosis, human immunodeficiency virus, herpes simplex virus, syphilis, and countless other pathogens. Moreover, qPCR is used to test the qualitative presence of microbes and quantify bacterial, fungal, and viral loads.

The sensitivity of diagnostic tools for mutations to oncogenes and tumor suppression genes has been improved at least 10,000-fold due to PCR, allowing for earlier diagnosis of cancers like leukemia. PCR has also enabled more nuanced and individualized therapies for cancer patients. Additionally, PCR can be used for tissue typing vital to organ implantation, and it has even been proposed as a replacement for antibody-based tests for blood type. PCR also has clinical applications in prenatal testing for genetic diseases and/or clinical pathologies. Samples are obtained either via amniocentesis or chorionic villus sampling.

In forensic medicine, short pieces of repeating, highly polymorphic DNA, coined short tandem repeats (STRs), are amplified and used to compare specific gene variations to differentiate individuals. [9] Primers specific for the loci of these STRs are used and amplified using PCR. Various loci contain STRs in the human genome, and the statistical power of this technique is enhanced by checking multiple sites.

10.6.2.3: Gene Synthesis

Artificial gene synthesis, sometimes known as **DNA printing,** is a method in synthetic biology used to create artificial genes in the laboratory. Based on solid-phase DNA synthesis differs from molecular cloning and polymerase chain reaction (PCR) in that it does not have to begin with preexisting DNA sequences. Therefore, making a completely synthetic double-stranded DNA molecule with no apparent limits on either nucleotide sequence or size is possible.

The method has generated functional bacterial or yeast chromosomes containing approximately one million base pairs. Creating novel nucleobase pairs in addition to the two base pairs in nature could greatly expand the genetic code.

Har Gobind Khorana and coworkers demonstrated the synthesis of the first complete gene, a yeast tRNA, in 1972. Synthesis of the first peptide- and protein-coding genes was performed in the laboratories of Herbert Boyer and Alexander Markham, respectively.

Commercial gene synthesis services are now available. Approaches are often based on a combination of organic chemistry and molecular biology techniques, and entire genes may be synthesized "de novo" without the need for template DNA. Gene synthesis is an important tool in many fields of recombinant DNA technology, including heterologous gene expression, vaccine development, gene therapy, and molecular engineering. The synthesis of nucleic acid sequences can be more economical than classical cloning and mutagenesis procedures. It is also a powerful and flexible engineering tool for creating and designing new DNA sequences and protein functions.

Gene optimization

While the ability to make increasingly long stretches of DNA efficiently and at lower prices is a technological driver of this field, increasing attention is being focused on improving the design of genes for specific purposes. Early in the genome sequencing era, gene synthesis was used as an (expensive) source of cDNAs predicted by genomic or partial cDNA information but was difficult to clone. This practice has become less urgent as higher-quality sources of sequence-verified cloned cDNA have become available.

Producing large amounts of protein from gene sequences can sometimes prove difficult. Many of the most interesting proteins are normally regulated to be expressed in very low amounts in wild-type cells. Redesigning these genes offers a means to improve gene expression in many cases. Rewriting the open reading frame is possible because of the degeneracy of the genetic code. Thus, it is possible to change up to about a third of the nucleotides in an open reading frame and still produce the same protein. The available number of alternate designs possible for a given protein is astronomical. For a typical protein sequence of 300 amino acids, over 10^{150} codon combinations will encode an identical protein. Codon optimization, or replacing rarely used codons with more common codons, sometimes has dramatic effects. Further optimizations, such as removing RNA secondary structures, can also be included. At least in the case of *E. coli*, protein expression is maximized by predominantly using codons corresponding to tRNA that retain amino acid charging during starvation. Computer programs are used to optimize this task. A well-optimized gene can improve protein expression 2 to 10-fold; in some cases, more than 100-fold improvements have been reported. Because of the large number of nucleotide changes made to the original DNA sequence, the only practical way to create the newly designed genes is to use gene synthesis.

Oligonucleotide synthesis

Oligonucleotides are chemically synthesized using building blocks called nucleoside phosphoramidites. These can be normal or modified nucleosides with protecting groups to prevent their amines, hydroxyl, and phosphate groups from interacting incorrectly. One phosphoramidite is added at a time, the 5' hydroxyl group is deprotected, and a new base is added, and the process is repeated. The chain grows in the 3' to 5' direction, which is backward relative to DNA biosynthesis *in vivo*. In the end, all the protecting groups are removed. Figure (\PageIndex{20}\) below shows the solid phase DNA synthesis reaction.

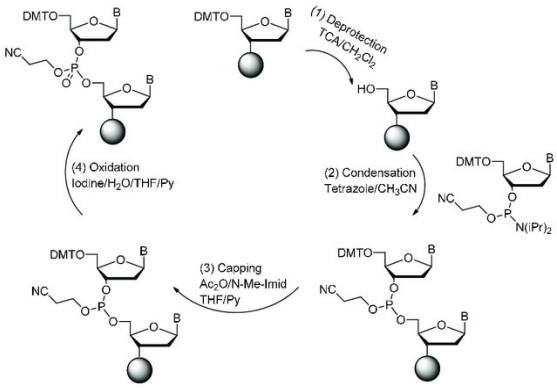


Figure (\PageIndex{20}\): **Four-step phosphoramidite oligodeoxynucleotide synthesis cycle.** The phosphoramidite method, pioneered by Marvin Caruthers in the early 1980s, and enhanced by the application of solid-phase technology and automation, is now firmly established as the method of choice. Phosphoramidite oligonucleotide synthesis proceeds in the 3 to 5 direction (opposite to the 5 to 3 direction of DNA biosynthesis in DNA replication). One nucleotide is added per synthesis cycle. The phosphoramidite DNA synthesis cycle consists of a series of steps outlined in the figure. *Image by Ni, S., et al (2017) International Journal of Molecular Sciences 18(8):1683*

Nevertheless, being a chemical process, several incorrect interactions occur, leading to some defective products. The longer the oligonucleotide sequence that is being synthesized, the more defects there are, thus this process is only practical for producing short sequences of nucleotides. The current practical limit is about 200 bp (base pairs) for an oligonucleotide with sufficient quality to be used directly for a biological application. HPLC can be used to isolate products with the proper sequence. Meanwhile, many oligos can be synthesized in parallel on gene chips. They should be prepared individually and on larger scales for optimal performance in subsequent gene synthesis procedures.

10.6.2.3.1: DNA synthesis and synthetic biology

The significant drop in the cost of gene synthesis in recent years due to increasing competition of companies providing this service has led to the ability to produce entire bacterial plasmids that have never existed. The field of synthetic biology utilizes the technology to produce synthetic biological circuits, which are stretches of DNA manipulated to change gene expression within cells and cause the cell to produce a desired product.

The ability to synthetically produce DNA will enable the development of environmental, medical, and commercially relevant products. For example, in 2015, Novartis, in collaboration with the Synthetic Genomics Vaccines Inc. and the US Biomedical Advanced Research and Development Authority, announced that they had effectively created a synthetic DNA influenza vaccine. New synthetic DNA vaccines promise to provide an alternative to current egg-produced conventional vaccines that can be plagued by low efficacy.

DNA vaccines can avoid many issues associated with egg-based vaccine production by generating viral proteins within host cells. To create a DNA vaccine, an antigen-encoding gene is cloned into a non-replicative expression plasmid, which is delivered to the host by traditional vaccination routes. Host cells that take up the plasmid express the vaccine antigen, which can be presented to immune cells via the major histocompatibility complex (MHC) pathways. CD4+ T helper cell activation following MHC class II presentation of secreted DNA vaccine protein is critical for the production of antigen-specific antibodies, as shown in Figure (\PageIndex{21}\) below.

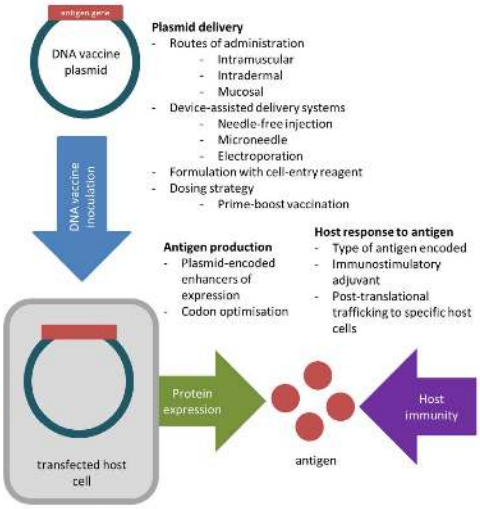


Figure by Lee, L.Y.Y., Izzard, L., and Hurt, A.C. (2018) Frontiers in Immunology, doi: 10.3389/fimmu.2018.01568

After two decades of research, DNA vaccine technology is gaining maturity—several veterinary DNA vaccines are currently licensed for West Nile virus and melanoma, and significantly, the first commercial DNA vaccine against H5N1 in chickens has recently been conditionally approved by the USDA. In addition, ongoing large animal trials of DNA vaccines against other diseases such as HIV, hepatitis, and Zika virus offer valuable insights that can be applied to influenza DNA vaccine design. Promising approaches have arisen from the numerous studies evaluating different DNA vaccine formulations and delivery systems. Still, a strategy that consistently elicits protection against influenza in large animal models has not yet emerged. Successful plasmid delivery and the use of appropriate adjuvants remain key challenges that need to be addressed before influenza DNA vaccines become effective for human use.

10.6.2.4: RNA Vaccines

Of course, amazing progress has been made in creating RNA vaccines, as demonstrated by the Moderna and Pfizer RNA vaccines to the spike protein of the SARS-CoV2 virus. Models show that just in the US through 2022, COVID-19 vaccines saved the US $1.15 trillion and 3 million lives. Just in the first year, the vaccines are estimated to have saved upwards of 20 million lives worldwide, an accomplishment worthy of every prize in the world! Many more could have been saved in the developing world if the vaccine was more widely available.

As we learned previously, RNA is much more labile to hydrolysis than DNA since RNA has a 2' OH group. Methods to stabilize RNA were required before an RNA vaccine could become a reality. More importantly, dsRNA is a danger signal that a viral infection may be present. dsRNA, a pathogen-associated molecular pattern (PAMP), binds to the Toll-like Receptor 3 (TLR3) and initiates an inflammatory response that would eliminate the RNA before it could be decoded to form a protein sequence that could elicit an antibody response, a requirement for a vaccine.

Katalin Karikó and Drew Weissman found that modifying the uracil bases to pseudouracil (Ψ) prevents the PAMP response and allows RNA to last long enough to create the protein sequence (i.e., it increases the stability of the RNA to hydrolysis). Pseudouracil (Ψ) is found in structural RNAs (transfer, ribosomal, small nuclear, and small nucleolar), and is the most common modification found in RNA. It even is metabolized by a naturally occurring pathway back to uracil.

Figure (\PageIndex{22}\) below shows the structures of uracil and the N1-methyl derivative of pseudouracil attached to ribose in an RNA and their base pairing to adenine. The N1-methylpseudouracil base still bases pairs with an adenine base. Hence, the modification does not affect the RNA structure and its functional ability to be decoded into a protein sequence. In short, mRNAs modified to contain pseudouracil have a much greater translational capacity and stability. In addition, methylation of uracil decreases the immunogenicity of RNAs that contain it.

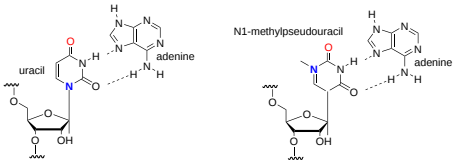


Figure (\PageIndex{22}\): Structural and base pairing features of uracil and N1-methylpseudouracil

Figure (\PageIndex{23}\) below shows a cartoon version of the modified mRNA (a) Covid-19 and its encapsulation into a lipid nanoparticle (b) for an mRNA vaccine.

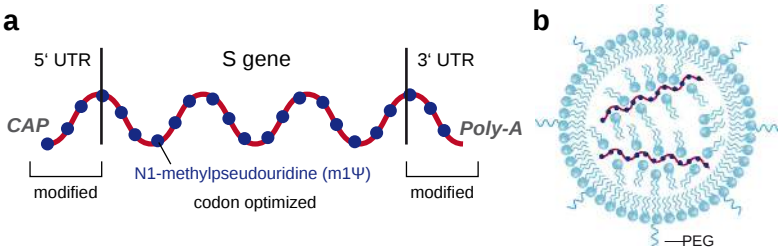


Figure (\PageIndex{23}\): Modified mRNA (a) Covid-19 and its encapsulation into a lipid nanoparticle (b) for an mRNA vaccine. Heinz, F.X., Stiasny, K. Distinguishing features of current COVID-19 vaccines: knowns and unknowns of antigen presentation and modes of action. *npj Vaccines* **6,** 104 (2021). https://doi.org/10.1038/s41541-021-00369-6. Creative Commons Attribution 4.0 International License

Panel (a) shows the generalized structure of the mRNA vaccine for the S gene, which encodes the virus's surface spike protein. Panel b shows the lipid nanoparticle that encapsulates and protects the mRNA vaccine. The lipids include proprietary mixtures of phospholipids, cholesterol, cationic (ionizable) lipids, and polyethylene glycol (PEGylated). Without the basic research of Karikó and Weissman (and many others), the world would not have had the Covid-19 vaccine in time to save so many lives. They were awarded the Nobel Prize in Medicine in 2023 for their work.

Biosynthesis of pseudouridine

The biosynthesis of pseudouridine, the most common modification of cellular RNA, is done after DNA transcription (i.e., post-transcriptionally) using the enzyme pseudouridine synthases (PUS), which is found in all kingdoms in life. The reaction involves

- cleavage of the C-N-glycosidic bond of uridine in RNA
- rotation of the cleaved uracil to align C5 of uracil and C1′ of the ribose
- formation of the C1′-C5 carbon–carbon bond.

These processes are illustrated in Figure (\PageIndex{24}\) below.

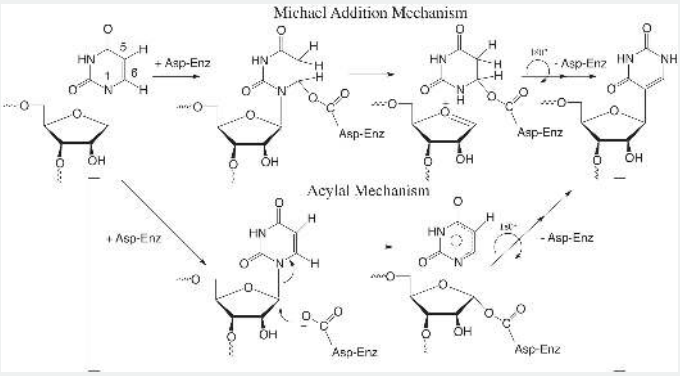


Figure (\PageIndex{24}\): Post-transcriptional modification of uridine to pseudouridine. Czudnochowski N. et al. The mechanism of pseudouridine synthases from a covalent complex with RNA, and alternate specificity for U2605 versus U2604 between close homologs. Nucleic Acids Res. 2014 Feb;42(3):2037-48. doi: 10.1093/nar/gkt1050. Epub 2013 Nov 7. PMID: 24214967; PMCID: PMC3919597. Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)

10.6.3: Summary

This chapter provides a comprehensive overview of the technologies and methodologies that underpin modern genomics and synthetic biology. It begins by distinguishing between genomic DNA (gDNA) and complementary DNA (cDNA). Genomic DNA represents the full chromosomal content of an organism—including both coding and non-coding regions—while cDNA is synthesized from mature mRNA, representing only the expressed (exonic) sequences. This distinction is crucial when cloning eukaryotic genes for expression in systems that lack the splicing machinery.

The chapter then reviews the flow of genetic information in eukaryotes, from transcription of gDNA into pre-mRNA, through RNA processing (splicing, 5′-capping, and polyadenylation), to translation. It emphasizes how the structural differences between prokaryotic and eukaryotic genes (such as the presence of introns and the organization of operons) dictate different strategies for gene cloning and expression.

Subsequently, the text covers DNA extraction and purification methods, detailing techniques ranging from traditional phenol/chloroform extraction to modern silica-based chromatography. It explains the chemical principles that allow nucleic acids to be selectively bound and eluted, ensuring the production of high-quality DNA for downstream applications.

The chapter then transitions into DNA sequencing technologies. It outlines the evolution from early Maxam-Gilbert and Sanger sequencing methods to next-generation sequencing (NGS) platforms such as Illumina, Roche 454, Ion Torrent, nanopore sequencing, and Single Molecule Real-Time (SMRT) sequencing. Each technology's workflow, advantages (e.g., speed, cost, throughput, and accuracy), and limitations are discussed, highlighting how advancements in sequencing have revolutionized fields ranging from genomics to personalized medicine.

In addition, the Polymerase Chain Reaction (PCR) and its variants (including real-time PCR and reverse-transcription PCR) are presented as indispensable techniques for amplifying and quantifying nucleic acids. The principles of PCR—denaturation, primer annealing, and elongation—are explained alongside common challenges such as primer dimer formation and non-specific amplification.

The latter sections of the chapter delve into DNA synthesis and synthetic biology. The process of artificial gene synthesis, including oligonucleotide synthesis via phosphoramidite chemistry, is discussed along with strategies for gene optimization through codon usage and secondary structure minimization. These innovations enable the design and construction of entirely synthetic genetic circuits, which have broad applications in research, vaccine development, and therapeutic interventions.

Finally, the chapter highlights the emergence of RNA vaccines, detailing how modifications such as the substitution of uracil with pseudouracil enhance mRNA stability and reduce immunogenicity. The encapsulation of modified mRNA in lipid nanoparticles is explained as a critical step in vaccine delivery, a technology that has been pivotal in the rapid development of COVID-19 vaccines.

In summary, this chapter ties together the foundational principles of DNA structure, extraction, and sequencing with cutting-edge applications in gene synthesis, synthetic biology, and vaccine technology, underscoring how advancements in molecular biology have transformed our ability to read, write, and edit genetic information for both scientific and clinical purposes.

This page titled 10.6: DNA Isolation, Sequencing, and Synthesis is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **9.1: DNA Isolation, Sequencing, and Synthesis** by Henry Jakubowski and Patricia Flatt has no license indicated.

10.7: Bioinformatics

10.7.1: Introduction

An unprecedented revolution has been observed in science with recent technological advances, which have provided a large amount of “omic” data. The crescent generation and availability of this information available in public databases were, and still are, a challenge for professionals from different areas. However, what is the challenge? In biology, the main challenge is to make sense of the enormous amount of structural data and sequences that have been generated at multiple levels of biological systems. Still, in bioinformatics, the development of tools is necessary (statistical and computational) capable of assisting in understanding the mechanisms underlying biological questions in the study. Besides, if we consider the complexity of science, this is a highly reductionist view. The era of a “new biology” emerges accompanied by the birth/development of other sciences, such as bioinformatics and computational biology, which have an integrated interface with molecular biology. Although considered recently, bioinformatics and genomics have evolved interdependently and promoted a historical impact on the available knowledge.

Bioinformatics is a hybrid science that links biological data with techniques for information storage, distribution, and analysis to support multiple areas of scientific research including biomedicine. It uses molecular biology and genetics, computer science, mathematics, and statistics to address data-intensive, large-scale biological problems are addressed from a computational point of view. Bioinformatics is fed by high-throughput data-generating experiments, including genomic sequence determinations and measurements of gene expression patterns. Database projects curate and annotate the data and then distribute it via the World Wide Web. Mining these data leads to scientific discoveries and in the end to the identification of new clinical applications.

A bioinformatics solution usually involves the following steps:

- Collect statistics from biological data
- Build a computational model
- Solve a computational modeling problem
- Test and evaluate a computational algorithm

It also addresses the following aspects:

- Types of biological information and databases
- Sequence analysis and molecular modeling
- Genomic analysis
- Systems biology

In the field of medicine in particular, many important applications for bioinformatics have been discovered. For example, it is used to identify correlations between gene sequences and diseases, to predict protein structures from amino acid sequences, to aid in the design of novel drugs, and to tailor treatments to individual patients based on their DNA sequences (pharmacogenomics). Using bioinformatics, we can now conduct global analyses of all the available data to uncover common principles that apply across many systems and highlight novel features.

Many bioinformatic tools are available online in the fields of genomics, comparative genomics, proteomics, drug discovery, cancer research, phylogenetics, forensic sciences, biodefense, nutrigenomics, gene expressions, protein structure/function, etc. Instead of endorsing any given tool, and constantly curating web addresses, which often change, we offer a series of links that readers can explore to find the tools appropriate to their needs.

- Bioinformatics resources - NCBI/NIH
- What is bioinformatics from EMBL-EMI: A course
- Bioinformatics tools from GenScript
- Bioinformatics tools - FDA
- Bioinformatics Resources - UC Berkeley
- NCBI Bioinformatics Resources: An Introduction: NCBI/NIH Homepage - UC Berkeley

This page titled 10.7: Bioinformatics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **9.2: Bioinformatics** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

11: Lipids

11.1: Introduction to lipids
11.2: Lipids Aggregates in Water - Micelles and Liposomes
11.3: Structure and Function - Lipids and Membranes

11: Lipids is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

11.1: Introduction to lipids

Learning Goals (ChatGPT o3-mini, 2/2/25)

- **Define Lipids and Their Key Properties:**
  • Describe what lipids are, focusing on their organic nature, hydrophobicity, and solubility in organic solvents versus aqueous solutions.
- **Classify Lipids Based on Structural Components:**
  • Distinguish between lipids that contain fatty acids and those that contain isoprenoid units, and explain the significance of each class.
- **Analyze Fatty Acid Structure and Conformation:**
  • Explain the structure of fatty acids, including chain length, saturation, and the impact of cis (and trans) double bonds on the conformation (zig-zag vs. kinked structures) and melting points.
- **Understand the Nomenclature and Stereochemistry of Fatty Acids:**
  • Interpret symbolic, systematic, and common names of fatty acids and explain concepts like prochirality and the stereospecific numbering (sn system) in glycerolipids.
- **Connect Lipid Structure to Function in Energy Storage and Membrane Architecture:**
  • Discuss how the highly reduced state of fatty acids in triacylglycerols makes them excellent energy storage molecules and compare this with the structure and hydration of polysaccharides like glycogen.
  • Explain how the amphiphilic nature of lipids leads to the formation of micelles and bilayers, which are critical for membrane structure and function.
- **Evaluate the Diversity of Lipid Types and Their Physiological Roles:**
  • Describe the roles of storage lipids (triacylglycerols) versus membrane lipids (glycerophospholipids, glyceroglycolipids, and sphingolipids), and how variations in fatty acid composition (saturated vs. unsaturated) influence their physical properties and biological functions.
- **Interpret the Functional Importance of Essential Fatty Acids:**
  • Discuss why certain fatty acids (n-6 and n-3 classes) are essential, their metabolic precursors, and their roles in health, including effects on inflammation and brain development.
- **Understand Isoprenoid Lipid Biosynthesis and Function:**
  • Explain how isoprenoid lipids are synthesized from DMAPP and IPP, and describe the roles of key isoprenoid derivatives such as cholesterol, terpenes, and vitamins in cellular physiology.
- **Relate Lipid Aggregation to Biological Function:**
  • Describe how fatty acids aggregate into micelles in aqueous environments, and how this self-association influences their function, including transport by fatty acid-binding proteins (e.g., albumin).
- **Integrate Lipid Structural Variability with Membrane Dynamics:**
  • Explain how differences in lipid shape (e.g., cone-like versus cylindrical) determine their self-assembly into specific structures like micelles or bilayers and how this impacts membrane fluidity, signaling, and vesicle formation.

These goals are intended to guide your exploration of lipid biochemistry, linking the detailed molecular structure of lipids with their diverse roles in energy storage, membrane architecture, and cellular signaling.

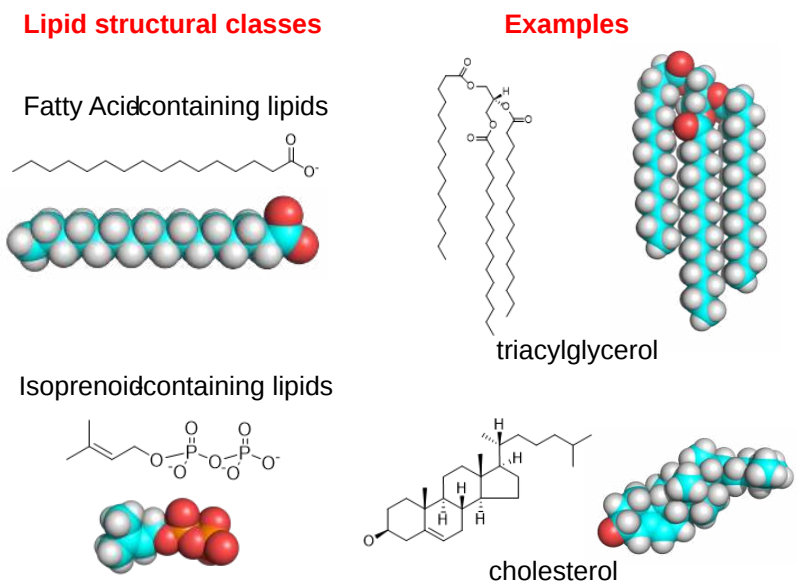
11.1.1: Introduction to Lipids

(Thanks to Rebecca Roston for providing a cohesive organizational framework and image templates)

Lipids are organic molecules soluble in organic solvents, such as chloroform/methanol, but sparingly soluble in aqueous solutions. These solubility properties arise since lipids are primarily hydrophobic. One type, triglycerides, is used for energy storage since they are highly reduced and get oxidized to release energy. Their hydrophobic nature allows them to pack efficiently through self-

association in an aqueous environment. Triglycerides are also used for insulation since they conduct heat poorly, which is good if you live in a cold climate but bad if you wish to dissipate heat in a hot one. Triglycerides also offer padding and mechanical protection from shocks (think walruses). Another type of lipid forms membrane bilayers, which separate cellular contents from the outside environment or separate intracellular compartments (organelles) from the cytoplasm. Some lipids are released from cells to signal other cells to respond to specific stimuli in a process called cell signaling. From a more molecular perspective, lipids can act as cofactors for enzymes, pigments, antioxidants, and water repellents. As we saw with proteins, lipid structure mediates their function. So, let's probe their structures.

Lipids can be split into structural classes in a variety of ways. An earlier classification divided them into those that release fatty acids under base-catalyzed hydrolysis to form soaps in a saponification reaction and those that don't. A much better and broader classification is based on whether the lipids contain fatty acids or isoprenoids, as shown in Figure **11.1.1** below.



11.1.1.1:

Figure **11.1.1**: Fatty acids and isoprenoid lipids

The nonpolar chains of the fatty acid are drawn in the figure above in the lowest energy zig-zag fashion, as we saw when we discussed the main chain conformation of proteins (Chapter 4.1). In that chapter, we started by exploring a long 12-C chain of carboxylic acid, dodecanoic acid. In the lowest energy conformation, the dihedral angles are all $\pm 180^0$ to minimize torsional strain in the molecule. Rotation around one C-C bond can produce a gauche form, which introduces a kink into the chain, as shown in Figure **11.1.2**.

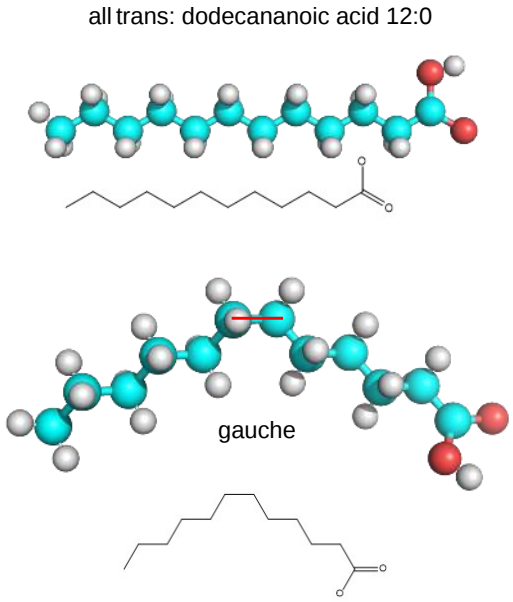


Figure 11.1.2: Trans and Gauche bonds in a fatty acids

11.1.2: Fatty Acids

Fatty acids can be free or covalently linked by ester or amide link to a base molecule like triglycerides or membrane lipids. The key principle we learned in our study of proteins, which is that structure determines function, also applies to lipids. The figure below shows three types of molecules: a free fatty acid, a wax with an esterified fatty acid, and a glycolipid with a fatty acid connected by an amide link in another type of lipid (glycosphingolipid). Each has different properties, leading to different functions. Waxes, for instance, are very nonpolar and water-insoluble. They are amorphous solids at room temperature but, depending on their structure, can melt to form high-viscosity liquids. They are used as a coating on the surfaces of leaves to help prevent water loss. The glycolipids (glyco- as they contain a monosaccharide group) are constituents of membrane bilayers. Figure 11.1.3 shows the general structures of fatty acid-containing waxes and other lipids.

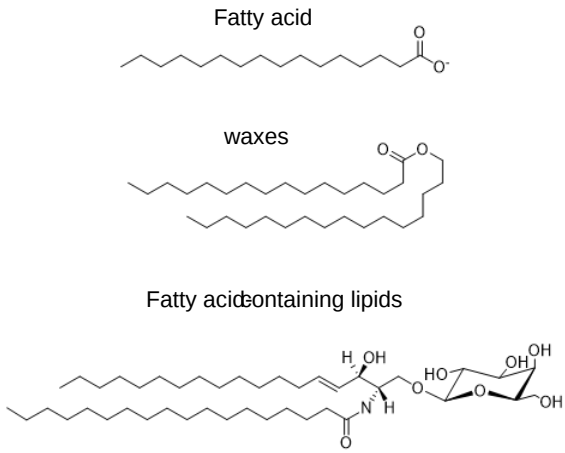


Figure 11.1.3: Fatty acids in waxes and glycolipids

Fatty acids vary in length, usually contain an even number of carbon atoms, and can be saturated (have no double bonds in the acyl chain) or unsaturated (with either one -monounsaturated - or multiple - polyunsaturated - cis double bond(s)). The double bonds are

NOT conjugated as they are separated by a methylene (-CH_2-) spacer. Fatty acids can be named in many ways.

- Symbolic name: given as x:y $^{\Delta a,b,c}$, where x is the number of carbon atoms in the chain, y is the number of double bonds, and a, b, and c are the positions of the start of the double bonds, counting from C1 - the carboxyl carbon. Double bonds are usually cis (Z).
- Systematic name using IUPAC nomenclature. The systematic name gives the number of carbon atoms in the chain (e.g., hexadecanoic acid for 16:0). If the fatty acid is unsaturated, the base name reflects the number of double bonds (e.g., octadecenoic acid for 18:1 $^{\Delta 9}$ and octadecatrienoic acid for 18:3$^{\Delta 9,12,15}$).
- Common name: (e.g., oleic acid, 18:1$^{\Delta 9}$), which is found in high concentration in olive oil.

They can be named most efficiently with a symbolic name. Figure **11.1.4** shows examples of fatty acids and their symbolic names.

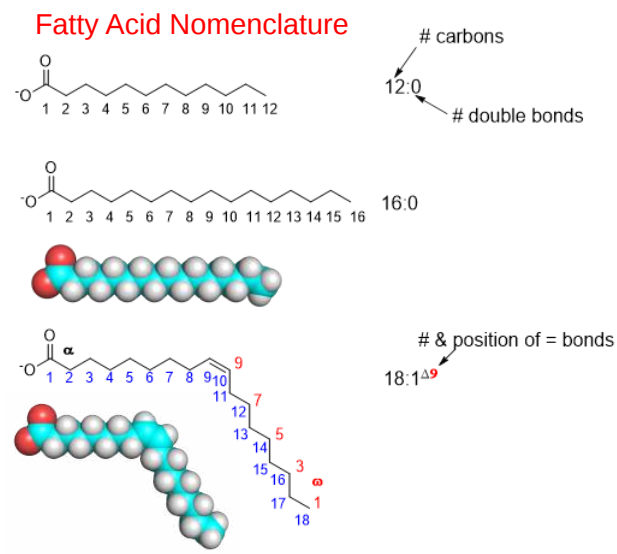


Figure **11.1.4**: Some fatty acids and their symbolic names.

There is an alternative to the symbolic representation of fatty acids, in which the carbon atoms are numbered from the distal end (the n or ω end) of the acyl chain (the opposite end from the alpha carbon). Hence, 18:3 $^{\Delta 9,12,15}$ could be written as 18:3 (ω -3) or 18:3 (n -3), where the terminal C is numbered one and the first double bond starts at C3.

The most common saturated fatty acids in biochemistry textbooks are listed in the table below. Note how the melting point increases with the length of the hydrocarbon chain. This arises from increasing noncovalent induced dipole-induced dipole attractions between the long chains. Heat must be added to lessen these attractions to allow melting. Table **11.1.1** below shows examples of fatty acids and their symbolic names.

Symbolic	common name	systematic name	structure	mp(C)
12:0	Lauric acid	dodecanoic acid	$CH_3(CH_2)_{10}COOH$	44.2
14:0	Myristic acid	tetradecanoic acid	$CH_3(CH_2)_{12}COOH$	52
16:0	Palmitic acid	Hexadecanoic acid	$CH_3(CH_2)_{14}COOH$	63.1
18:0	Stearic acid	Octadecanoic acid	$CH_3(CH_2)_{16}COOH$	69.6
20:0	Arachidic acid	Eicosanoic acid	$CH_3(CH_2)_{18}COOH$	75.4

Table **11.1.1**: Examples of fatty acids and their symbolic names

Table **11.1.2** below shows common unsaturated fatty acids. Arachidonic acid is an (ω -6) fatty acid, while docosahexaenoic acid is an (ω -3) fatty acid. Note the decreasing melting point for the 18:X series with increasing double bonds.

Symbol	common name	systematic name	structure	mp(C)
$16:1^{\Delta 9}$	Palmitoleic acid	Hexadecenoic acid	$CH_3(CH_2)_5CH=CH-(CH_2)_7COOH$	-0.5
$18:1^{\Delta 9}$	Oleic acid	9-Octadecenoic acid	$CH_3(CH_2)_7CH=CH-(CH_2)_7COOH$	13.4
$18:2^{\Delta 9,12}$	Linoleic acid	9,12 -Octadecadienoic acid	$CH_3(CH_2)_4(CH=CHCH_2)_2(CH_2)_6COOH$	-9
$18:3^{\Delta 9,12,15}$	α-Linolenic acid	9,12,15 -Octadecatrienoic acid	$CH_3CH_2(CH=CHCH2)3(CH2)6COOH$	-17
$20:4^{\Delta 5,8,11,14}$	arachidonic acid	5,8,11,14- Eicosatetraenoic acid	$CH3(CH2)_4(CH=CHCH_2)_4(CH_2)_2COOH$	-49
$20:5^{\Delta 5,8,11,14,17}$	EPA	5,8,11,14,17- Eicosapentaenoic- acid	$CH_3CH_2(CH=CHCH_2)_5(CH_2)_2COOH$	-54
$22:6^{\Delta 4,7,10,13,16,19}$	DHA	Docosahexaenoic acid	22:6w3	

Table **11.1.2**: Common unsaturated fatty acids

Let's consider how double bonds in fatty acids influence their melting points. Figure **11.1.5** shows common variants of fatty acids, each with 18 carbon atoms. Compare the Lewis structure and spacefill models below. What a difference a cis double bond makes!

The double bonds in fatty acids are cis (Z), which introduces a "permanent" kink into the chain, similar to the"temporary" kink in the saturated dodecanoic acid carboxylic acid with a single gauche bond (Figure 2). The kink is permanent since there is no rotation around the double bond unless it is broken (which can happen through photoisomerization reactions). The more double bonds, the greater the kinking. The more kinks, the less chance for van der Waals contacts between the acyl chains, and the reduced induced-dipole-induced dipole interactions between the chains lead to lowered melting points.

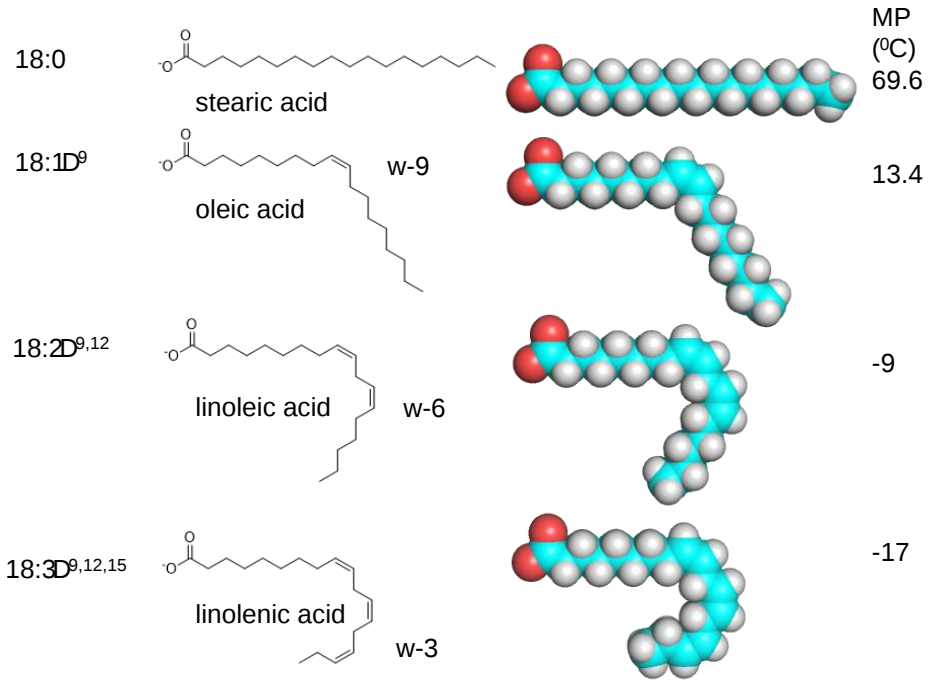


Figure **11.1.5**: Fatty acids with 18 carbon atoms

In your mind, replace the cis double bond in oleic acid with a trans double bond. You should now see a long "zig-zag" shaped molecule with no kinks. Trans fatty acids are rare in biology but are produced in the industrial partial hydrogenation of fats, which decreases the number of double bonds and makes the fats more solid-like and tastier while decreasing rancidity. These trans fatty

acids would pack closer together and affect the structure and function of the lipid in a given environment (such as a membrane bilayer). Increased consumption of trans fatty acids is associated with an increased risk of cardiovascular disease.

Table **11.1.3**: shows the percentages of fatty acids in different oils/fats

FAT	<16:0	16:1	18:0	18:1	18:2	18:3	20:0	22:1	22:2	.
Coconut	87	.	3	7	2	.	.	.	.	.
Canola	3	.		11	13	10	.	7	50	2
Olive Oil	11	.	4	71	11	1	.	.	.	.
Butter-fat	50	4	12	26	4	1	2	.	.	.

Table **11.1.3**: Percentages of fatty acids in different oils/fats

The fatty acid composition differs in different organisms:

- animals have 5-7% of fatty acids with 20-22 carbons, while fish have 25-30%
- animals have <1% of their fatty acids with 5-6 double bonds, while plants have 5-6% and fish 15-30%

As there are essential amino acids that humans cannot synthesize, there are also essential fatty acids that the diet must supply. There are only two, one each, in the n-6 and n-3 classes:

- n-6 class: α-linoleic acid (18:2 n-6, or $18{:}2^{\Delta 9,12}$) is a biosynthetic precursor of arachidonic acid (20:4 n-6 or $20{:}4^{\Delta 5,8,11,14)}$
- n-3 class: linolenic acid (18:3 n-3, or $18{:}3^{\Delta 9,12,15}$) is a biosynthetic precursor of eicosapentaenoic acid (EPA, 20:5 n-3 or $20{:}5^{\Delta 5,8,11,14,17}$) and to a much smaller extent, docosahexaenoic acid (DHA, 22:6 n-3 or $22{:}6^{\Delta 4,7,10,13,16,19}$).

These two fatty acids are essential since mammals cannot introduce double bonds in fatty acids beyond carbon 9. These essential precursor fatty acids are substrates for intracellular enzymes such as elongases and desaturases (to produce 20:4 n-6, 20:5 n-3, and 22:6 n-3 fatty acids), and beta-oxidation type enzymes in the endoplasmic reticulum and another organelle, the peroxisome. The peroxisome is involved in the oxidative metabolism of straight-chain and branched fatty acids, peroxide metabolism, and cholesterol/bile salt synthesis. Animals fed diets high in plant 18:2(n-6) fats accumulate 20:4(n-6) fatty acids in their tissues, while those fed diets high in plant 18:3(n-3) accumulate 22:6(n-3). Animals fed diets high in fish oils accumulate 20:5 (EPA) and 22:6 (DHA) at the expense of 20:4(n-6).

Many studies support the claim that diets high in fish containing abundant ω-3 fatty acids, particularly EPA and DHA, reduce inflammation and cardiovascular disease. ω-3 fatty acids are abundant in high-oil fish (salmon, tuna, sardines) and lower in cod, flounder, snapper, shark, and tilapia.

Some suggest that, contrary to images of early hominids as hunters and scavengers of meat, human brain development might have required the consumption of fish, which is highly enriched in arachidonic and docosahexaenoic acids. The brain contains lipids that are highly enriched in these two fatty acids. These fatty acids are necessary for the proper development of the human brain. Deficiencies in these might contribute to ADHD, dementia, and dyslexia. These fatty acids are essential in the diet. The mechanism for the protective effects of n-3 fatty acids in health will be explored later in the course when we discuss prostaglandin synthesis and signal transduction.

Aggregates of Fatty Acids in Aqueous Solution: Micelles

Structure determines both properties and function. It should be evident that free, unesterified fatty acids are very (but not completely) insoluble in water. When added to water, they saturate the solution at very low concentrations and then undergo a phase separation to produce aggregates called **mice**lles. The micelle structure formed from dodecyl sulfate, a common detergent with a sulfate instead of a carboxylate head group, is shown below. All the nonpolar Cs and Hs of the long alkyl chains are "buried" and are not exposed to water, whereas the sulfate head groups are solvent-exposed.

Figure **11.1.6** shows an interactive iCn3D model of image

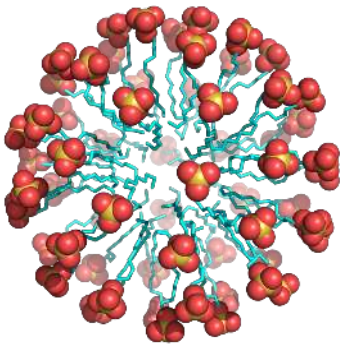


Figure 11.1.6: Sodium dodecylsulfate micelle (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...KECTiGaPPScED7

Fatty acids (carboxylates and sulfates) are **amphiphilic**, with a larger polar/charged head group that tapers down to a hydrophobic tail, forming a cone-like structure. This cone structure allows the packing of many of these single-chain amphiphiles into a micelle, as shown in Figure 11.1.7.

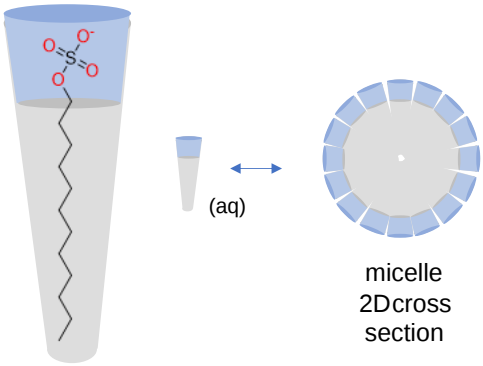


Figure 11.1.7: Cone-like representation of sodium dodecyl sulfate free and in a micelle

Free fatty acids are transported in the cell and the blood, not in micelles, but by fatty acid-binding proteins. The most abundant protein in the blood, albumin, binds and transports fatty acids. Figure 11.1.8 shows an interactive iCn3D model of hexadecanoic acid bound to human albumin (1E7H). The fatty acids are shown in colored spacefill rendering.

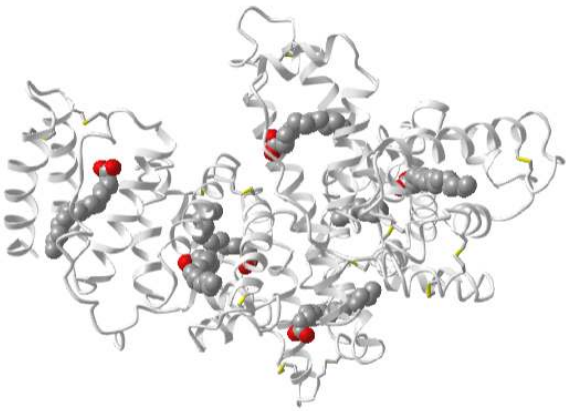


Figure 11.1.8: Hexadecanoic acid bound to human albumin (1E7H). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CGUdMj11vNevF6

11.1.3: Waxes

You are familiar with earwax and also the waxy surface of plants. Ear wax contains long-chain fatty acids (saturated and unsaturated) and alcohols derived from them. (They also contain isoprenoid derivatives like squalene and cholesterol, discussed below). The waxy cuticle surface layers of plants contain very long-chain fatty acids (C20–C34) and their derivatives, including alkanes, aldehydes, primary and secondary alcohols, ketones, and esters. We will consider waxes as very long-chain fatty acids and their derivatives, as shown in Figure **11.1.9**.

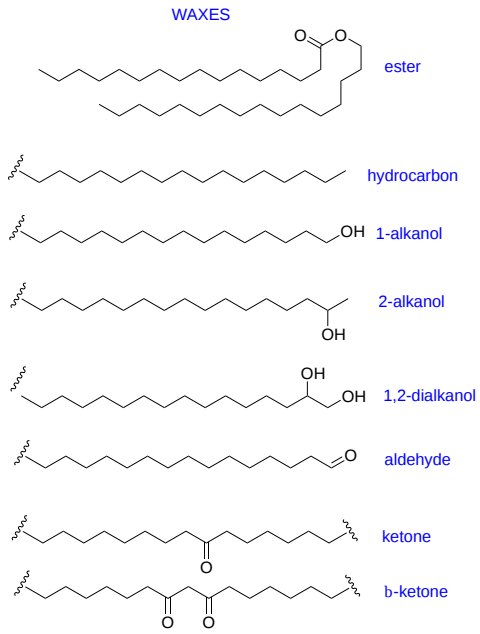


Figure **11.1.9**: Waxes

These molecules are extremely nonpolar and, as such, make great barriers preventing water loss through leaves and water penetration into the ear. This group of molecules clearly shows that the properties (insolubility, high melting point) and function (hydrophobic barrier/protection) arise from structure (very long chain carbon molecules, few electronegative atoms, and lack of C=C double bonds).

11.1.4: Fatty Acid-Containing Lipids

We can categorize these lipids based on function or structure (even though these are related).

Function

- storage lipids - triacylglycerols
- membrane lipids - many different lipids

Structure

- glycerolipids, which use glycerol as a backbone for fatty acid attachment
- sphingolipids, which use sphingosine as a backbone

The structures of glycerol and sphingosine are shown in Figure **11.1.10**. Fatty acids are connected to these two "backbone" structures by either ester (mostly) or amide links.

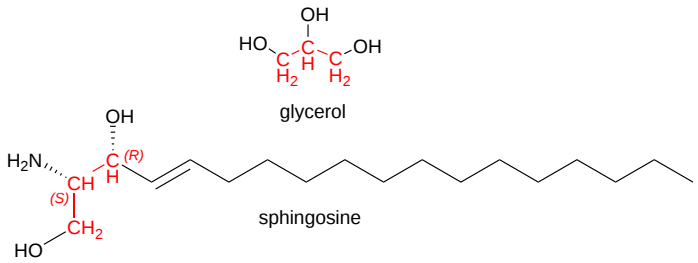


Figure **11.1.10**: Structures of glycerol and sphingosine

Let's explore the classes of fatty acid-containing lipids that use these two "backbone" structures.

11.1.5: Storage Lipids - Triacylglycerols (TAGs)

Most fatty acids in species that store fatty acids for energy are found in **triacylglycerols**. You will often see them named triglycerides or triacylglycerides, but this term is used more in clinical chemistry and industry (and often in the media). Figure **11.1.11** shows a schematic diagram of the glycerol backbone with three fatty acids esterified. They are glycerolipids as they contain a glycerol base.

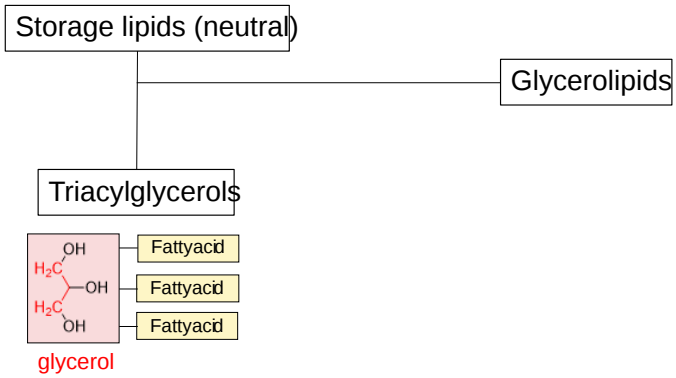


Figure **11.1.11**: Structural features and nomenclature for triacylglycerols

Glycerol is not chiral, but given the incredible diversity of fatty acids, glycerols likely have three different fatty acids esterified to them, making them chiral. If triacylglycerols contain predominantly saturated fatty acids, they are solids at room temperature and are called **fats**. Those with multiple double bonds in the fatty acids are likely liquids at room temperature. These are called **oils**. Triacylglycerols are even more insoluble than fatty acids, which contain a polar and mostly charged carboxylate.

Figure **11.1.12** shows a triacylglycerol containing all 16:0 saturated fatty acids (left) and one containing all $16{:}2^{\Delta 9,12}$ polyunsaturated fatty acids. The figures were constructed with a specific set of dihedral angles to illustrate a point that polyunsaturated fats in a triacylglycerol don't pack as tightly and have lower induced dipole-induced dipole attractions between the acyl chains than is possible with saturated fatty acyl chains. Hence, the melting point of triacylglycerols containing polyunsaturated fatty acids is lower than that of those with saturated ones.

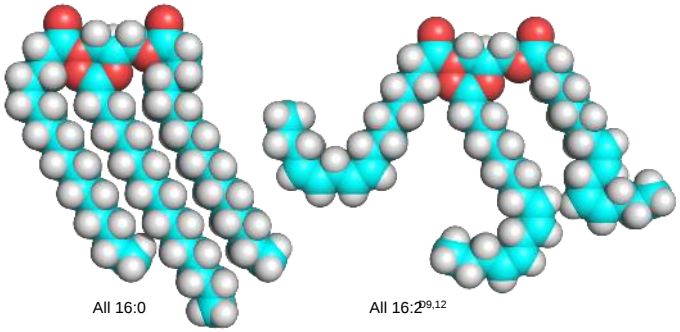


Figure 11.1.12: Triacylglycerol containing all 16:0 saturated fatty acids (left) and one containing all 16:2$^{\Delta 9,12}$ polyunsaturated fatty acids.

Let's repeat the key mantra: the structure of lipids determines their function. Consider the very insoluble triacylglycerol used as the body's predominant storage form of chemical energy. In contrast to polysaccharides such as glycogen (a glucose polymer), the carbon atoms in the acyl chains of the triacylglycerol are in a highly reduced state. The main energy source that drives our bodies and society is oxidizing carbon-based molecules to carbon dioxide and water in a highly exergonic and exothermic reaction. Sugars are already partway down the free energy "hill" since each carbon is partially oxidized. 9 kcal/mol (38 kJ/mol) can be derived from the complete oxidation of fats, in contrast to 4.5 kcal/mol (19 kJ/mol) from proteins or carbohydrates.

In addition, glycogen is highly hydrated. For every 1 g of glycogen, 2 g of water is H-bonded. Hence, storing the equivalent mass of carbohydrates would take 3 times more weight than triacylglycerol, which is stored in anhydrous lipid "droplets" within cells. In addition, fats are more flexible, given the large number of conformations available to the acyl chain C-C bonds by simple rotation around the C-C bonds. Polysaccharides have monomeric cyclohexane-like chair structures and are much more rigid.

Triacylglycerols are predominantly stored in special cells called adipocytes, which comprise adipose or fat tissue and must be mobilized by signaling agents and transported as fatty acids to cells for utilization. Again, triacylglycerols don't form membranes, which separate the outside and inside aqueous environments. They are simply so insoluble that they phase-separate into lipid droplets. Their formation and structure are a bit more complex than that, though, and we will discuss lipid droplets more in the next section.

11.1.6: Membrane Lipids

Membranes are bilayers of amphiphilic lipids that separate the outside and inside (cytoplasm) aqueous environments in cells and the cytoplasm and interior contents of organelles within cells. In general, single-chain lipid amphiphiles form micelles. Amphiphilic membrane lipids typically have **two** nonpolar tails connected to a polar head, giving them a less conical and more cylindrical shape that disallows micelle formation while favoring bilayer formation. Figure 11.1.13 shows a short section of a bilayer membrane made from lipids with a polar (and charged) head group (phosphocholine) and two 16:0 chains. The red and blue spheres represent the O and N atoms of the head groups, which are sequestered to the exterior parts of the bilayer, where they interact with water. The nonpolar 16:0 tails are shown in cyan, clearly illustrating the nonpolar nature of the interior of the bilayer.

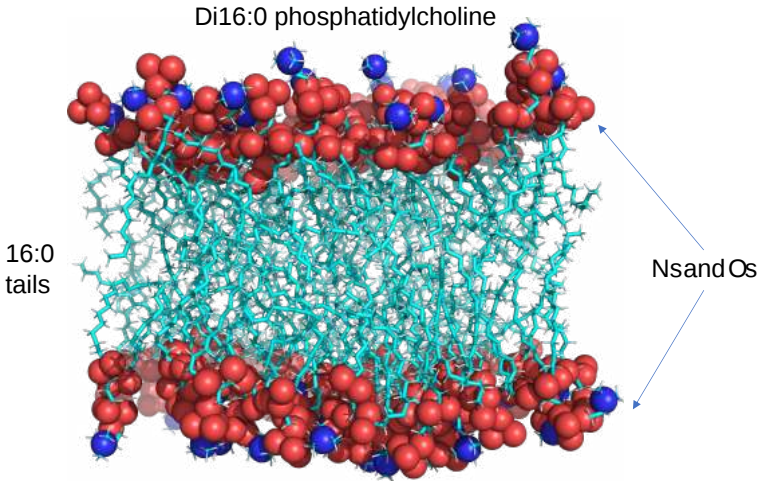


Figure 11.1.13: Bilayer membrane made from lipids with a polar (and charged) head group (phosphocholine) and two 16:0 chains.

Figure 11.1.14 shows an interactive iCn3D model of a hydrated bilayer of the di16:0 phosphatidylcholine bilayer.

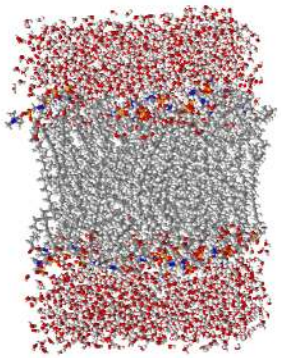


Figure 11.1.14: Hydrated di16:0 phosphatidycholine bilayer (Copyright; author via source).
Click the image for a popup.

We will consider two general types of fatty acid-containing membrane lipids**: glycerolipids**, with two fatty acids esterified to a glycerol base, and **sphingolipids,** with one fatty acid in amide linked to a different base, sphingosine. Sphingosine has a built-in long alkyl chain that provides the "second" nonpolar chain. There are many different polar/charged head groups for these membrane lipids.

Let's look more at the double-chain amphiphiles comprising these membrane bilayers.

11.1.7: Glycerolipids

There are two main glycerolipids, glycerophospholipids and glyceroglycolipids, the most common lipids in membranes.

Glycerophospholipids

Figure 11.1.15 shows glycerophospholipids' structural features and nomenclature.

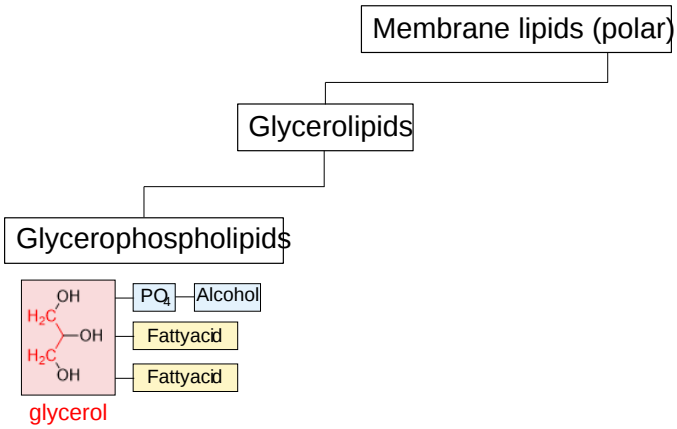


Figure 11.1.15: Structural features and nomenclature for glycerophospholipids

These lipids have enormous structural variability, given the large number of different fatty acids (both saturated and unsaturated) and head groups that can be attached to a phosphate attached to carbon 3 of glycerol. Figure 11.1.16 shows the structures of the most common glycerophospholipids.

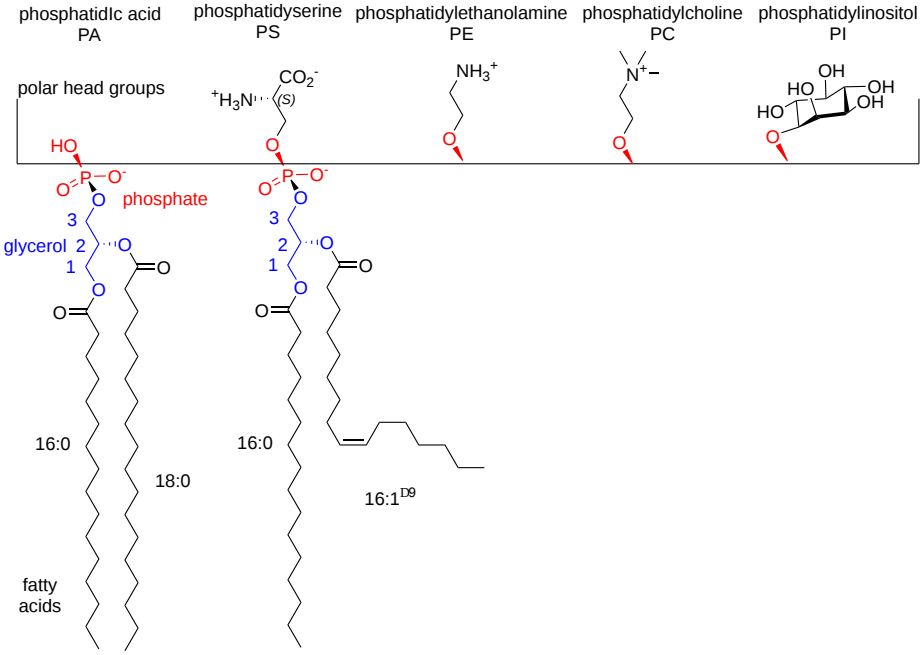


Figure 11.1.16: Common glycerophospholipids

Phosphatidylcholine (PC) is commonly called lecithin, while phosphatidylserine (PS) is called cephalin.

Note that the head groups have charges since they all have a negatively charged phosphate. PS has two additional charged atoms, which would effectively cancel out. PE has a charged amine but could become uncharged at pH values approaching its pKa. PC has a quaternary amine, which is positively charged at all pHs. Hence, PC has a positive and negative charge but a zero net charge.

Glyceroglycolipids

These do not have a phosphate group attached to the oxygen on C3 of glycerol. Rather, they have a mono- or oligosaccharide or, more loosely, a betaine group, each attached by an ether linkage to the glycerol C3 carbon. Figure 11.1.17 Structural features and nomenclature for glyceroglycolipids.

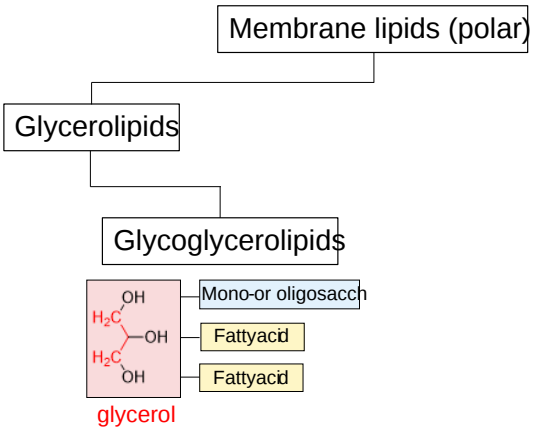


Figure **11.1.17**: Structural features and nomenclature for glyceroglycolipids

Figure **11.1.18** shows some examples of glyceroglycolipids.

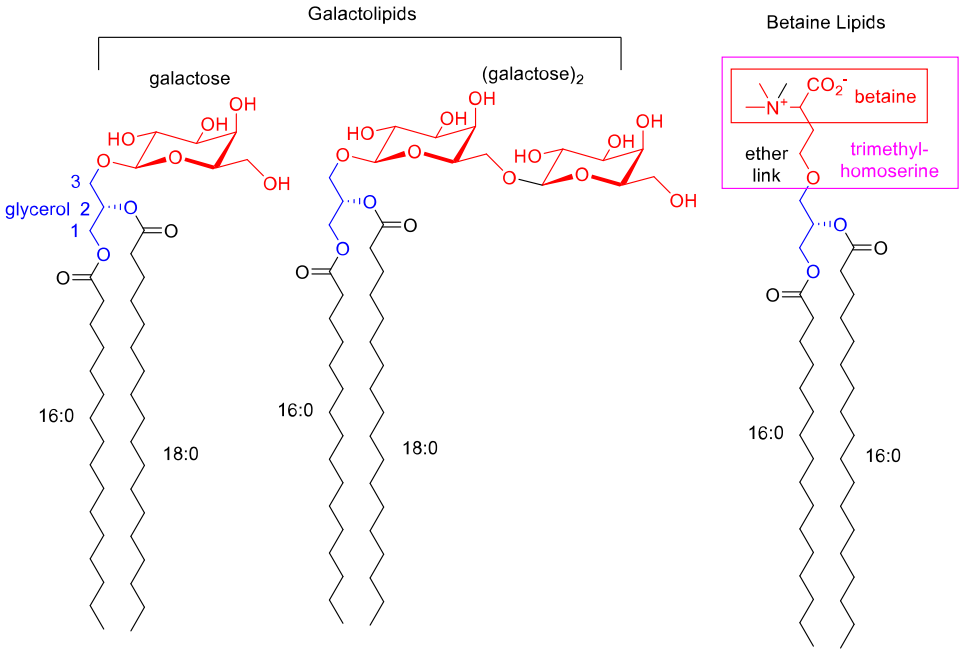


Figure **11.1.18**: Examples of glycoglycerolipids.

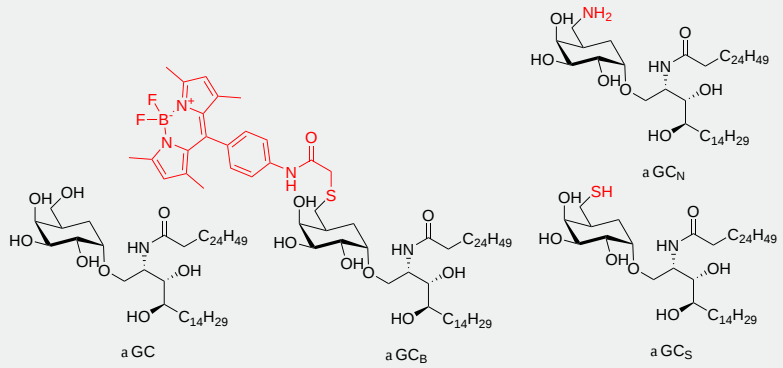
Again, there are an enormous number of different glycoglycerolipids, owing to the diversity of head groups and fatty acids esterified to glycerol at C1 and C2.

The betaine lipids (top right in the figure above) have neither a phosphate nor a sugar in the head group. The head group is betaine (trimethylglycine or any N-trimethylated amino acid). Betaine glycerolipids are found in lower eukaryotic organisms (algae, fungi, and some protozoa), photosynthetic bacteria, and spore-producing plants like ferns. Some would call these lipoamino acids.

A new malarial vaccine with a glycoglycerolipid adjuvant

It is extremely difficult to develop a vaccine against malaria. The resident time of the Plasmodium parasite in the blood after a mosquito bite is quite short (just a few hours), so any immune response against it must be very fast. The sporozoite parasite soon migrates to the liver, proliferating before leaving the liver and infecting red blood cells. The parasite is more sequestered in the liver from the antibody response engendered by typical vaccination, which promotes B-cell differentiation and proliferation into antibody-producing cells. Another type of lymphocyte, T cells, can also be activated in an immune response.

In particular, CD8 memory T cells (Trm cells) are produced in the liver against parasites, but high levels are needed. Mouse liver cells make modest levels of Trm cells when vaccinated with mRNAs encoding the Plasmodium's large ribosomal subunit protein L6 (RPL6). This is found in the liver for about a week after infection. However, when an "adjuvant" (something that increases the immune response, usually in a nonspecific way) containing α-galactosylceramide (αGC) is used, there is an increase in the production of liver Trm cells. Chemically modifying the αGC adjuvant greatly increases the production of Trm cells. The structure of αGC and its derivatives is shown below.



The adjuvant interacts as an agonist with type I natural killer T cells, activating them and amplifying the liver's immune response. Mice vaccinated with an mRNA vaccine containing the modified αGC adjuvant were afforded protection against the Plasmodium parasite. The protection extended to mice previously exposed to the parasite as well.

11.1.8: Membrane Lipids - Sphingolipids

Sphingolipids contain a sphingosine backbone and a fatty acid linked through an amide bond. The simplest sphingolipids are the ceramides, which are double-chain amphiphiles but without further modification of the functional groups on the polar head of sphingosine. The structures of both sphingosine and ceramide are shown below in Figure 11.1.19:

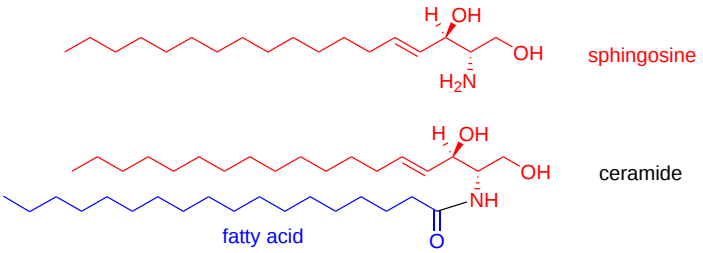


Figure 11.1.19: Structures of sphingosine and ceramide

Ceramides are abundant in the skin and help provide a protective barrier. Skin creams also contain ceramides. These lipids comprise only about 1% of body lipids, and over 200 different ceramides in humans arise from the different fatty acids linked to sphingosine. They have a key role in cell signaling and can also affect cardiovascular health in ways that are just being appreciated. Since they are double-chain amphiphiles, they are found in membranes where they alter the properties of the bilayer (including membrane fluidity). Ceramides with long acyl tails (especially those with 16, 18, or 24 carbons) are especially abundant in the skin.

Phosphosphingolipids and glycosphingolipids

These membrane lipids have a ceramide base but also contain modifications at the polar functional groups of the sphingosine head. Let's consider the phosphosphingolipids and the glycosphingolipids together. These groups do not use glycerol as a base for the attachment of fatty acids and head groups. Rather, they use the molecule sphingosine. Figure 11.1.20 shows sphingolipids' structural features and nomenclature.

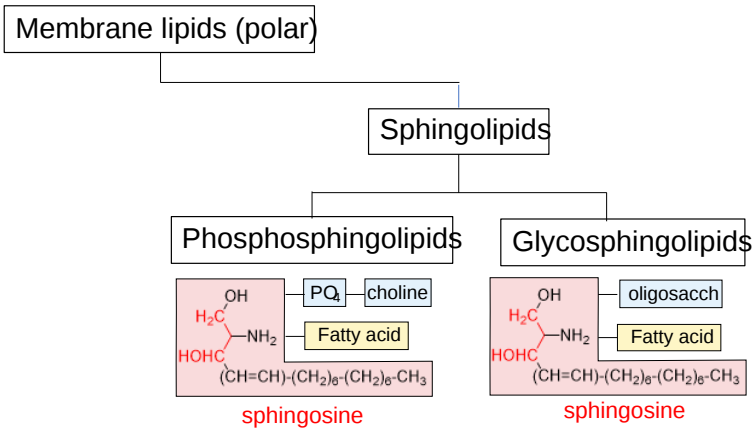


Figure 11.1.20: Structural features and nomenclature for sphingolipids

Examples of both classes are shown in Figure 11.1.21. Note that the base sphingosine (in red) provides an amine to attach a fatty acid through an amide bond and an OH for attachment to the head group.

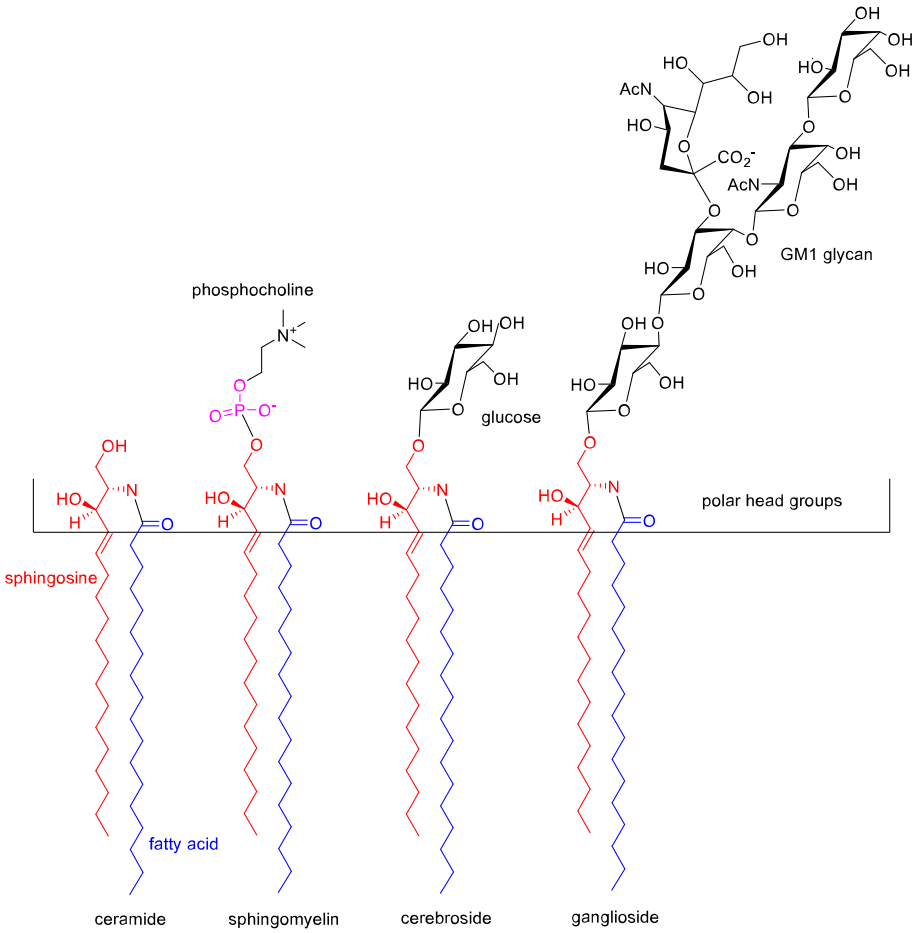


Figure 11.1.21: Examples of phosphosphingolipids and glycosphingolipids

Sugar-containing glycosphingolipids are **found largely in the outer face of plasma membranes.** The primary lipid of myelin, which coats neuronal axons and insulates them from loss of electrical signaling down the axon, is galactocerebroside

Figure 11.1.22 shows a summary of all of the different types of fatty acid-containing lipids

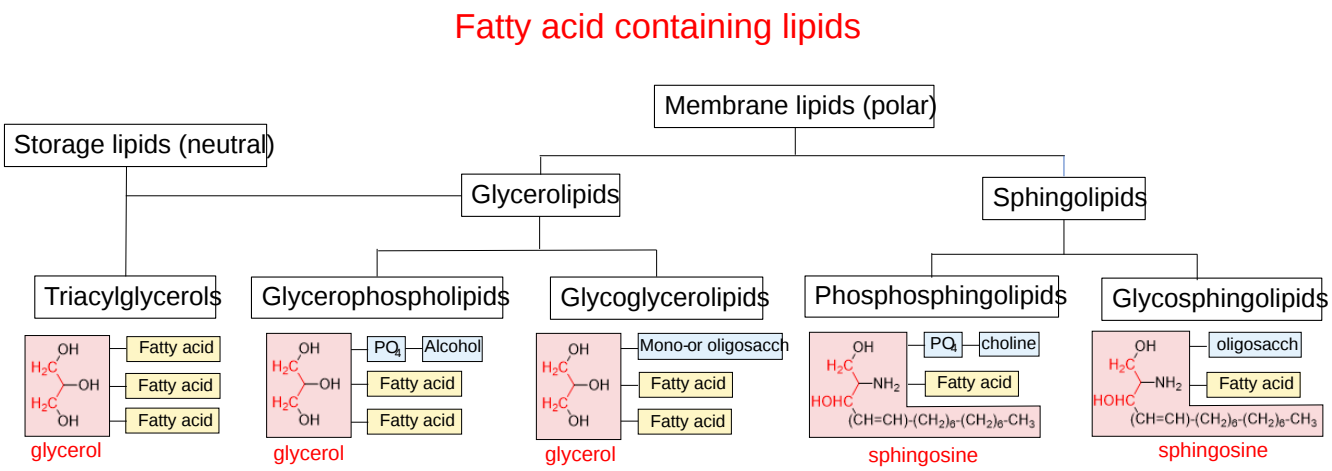


Figure 11.1.22: Summary of structural features and nomenclature for fatty acid-containing lipids

The complexity of biological lipids can be overwhelming. They are derived from two building blocks, ketoacyl and isoprene groups, which are modified with other groups. The Lipid Metabolites and Pathways Strategy (LIPID MAPS) consortium has offered a very detailed and encompassing definition of lipids: "hydrophobic or amphipathic small molecules that may originate entirely or in part by carbanion-based condensations of thioesters (fatty acids, glycerolipids, glycerophospholipids, sphingolipids, saccharolipids, and polyketides) and/or by carbocation-based condensations of isoprene units (prenol lipids and sterol lipids)." Eight different categories of lipids are listed in the parentheses above.

11.1.9: Shapes of membrane lipids

Let's look at the general shape of the double-chain amphiphiles that make bilayers. We saw the long-chain fatty or sulfate acids form conical structures that fit nicely together when they self-aggregate to form micelles. In contrast, membrane-forming double-chain amphiphiles have more cylindrical shapes that can't be fitted together in micelles but form a less curved bilayer structure, as shown in Figure 11.1.23. We will consider the variety of membrane structures in the next section.

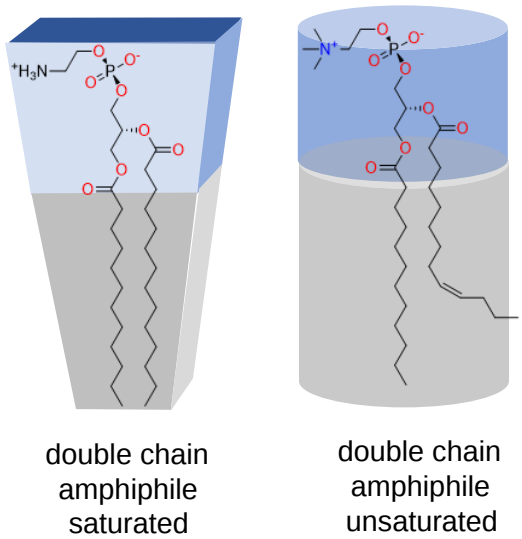


Figure 11.1.23: General shape of the double-chain amphiphiles that make bilayers

11.1.10: Triacylglycerol/Phospholipid Stereochemistry

Glycerol is an achiral molecule since C2 has two identical substituents, $-CH_2OH$. Biologically, glycerol can be chemically converted to chiral phospholipids (PL) of just one enantiomer. How can this be possible if the two CH_2OH groups on C2 of

glycerol are identical? It turns out that even though these groups are stereochemically equivalent, we can differentiate them as described in the figure below. Let's replace the -CH_2OH in one of the end carbons with -CH_2OD. With this simple change, glycerol is now chiral. Look at the top half of Figure **11.1.24**.

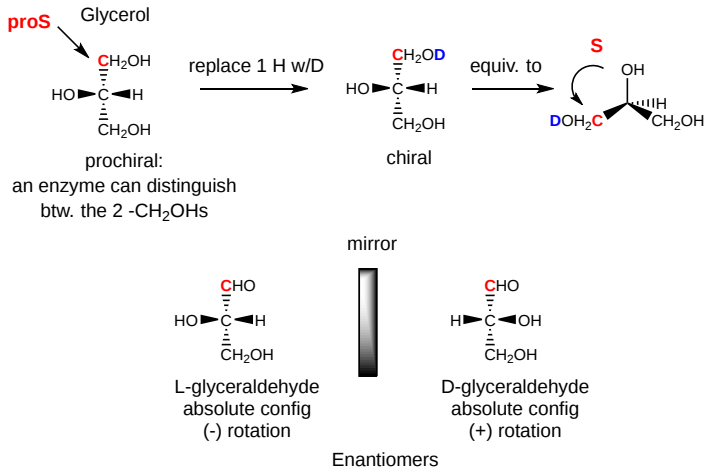


Figure **11.1.24**: Potential chirality of Glycerol - ProS

Glycerol is oriented with the OH on C2 (the middle carbon) pointing to the left. The OH of the top carbon in this orientation, C1, is replaced with OD, where D is deuterium, to make the molecule chiral (four different groups attached to C2). By rotating the molecule such that the H on C2 points to the back and assigning priorities to the other substituents on C2 (OH =1, $DOCH_2$ =2, and CH_2OH = 3), it can be seen that the resulting molecule is in the S configuration. We simply name the C1 carbon, which we modified with deuterium, the **proS** carbon. Likewise, if we replaced the OH on C3 with OD, we would form the R enantiomer. Hence, C3 is the proR carbon. This shows that we can actually differentiate between the two identical CH_2OH substituents. We say that glycerol is not chiral, but **prochiral**. (Think of this as glycerol has the potential to become chiral by modifying one of two identical substituents.)

In the bottom half of Figure **11.1.23**, we can relate the configuration of glycerol above (when OH on C2 is pointing to the left) to the absolute configuration of L-glyceraldehyde, a simple sugar (a polyhydroxyaldehyde or ketone), another 3C glycerol derivative. This molecule is chiral, with the OH on C2 (the only chiral carbon) pointing to the left. It is easy to remember that any **L** sugar has the OH on the **L**ast chiral carbon pointing to the **L**eft. The enantiomer (mirror image isomer) of L-glyceraldehyde is D-glyceraldehyde, in which the OH on C2 points to the right. Biochemists use L and D for lipid, sugar, and amino acid stereochemistry instead of the R and S nomenclature used in organic chemistry. The stereochemical designation of all the sugars, amino acids, and glycerolipids can be determined from the absolute configuration of L- and D-glyceraldehyde.

Now let's see how an enzyme can take a prochiral molecule like glycerol and phosphorylate only one of the -CH_2OHs to make one specific isomer, glycerol-3-phosphate, a key intermediate in the biosynthesis of phosphatidic acid (PA), a glycerophospholipid, as well as chiral triacylglycerols, shown in Figure **11.1.25**. The far left part of the pathway shows how the proR CH_2OH of glycerol is phosphorylated to produce one specific enantiomer, L-glycerol-3-phosphate. (The top part of the figure shows another way to make this molecule from glucose through the glycolytic pathway, which we will encounter in a future chapter.

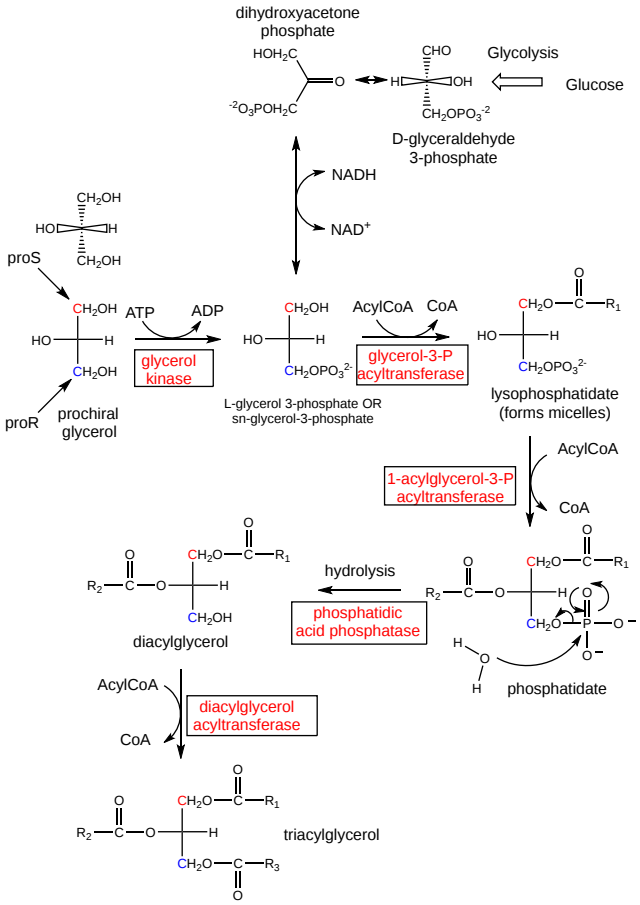
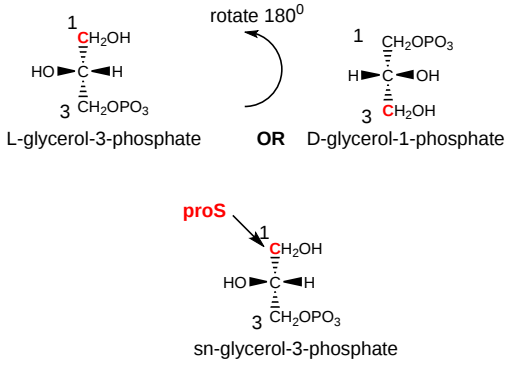


Figure 11.1.25: Synthesis of derivatives from prochiral glycerol

The first step (above figure) involves the phosphorylation of the OH on C3 by ATP (a phosphoanhydride similar in structure to acetic anhydride, an excellent acetylating agent) to produce the chiral glycerol phosphate. Based on the absolute configuration of L-glyceraldehyde and using this to draw glycerol (with the OH on C2 pointing to the left), we can see that the phosphorylated molecule can be named L-glycerol-3-phosphate. By rotating this molecule 180 degrees, without changing the stereochemistry of the molecule, we don't change the molecule at all, but using the D/L nomenclature above, we would name the rotated molecule D-glycerol-1-phosphate. This is illustrated in Figure 11.1.26.



Assign 1 position to group occupying the proS position of a prochiral center (glycerol). Orient molecule with C1 at top.

Figure 11.1.26: Stereospecific numbering system (sn) for glycerol

We can't give the same molecule two different names. Hence, biochemists have developed the **stereospecific numbering system (sn)**, which assigns the 1-position of a prochiral molecule to the group occupying the proS position. The proS C1 is, hence, at the sn-1 position. With that designation, C2 is at the sn-2 position, and C3 is at the sn-3 position. Using this nomenclature, we can see

that the chiral molecule described above, glycerol-phosphate, can be unambiguously named sn-glycerol-3-phosphate. The hydroxyl substituent on the proR carbon was phosphorylated.

Interestingly, archaea use isoprenoid chains linked by ether bonds to *sn*-glycerol 1-phosphate in their synthetic pathways. As noted above, bacteria and eukaryotes use fatty acids attached by ester bonds to *sn*-glycerol 3-phosphate

The enzymatic phosphorylation of the proR CH_2OH of glycerol to form sn-glycerol-3-phosphate is illustrated in Figure 11.1.27. As we could differentiate the two identical CH_2OH substituents as containing either the proS or proR carbons, so can the enzyme. The enzyme can differentiate identical substituents on a prochiral molecule if the prochiral molecule interacts with the enzyme at three points. Another example of a prochiral reactant/enzyme system involves the oxidation of the prochiral molecule ethanol by the enzyme alcohol dehydrogenase, in which only the proR H of the 2 H's on C2 is removed. (We will discuss this later.)

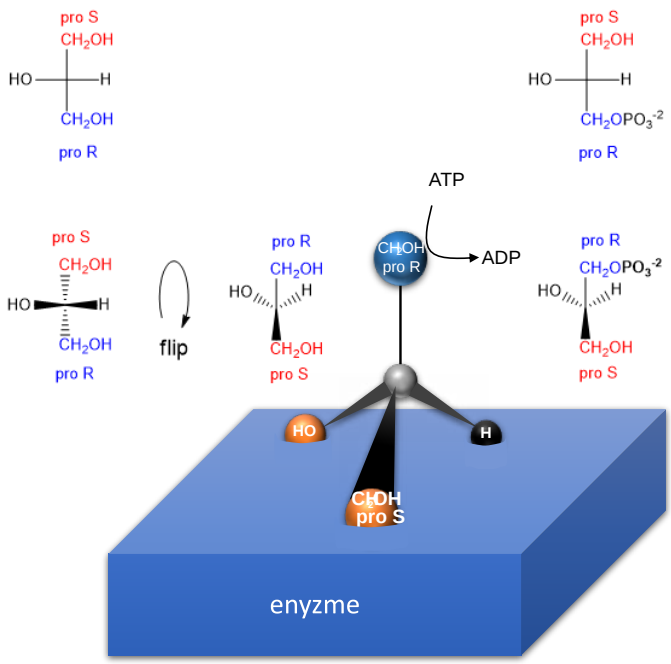


Figure 11.1.27: Enzymatic phosphorylation of the proR CH_2OH of glycerol to form sn-glycerol-3-phosphate

11.1.11: Isoprenoid-containing lipids

This is the last class of lipids we will consider. They do not contain fatty acids. Rather, they contain isoprene, a small branched alkadiene, which can polymerize into larger molecules containing isoprene monomer to form isoprenoids, often called terpenes. Instead of using isoprene as the polymerization monomer, either dimethylallyl pyrophosphate (DMAPP) or isopentenylpyrophosphate (IPP) is used biologically.

Figure 11.1.28 shows how DMAPP and IPP (both containing 5Cs) are used in a polymerization reaction to form geranyl-pyrophosphate (C10), farnesyl pyrophosphate (C15), and geranyl-geranyl pyrophosphate (C20).

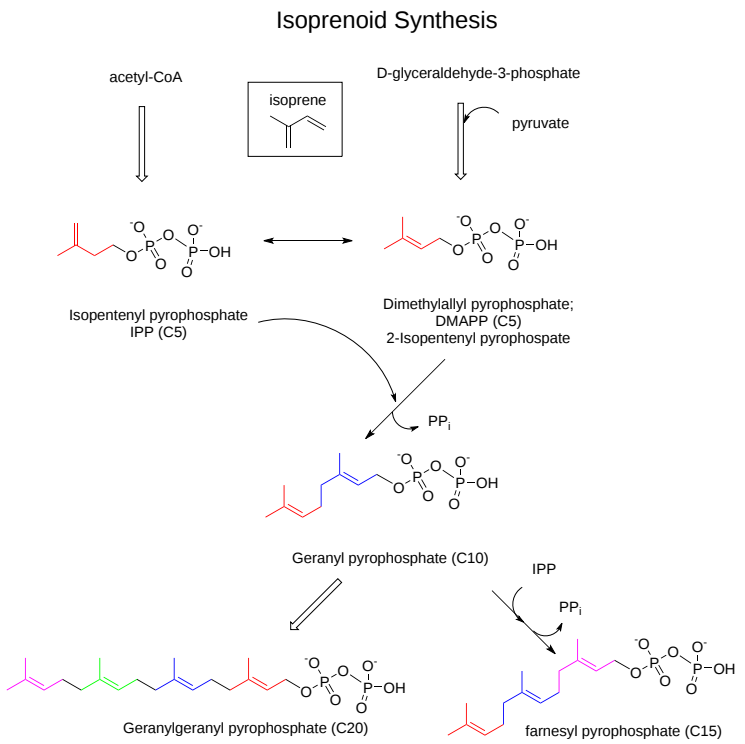


Figure 11.1.28: Synthesis of geranyl-pyrophosphate (C10), farnesyl pyrophosphate (C15) and geranyl-geranyl pyrophosphate (C20)

Many isoprenoid lipids are made from farnesyl pyrophosphate. For membrane purposes, the most important of these is cholesterol. Figure 11.1.29 shows an overview of the synthesis of cholesterol from two farnesyl pyrophosphates linking together in a "tail-to-tail" reaction to form squalene, a precursor of cholesterol. Each isoprene unit (5Cs) is shown in different colors for easier visualization.

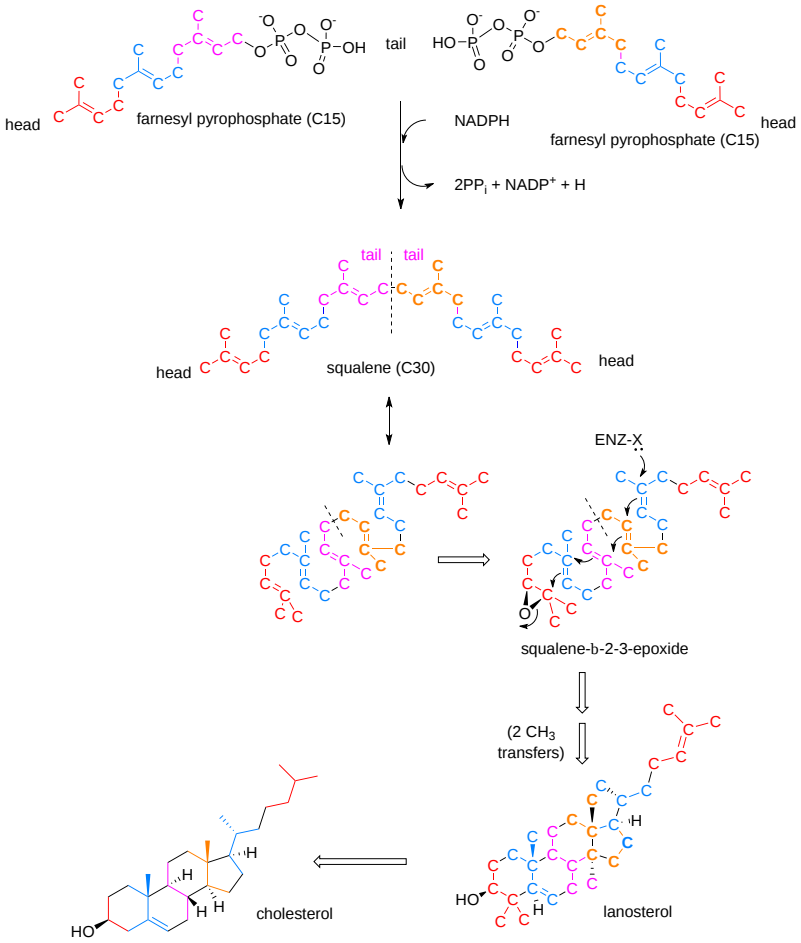


Figure 11.1.29: Synthesis of squalene from isoprene units

Figure 11.1.30 shows other biologically important isoprenoid-containing vitamins.

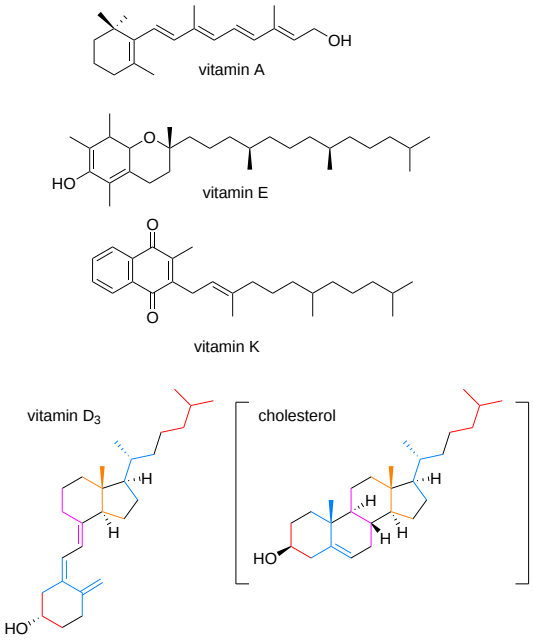
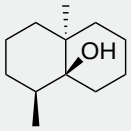


Figure 11.1.30: Isoprenoid vitamins

The smell of fresh rain

Everyone who has walked in the woods after a fresh rain knows the smell of a terpene called **geosmin,** "an Earthy-smelling substance" found in abundance in a group of substances collectively called petrichor, derived from the Greek petra (rock) and ichor (blood of gods). The structure of geosmin is shown below.



These substances are released when water falls on soils and rocks and seep into pores from which aerosols are released. Too much rain will saturate the pores in rocks and prevent the release of geosmin. Some insects, like the springtail, are attracted to geosmin. The synthesis of geosmin by bacteria and blue-green algae is controlled by transcription factors, which also affect bacterial spore formation. The attracted insects carry away the spores. Flies are repelled by geosmin, which is detected by a receptor at very low geosmin concentrations. In contrast, the *Aedes aegypti* mosquito is attracted to geosmin, which parallels the fact that the mosquito is unaffected by bacterial toxins.

11.1.12: One last look at lipid structure and shapes

With the exclusion of waxes and triacylglycerols, the other lipids we have discussed, including the mostly planar molecule cholesterol, are amphiphilic. Single-chain fatty acids form micelles, while lipids with two nonpolar chains and a polar/charged head group form bilayers. Given the relative size of the head group and the degree of unsaturation of the double bonds in fatty acids, the overall shapes of the membrane-forming lipids differ, as illustrated below. They are arranged like dominoes in the membrane based on their geometric volumes. Preferential clustering of identical types can cause local and extended changes in a prototypical bilayer structure.

Membranes and their components must be dynamic to enable all the functions and activities of a membrane. Ligands bind to membrane receptors (usually proteins), which can invaginate and pinch off to form an intracellular vesicle containing the receptor for processing. Likewise, vesicles can pinch off into the extracellular space. Cells must divide. Think of the membrane changes necessary for that! Also, consider that the length of cholesterol is half that of a typical double-chain amphiphile. Hence, it fits into just one of the bilayer lipid leaflets, where it modulates lipid bilayer properties. Figure **11.1.31** shows a series of lipids and their shape profiles.

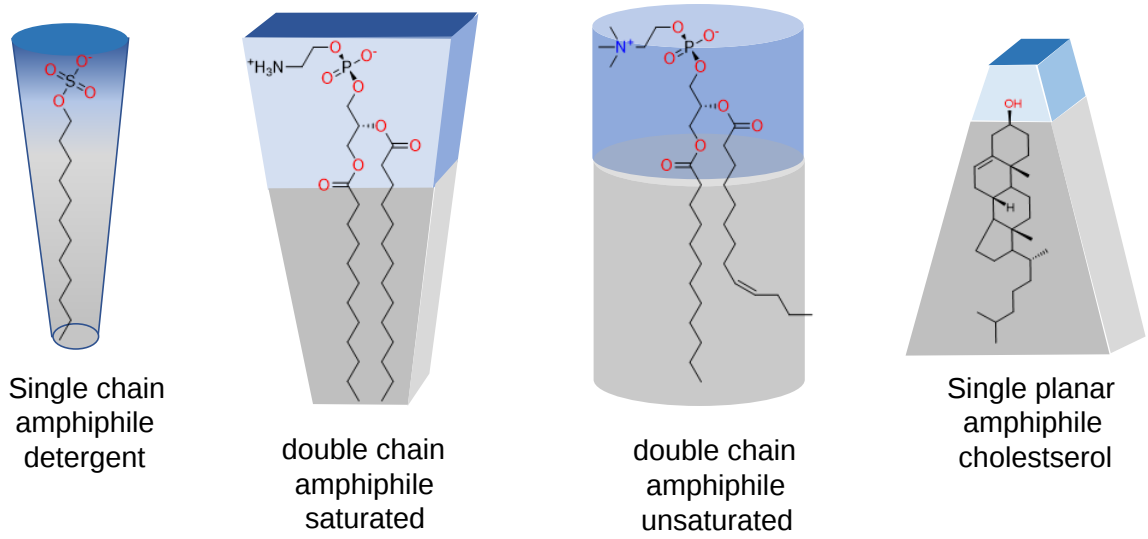


Figure **11.1.31**: A series of lipids and their shape profiles

We will consider membranes in greater detail in Section 10.3. Next, we will explore the properties of micelle and lipid droplet systems in more detail before addressing the more structurally complicated lipid bilayers.

11.1.13: Summary

This chapter explores the structural diversity and multifaceted functions of lipids in biological systems, emphasizing the fundamental principle that structure determines function. Lipids are defined as organic molecules that are soluble in organic solvents but largely insoluble in water due to their hydrophobic nature. This property underlies their roles in energy storage, insulation, mechanical protection, membrane formation, and cellular signaling.

The chapter begins by classifying lipids based on their chemical building blocks into those derived from fatty acids and those composed of isoprenoids. Fatty acids, which can be saturated or unsaturated, exhibit structural variations such as chain length and the presence of cis or trans double bonds. These variations significantly influence their physical properties—for example, melting points—by affecting the packing of acyl chains. Essential fatty acids from the n-6 and n-3 families are discussed, highlighting their importance in health, including their roles in brain development and cardiovascular protection.

In addition to free fatty acids, the chapter reviews complex lipids. Storage lipids, particularly triacylglycerols, serve as anhydrous energy reserves, contrasting with hydrated carbohydrates like glycogen. Membrane lipids are predominantly amphiphilic, with two distinct classes highlighted: glycerolipids (including glycerophospholipids and glyceroglycolipids) and sphingolipids. These lipids self-assemble into bilayers, forming the structural basis of cellular membranes. The stereochemistry of glycerol, although initially achiral, becomes crucial upon esterification, leading to specific configurations (using the stereospecific numbering system) that affect membrane function.

The chapter also examines isoprenoid lipids, which are synthesized from five-carbon isoprene units. These lipids, including cholesterol and various terpenes, play key roles in membrane structure, serve as precursors for vitamin synthesis, and contribute to sensory properties like the characteristic earthy smell of geosmin.

Finally, the text discusses the self-assembly behavior of lipids in aqueous environments. It explains how amphiphilic lipids form micelles and bilayers based on the geometric differences between single-chain (conical) and double-chain (cylindrical) structures, which are critical for processes such as membrane fluidity and vesicle formation.

Overall, the chapter illustrates how the unique chemical structures of lipids enable them to fulfill diverse biological roles—from energy storage to forming the dynamic boundaries of cells—and sets the stage for deeper exploration of membrane biochemistry and lipid signaling pathways in later sections.

This page titled 11.1: Introduction to lipids is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **10.1: Introduction to lipids** by Henry Jakubowski and Patricia Flatt is licensed CC BY-NC-SA 4.0.

11.2: Lipids Aggregates in Water - Micelles and Liposomes

Learning Goals (ChatGPT o3-mini, 2/2/25)

- **Describe the Behavior of Single-Chain Amphiphiles in Water:**
  • Explain how single-chain amphiphiles interact with water, including their ability to dissolve as monomers, form a surface monolayer, or self-aggregate into micelles when their concentration exceeds the critical micelle concentration (CMC).
- **Explain Micelle Formation Mechanisms:**
  • Describe the noncovalent interactions (e.g., induced dipole-induced dipole forces, ion-dipole interactions, and hydrogen bonding) that stabilize the interior of micelles and the exposed polar head groups, and explain how "like dissolves like" underpins these processes.
- **Analyze Thermodynamics of Self-Assembly:**
  • Apply the concepts of Gibbs free energy (ΔG), enthalpy (ΔH), and entropy (ΔS) to explain why micelle formation is enthalpically favored but entropically disfavored, and how the balance of these forces determines the spontaneous formation of micelles.
- **Differentiate Between Single-Chain and Double-Chain Amphiphile Aggregation:**
  • Compare the structural differences between single-chain amphiphiles (forming micelles) and double-chain amphiphiles (forming bilayers and vesicles), and explain how the number of hydrophobic chains and the area per head group influence aggregate shape and stability.
- **Understand Phase Behavior and Aggregate Transitions:**
  • Discuss how changes in conditions (e.g., concentration, temperature, pH) can lead to transitions in aggregate shape—from spherical micelles to cylindrical micelles and eventually to planar bilayers—and relate these transitions to changes in head group repulsion and chain packing.
- **Integrate Laboratory and In Vivo Perspectives:**
  • Relate the insights obtained from in vitro experiments with simple lipid solutions to the more complex in vivo environment, where lipids interact with proteins and other molecules, influencing membrane formation, dynamics, and cellular function.

These goals aim to deepen your understanding of how lipid structure governs self-assembly, the underlying thermodynamics of micelle and bilayer formation, and the broader relevance of these principles to biological membranes.

11.2.1: Single Chain Amphiphiles and Micelles

Understanding lipids in simple solutions is essential to understanding them *in vivo*. The same physical-chemical constraints would apply to the complex environment of the cell. What is different in the cell is that lipids are found in a cellular environment crowded with proteins that bind, synthesize, and break down lipids. Nevertheless, we can apply what we know from the test tube experiments to the cell.

To understand how molecules might react, it helps to pretend you are a molecule and ask yourself what you would do! We want to know how lipid molecules, specifically single and double-chain amphiphiles, interact with each other and the solvent when added to water. Before you read the answer, look at the image below and ask yourself: What would I do if I were a single chain amphiphile and jumped into water as shown in Figure 11.2.1?

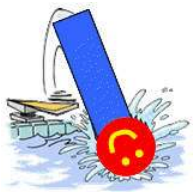


Figure 11.2.1: A cartoon single-chain amphiphile diving into a pool

Here is what they do. When added to water, some single-chain amphiphiles dissolve while others form a monolayer on the water's surface. If enough enter the solution and exceed their net solubility, they self-aggregate to form micelles. These outcomes are shown in Figure 11.2.2.

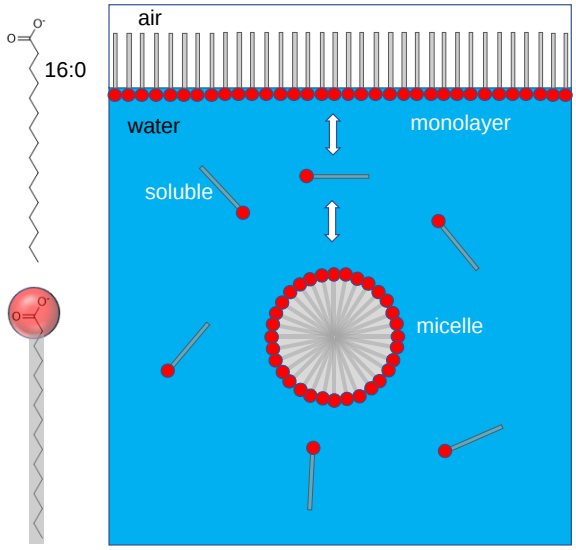
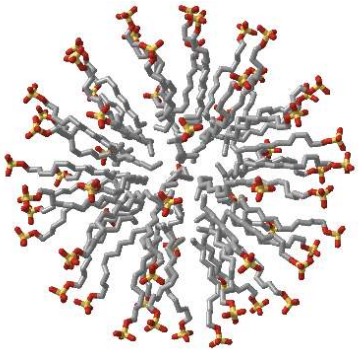


Figure 11.2.2: Distribution of single-chain amphiphiles in an aqueous solution

Figure 11.2.3 shows an interactive iCn3D model of an sodium dodecylsulfate (SDS) micelle



NCBI iCn3D Figure 11.2.3: Sodium dodecyl sulfate (SDS) micelle (Copyright; author via source).
Click the image for a popup or use this external link: not available

Double-chain amphiphiles, in contrast, form bilayers instead of micelles. (Note: single and double-chain amphiphiles can also form other multimolecular aggregate structures, but micelles and bilayers are the most common and the only ones we will consider.)

The micelle interior is completely nonpolar. Spherical bilayers that enclose an aqueous compartment are called vesicles or liposomes. Micelles and bilayers, formed from single and double-chain amphiphiles, respectively, represent noncovalent aggregates and are formed by an entirely physical process. No covalent steps are required.

Common single-chain amphiphiles that form micelles are detergents (like sodium dodecyl sulfate - SDS) and fatty acids. Sodium hydroxide feels slippery on your skin since the base hydrolyses the fatty acids esterified to skin lipids. The free fatty acids then aggregate spontaneously to form micelles that act like detergents and are slippery.

Micelles/detergents in water are an example of an emulsion of two materials that are generally immiscible in each other unless one is dispersed into small droplets in the other. Fine oil drops can be dispersed in water, and fine aqueous drops can be dispersed in a nonpolar liquid. Many vaccines are formulated as this latter type of emulsion. Grease and oil in your clothes can be carried away by

"diving" into the nonpolar part of the detergent micelle, which is dispersed in water as an emulsion. Another example of an emulsion or a colloid is a cloud, a dispersion of liquid water droplets in a solvent, the atmosphere.

The formation of these structures can be understood from the study of noncovalent interactions, as well as through thermodynamics. In a micelle, the buried acyl chains can interact and be stabilized by induced dipole-induced dipole forces as the nonpolar carbons and hydrogen are in van der Waals contact. They are sequestered from water. This view fits our simple axiom of "like-dissolves like." The polar head groups can be stabilized by ion-dipole bonds between charged head groups and water. Likewise, H-bonds between water and the head group stabilize the exposed head groups in water. Repulsive forces may also be involved. Head groups can repel each other through steric factors or ion-ion repulsion from like-charged head groups. The attractive forces must be greater than the repulsive forces, which lead to these molecular aggregates.

From a thermodynamic approach, one problem arises with this simple explanation. For a micelle or bilayer to form, many monomers must aggregate to form a single micelle or vesicle, which is entropically disfavored! So, let's delve into the thermodynamics of micelle formation.

ΔG, the free energy change for a reaction, determines the spontaneity and extent of a chemical or physical reaction. The system's free energy depends on three variables: temperature T, pressure P, and n, the number of moles of each substance. For the latter, think of solute X on two different sides of a permeable membrane. If the concentration of X is the same on each side, as shown in the system below, the system is in equilibrium, as shown in Figure **11.2.4**.

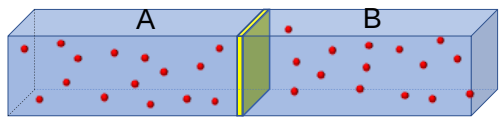


Figure **11.2.4**: Equilibrium state of a species distributed in solutions separated by a semipermeable membrane.

If the system is composed of two different parts, A and B, the system is at equilibrium (ΔG=0) if $T_A = T_B$, $P_A = P_B$, and the change in the absolute free energy per mole of A is $\Delta G_A/\Delta n = \Delta G_B/\Delta n$. More precisely, using simple calculus, we would discuss incremental changes in absolute free energy/mol, dG_A/dn for A (often called the chemical potential of A, μ_A), and dG_B/dn or μ_B)for B. At equilibrium, $dG_A/dn = dG_B/dn$. (We will use the symbol G here instead of μ). G is the absolute free energy/mol (again often called the chemical potential), where $G=G^o + RTln[A]$. The equations you used in introductory chemistry can be written.

$$\begin{aligned} \Delta G &= \Delta G^0 + RTInQr \\ \Delta G &= \Delta H - T\Delta S \\ \Delta G^0 &= \Delta H^0 - T\Delta S^0 \\ \Delta G^0 &= -RTInKeq \end{aligned} \tag{11.2.1}$$

Now, let's apply this to the chemical equation for micelle formation:

n SCA <==> 1 micelle

where SCA represents a single-chain amphiphile. At first glance, we might suspect that:

- $\Delta H^0 < 0$ since the induced dipole-induced dipole interactions among the buried acyl chains in the micelle would be much more favorable than the water-acyl interactions for the monomeric amphiphile in the solution. Our aphorism supports this notion, "like dissolves like." Of course, we couldn't ignore polar interactions (H bonding, for example) among the head groups and water. Still, we might expect these to be equally favorable in the monomeric and micellar states.
- $\Delta S^0 < 0$ since we are forming a very ordered state (a single micelle) with much less entropy from a state (single chains of amphiphiles dispersed in solution) with much more entropy.

Hence, it would appear that micelle formation is enthalpically favored but entropically disfavored. Let's examine this issue more closely. First, we need to obtain a greater understanding of ΔG^o, which should give us a clue as to where an SCA would "want" to be in this mixture. Remember, ΔG^0 is a constant at a given T, P, and solvent conditions and depends only on the relative stability of a molecule in a given environment and not its concentration.

Traube, in 1891, noticed that single-chain amphiphiles tend to migrate to the surface of the water and decrease its surface tension (ST). He observed that the decrease in ST is directly proportional to the amount of amphiphile added until a certain point, at which

point the added amphiphile has no additional effect. In other words, the response of ST saturates at some point.

We are more interested in what happens to amphiphiles in bulk water, not at the surface. As we showed in Figure 2 above, monomeric single-chain amphiphiles are in equilibrium with single-chain amphiphiles in micelles. Assume you have a way to measure a monomeric single-chain amphiphile in solution. What happens to its concentration as you add more and more SCA to the mixture? You observe the same effect that Traube noted with changes in surface tension. This explanation goes like this: as more amphiphile is added, more goes into the bulk solution as monomers. At some point, enough amphiphiles are added to form micelles. After this point, added amphiphiles form more micelles, and no further increases in monomeric single-chain amphiphiles are noted. The concentration of amphiphile at which this occurs is the critical micelle concentration (CMC). Figure 11.2.5 shows a graph of a monomeric single-chain amphiphile in solution versus the concentration added to the solution.

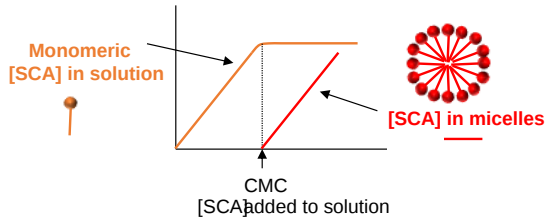


Figure 11.2.5: Monomeric single-chain amphiphile in solution versus the concentration added to the solution

This saturation effect can be observed with other systems as well.

- Consider the amount of NaCl(aq) in the solution as more NaCl(s) is added to water. The water is saturated with dissolved NaCl at some point, and no further increase in NaCl (aq) occurs.
- Consider the amount of a sparingly soluble hydrocarbon (HC) in water. After saturation, phase separation occurs.

Now consider the addition of a drop of a slightly soluble hydrocarbon liquid (HC_L) into water, as pictured in the diagram below. At t=0, the system is not at equilibrium, and some of the HC will transfer from the pure liquid to water, so at time t=0, $\Delta G_{TOT} < 0$. This is illustrated in Figure 11.2.6.

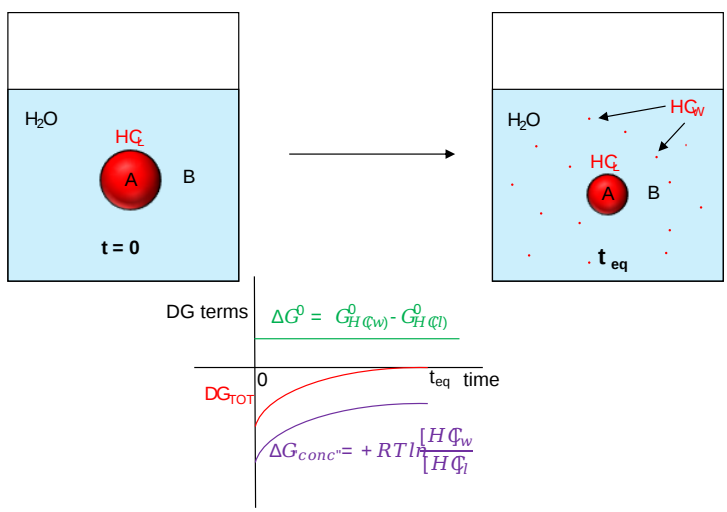


Figure 11.2.6: Addition of slightly soluble liquid hydrocarbon to water - ΔG vs time

The following equations can be derived.

$$\begin{gathered}\Delta G_{TOT} = (G_{HC-W}) - (G_{HC-L}) = G^0_{HC-W} + RT\ln[HC]_W - (G^0_{HC-L} + RT\ln[HC]_L) = \\ \Delta G_{TOT} = (G^0_{HC-W} - G^0_{HC-L}) + RT\ln([HC]_W - \ln[HC]_L) = \\ \Delta G_{TOT} = \Delta G^0 + RT\ln\frac{[HC]_W}{[HC]_L}\end{gathered} \tag{11.2.2}$$

Now, add a bit more complexity to the last example. Add a hydrocarbon x to a biphasic system of water and octanol and shake it vigorously as shown in Figure 11.2.7. At equilibrium, x would have "partitioned" between the two mostly immiscible phases.

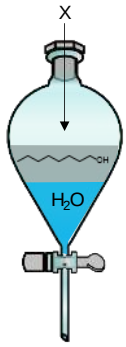


Figure 11.2.7: Addition of a hydrocarbon x to a biphasic system of water and octanol

A simple favorable reaction can be written for this system: x aq → x oct.

If X is a hydrocarbon, $\Delta G < 0$. Also, $\Delta G^o < 0$, since this term is independent of concentration and depends only on the intrinsic stability of x in water compared to that of octanol. This simple equation holds:

$$\Delta G_{TOT} = (G^0_{X-oct} - G^0_{X-w}) + RT \ln \frac{[X]_{oct}}{[X]_w} = \Delta G^0 + RT \ln \frac{[X]_{oct}}{[X]_w} \tag{11.2.3}$$

At equilibrium, $\Delta G^0=0$ and the equation can be rewritten as:

$$\Delta G^0 = -RT \ln \frac{[X]_{oct}}{[X]_w} = -RT \ln K_{part} \tag{11.2.4}$$

K_{part} is the equilibrium partition coefficient for X in octanol and water. K_{part} can readily be determined in the lab. Just shake a separatory flask with a biphasic system of octanol and water after injecting a bit of X. Then separate the layers and determine the concentration of X in each phase. Plug these numbers into the last equation. You should be able to predict the sign and relative magnitude of ΔG^o since it does not depend on concentration but only on the intrinsic stability of the molecules in the different environments. K_{paft} values are often determined for drugs since they often must diffuse across cell membranes to move into the cytoplasm, where they can act. Drugs, hence, must have a reasonable Kpart to pass through the membrane but not so high that they are insoluble.

11.2.2: Double Chain Amphiphiles and Bilayers

In contrast to single-chain amphiphiles, double-chain amphiphiles added to water form monolayers and vesicles called liposomes, as shown in Figure 11.2.8.

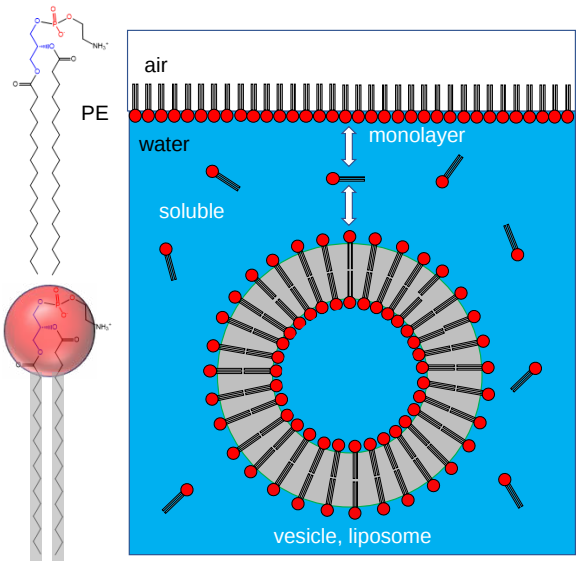


Figure 11.2.8: Distribution of double-chain amphiphiles in an aqueous solution

They can be unilamellar, which consists of a single bilayer surrounding the internal aqueous compartment, or multilamellar, which consists of multiple bilayers surrounding the enclosed aqueous solution. Figure 11.2.9 below shows images of a cross-section of a liposome/vesicle (diameter around 25 nm, so it is considered a small unilamellar vesicle or SUV). The bilayer is composed of 3298 DMPC (dimyristoylphosphatidylcholine) double-chain amphiphiles. The left image shows water molecules (over 9500) inside and outside the vesicle and the DMPCs as sticks, while the right image shows them in spacefill. Embedded in the top of the bilayer is the transmembrane helix of the cytokine receptor common subunit beta (PDB ID 2NA8), shown in red spacefill.

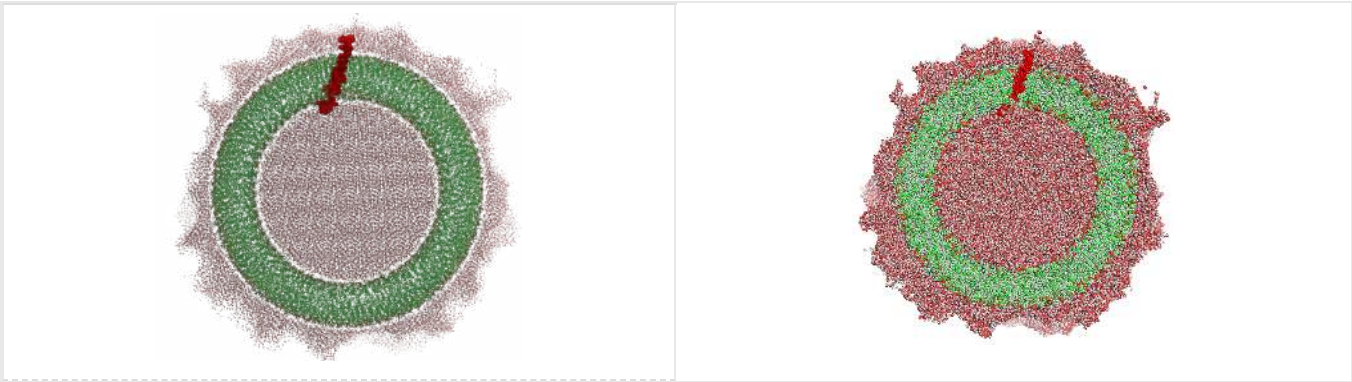


Figure 11.2.9: Cross-section of a liposome/vesicle of diameter around 25 nm comprised of DMPC (dimyristoylphosphatidycholine), the transmembrane helix of the cytokine receptor common subunit beta (PDB ID 2NA8), and water molecules both inside and outside of the bilayer. PDB files and the images were made using MolCube and Pymol, respectively.

Figure 11.2.10 shows an interactive iCn3D model of 2NA8 with the bilayers shown as layers of dummy atoms and with no water molecules.

Figure 11.2.10: 2NA8 with the bilayer represented with dummy atoms and with no water molecules. (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vzJVvf3kSVuYb8

You can imagine that multilamellar vesicles resemble an onion with multiple layers. Cartoons of unilamellar and multilamellar liposomes are shown in Figure 11.2.11, where each concentric circle represents a bilayer.

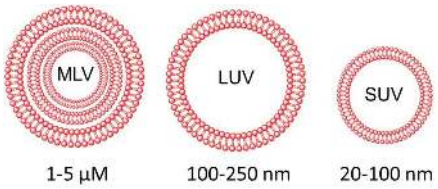


Figure 11.2.11: Classification of liposomes based on the lamellarity. Pandey, H., Rani, R., & Agarwal, V. (2016). Liposomes and Their Applications in Cancer Therapy. Brazilian Archives of Biology and Technology, 59, e16150477. https://doi.org/10.1590/1678-4324-2016150477
Creative Commons Attribution License

Liposomes vary in diameter. They can be generally categorized into small (S, diameter < 25 nm), intermediate (I, diameter around 100 nm), and large (L, diameter from 250-1000 nm). If these vesicles are unilamellar, they are abbreviated as SUV, IUV, and LUV.

The chemical composition of liposomes made in the lab can be widely varied. Most contain neutral phospholipids like phosphatidylcholine, phosphatidylethanolamine (PE) or sphingomyelin (SM), supplemented, if desired, with negatively charged phospholipids, like phosphatidyl serine (PS) and phosphatidyl glycerol (PG). In addition, single-chain amphiphiles like cholesterol (C) and detergents can be incorporated into the bilayer membrane, which modulates the fluidity and transition temperature (Tm) of the bilayer. If present in too great a concentration, single-chain amphiphiles like detergents, which form micelles, can disrupt the membrane so completely that the double-chain amphiphiles become incorporated into detergent micelles, now called mixed micelles, in a process that effectively destroys the membrane bilayer.

The properties of liposomes (charge density, membrane fluidity, and permeability) are determined by the lipid composition and size of the vesicle. The desired properties will be, in turn, determined by the use of the particular liposome. The vesicles offer wonderful, simple models to study the biochemistry and biophysics of natural membranes. Membrane proteins can be incorporated into the liposome bilayer using the exact method you will be using. However, apart from these purposes, liposomes can encapsulate water-soluble molecules such as nucleic acids, proteins, and toxic drugs. These liposomes can be targeted to specific cells if antibodies or other molecules that will bind specifically to the target cell can be incorporated into the bilayer of the vesicle. Intraliposomal material may then be transferred into the cell by fusion of the vesicle with the cell or by endocytosis of the vesicle.

Liposomes could also be called lipid nanoparticles, with sizes ranging up to 1000 nm. Liposomes are vesicular, with small aqueous-filled compartments. They can be made with encapsulated drugs for delivery to target cells through blood transport. They act as an emulsion in water.

Lipid nanoparticles can also be particulate (insoluble), which slowly degrade and release their contents in situ. Most recently, particulate lipid nanoparticles have helped save the world by being used to encapsulate messenger RNA (mRNA) for the spike protein from the SARS-CoV-2 virus, which causes the COVID-19 pandemic. These lipid nanoparticles are used in the vaccine against the coronavirus. The mRNA that encodes part of the spike protein is "encapsulated" in the lipid nanoparticle. The mRNA

contains specially modified nucleotides to increase their stability. The lipid nanoparticles also contain positively charged lipids, which help stabilize the negatively charged mRNA from degradation. The nanoparticles are endocytosed into cells, where the mRNA can be translated into the SARS-CoV-2 spike protein, required for whole virus entry into the cell. The immune system then recognizes the spike protein.

The lipids used in the formulation of these nanoparticles include fatty acids, mono-, di-, and triglycerols, glycerophospholipids, waxes (like cetyl palmitate), and other positively charged lipids including stearyl amine, benzalkonium chloride, cetrimide, cetyl pyridinium chloride, and dimethyldioctadecylammonium bromide. These are shown in Figure 11.2.12

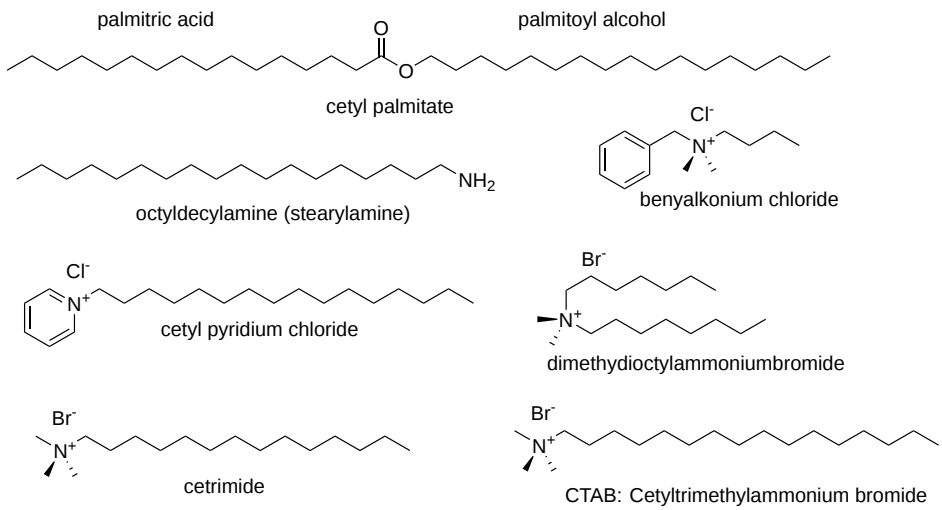


Figure 11.2.12: Amphiphiles used to make lipid nanoparticles

11.2.3: Why Micelles and Bilayers?

Micelles and liposomes form spontaneously - i.e., $\Delta G < 0$. But why do single-chain amphiphiles form micelles and double-chain amphiphiles form bilayers? Let's think about this from a thermodynamic and structural point of view.

As the number of Cs in the nonpolar carbon (NC) chain increases, the ΔG for transferring into a micelle, or by analogy, for a single chain amphiphile entering a micelle, becomes more and more negative (i.e., more favored). The following equation seems to apply to the transfer of a single-chain amphiphile into a micelle:

$\Delta G^o = G^o$ (mic) - G^o (aq) = + number - 709 NC

$$\Delta G^\circ = G^\circ\ (\text{mic})\ - G^\circ(\text{aq}) = +\ \text{number}\ - 709\text{NC} \tag{11.2.5}$$

where NC is the number of carbon atoms in the chain. The first positive term depends on the nature of the head group, while the second negative term is independent of the head group. These + and - terms bring us back to the principle of opposing forces we discussed when we looked at the noncovalent interactions involved in micelle and bilayer formation.

There are attractive interactions, including induced dipole-induced dipole interactions among the chains and dipole-ion and H-bond interactions with water and the head groups. Similar repulsive interactions arise from steric hindrance with bulky heady groups and ion-ion repulsions. Of course, there are also entropic considerations. Let us now consider these factors as we explore what might happen to a preformed micelle as we try to put more single-chain amphiphiles (**sca**) into it.

As we increase the number of Cs in the SCA, the micelles would have a larger radius. For a given SCA with a fixed number of Cs, once a spherical micelle is formed, it can no longer retain its spherical shape if more SCAs are added. Imagine increasing the diameter of a spherical micelle 10x. A large part of the inside would have no atoms or be filled with water, which would not be favorable. Therefore, if the micelle is to grow, it can do so only by changing shape to something other than a sphere. By squeezing a tennis ball, one can imagine that the shape could distort to a circular cylinder with end caps. In this way, the acyl chains can still interact. The only problem is that head groups will now be closer than they were in the sphere. This is simplistically illustrated in Figure 11.2.13.

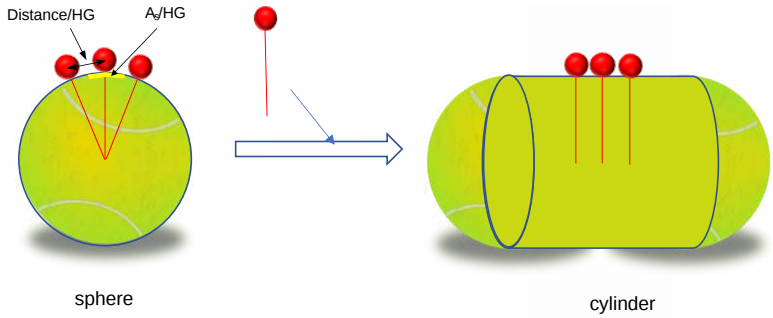


Figure 11.2.13: Hypothetical change in head group interactions on the conversion of a spherical vesicle to a cylindrical vesicle

Imagine that in a sphere, the head groups radiate perpendicularly from the sphere's surface. As the sphere is distorted into a cylinder, the head groups come closer together, so they will experience more steric interference. If a cylinder can be formed, however, it could continue to grow as long as needed with no further compression required. Imagine you further compressed the cylinder into a planar "bilayer" structure. The head groups would be even closer and experience even more repulsion. This bilayer will not form since growth can occur in the cylindrical phase without the added repulsion.

Now consider a double-chain amphiphile (DCA). In the case of an SCA, the number (N) of head groups (HG) = the number of acyl chains (CH). Hence, the surface area per HG equals the area per HC. Or: As/N HG = As/ N CH. For a DCA, N HG = N CH/2, therefore As/N HG = 2As/NCH. There is twice the surface area available per head group compared to the SCA. Therefore, the DCA can tolerate more compression. It can easily be compressed to a bilayer, which, as we saw, has much less As/HG. The cylindrical form has too much space per head group since water can enter the structure. The extra closeness of head groups in the bilayer can be tolerated even more since the ΔG^o for transfer of a DCA into a micelle is 60% more negative than that of an SCA. The As/HG for closed vesicles differs slightly from that of a truly planar bilayer since the vesicles are so large compared to a micelle.

Once again, we have discovered that structure mediates function. We can account for the fact that SCA and DCA form micelles and bilayers, respectively, by understanding the structure of the monomers!

In reality, things are more complicated.

The general rule holds that single-chain amphiphiles form micelles, and double-chain amphiphiles form bilayers. However, single-chain fatty acids can form bilayers under the right conditions if the pH is low enough that the head group is protonated and uncharged. Why would that make a difference? Fatty acids like oleic acid would be a candidate for components of the membranes of protocells in the evolution of life from abiotic conditions. Likewise, short double-chain amphiphiles can make micelles. A combination of double-chain amphiphiles with either short double-chain amphiphiles or single-chain amphiphiles can form a **bicelle** (a single structure with properties of both a bilayer and a micelle). In addition, other lipid phases can be observed. Which aggregates or phases ultimately form depends on the structure of the lipid, the solvent conditions, and the temperature. These include the following phases:

- lamellar gel (Lb) and lamellar liquid crystalline (La) phases
- hexagonal HI (cylinders packed in the shape of a hexagon with polar heads facing out into the water
- hexagonal HII (cylinders packed in the shape of a hexagon with acyl chains pointing out as in reverse micelles
- micellar (M).

We will discuss them in more detail in the next section.

11.2.4: Summary

This chapter explores how the physical-chemical properties of lipid molecules dictate their self-assembly in aqueous environments. This essential concept underpins our understanding of biological membranes and detergent function. Focusing on both single-chain and double-chain amphiphiles, the text explains how these molecules interact with water and with each other to form well-defined supramolecular structures.

Key Concepts:

- **Self-Assembly of Amphiphiles:**
  When introduced into water, single-chain amphiphiles can exist as dissolved monomers or form a monolayer at the water surface. Above a specific concentration, known as the critical micelle concentration (CMC), these molecules aggregate into micelles—spherical structures in which the hydrophobic acyl chains are sequestered away from water, while the polar head groups remain exposed to the aqueous environment. This behavior exemplifies the principle of "like dissolves like" and is governed by a balance of noncovalent interactions, such as van der Waals forces, ion-dipole interactions, and hydrogen bonding.
- **Thermodynamics of Micelle Formation:**
  The chapter delves into the thermodynamics driving self-assembly, using Gibbs free energy (ΔG) to describe the spontaneity of micelle formation. While aggregation is enthalpically favored (due to favorable interactions among hydrophobic chains), it is entropically disfavored because it leads to a decrease in molecular disorder. The net result—governed by the balance between these opposing forces—determines the stability and size of the micelle.
- **Comparison with Double-Chain Amphiphiles:**
  In contrast to single-chain amphiphiles, double-chain amphiphiles (such as those found in membrane lipids) form bilayers rather than micelles. The presence of two hydrophobic chains per molecule allows for tighter packing and reduced steric repulsion between head groups, favoring the formation of planar bilayers that can assemble into vesicles (liposomes) with an internal aqueous compartment.
- **Aggregate Phase Behavior:**
  The chapter also introduces the idea that the shape and size of lipid aggregates depend on the interplay between attractive forces (e.g., hydrophobic interactions) and repulsive forces (e.g., steric and ionic repulsions). As conditions change (e.g., concentration, temperature, pH), micelles can transition into other structures—such as cylindrical micelles or bilayers—demonstrating the versatile nature of lipid self-assembly. Various phase structures (lamellar, hexagonal, and micellar) are briefly mentioned as part of this dynamic behavior.
- **Biological Relevance:**
  While these self-assembly processes are illustrated with simple in vitro systems, the same principles apply in the crowded, complex environment of cells. Here, lipids interact with proteins that bind, synthesize, or degrade them, yet the underlying thermodynamic and structural principles remain fundamentally important for understanding membrane formation and function.

In summary, the chapter ties together the structural characteristics of amphiphilic molecules with the thermodynamic forces driving their assembly into micelles and bilayers. This knowledge provides a critical foundation for understanding the behavior of lipids in both laboratory settings and living cells, and sets the stage for deeper exploration into membrane dynamics and lipid-protein interactions in subsequent chapters.

This page titled 11.2: Lipids Aggregates in Water - Micelles and Liposomes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **10.2: Lipids Aggregates in Water - Micelles and Liposomes** by Henry Jakubowski and Patricia Flatt has no license indicated.

11.3: Structure and Function - Lipids and Membranes

Source: BiochemFFA_2_7.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Lipids are a diverse group of molecules that all share the characteristic that at least a portion of them is hydrophobic. Lipids play many roles in cells, including serving as energy storage (fats/oils), constituents of membranes (glycerophospholipids, sphingolipids, cholesterol), hormones (steroids), vitamins (fat soluble), oxygen/ electron carriers (heme), among others. For lipids that are very hydrophobic, such as fats/ oils, movement and storage in the aqueous environment of the body requires special structures. Other, amphipathic lipids, such as glycerophospholipids and sphingolipids spontaneously organize themselves into lipid bilayers when placed in water. Interestingly, major parts of many lipids can be derived from acetyl-CoA.

Fatty acids

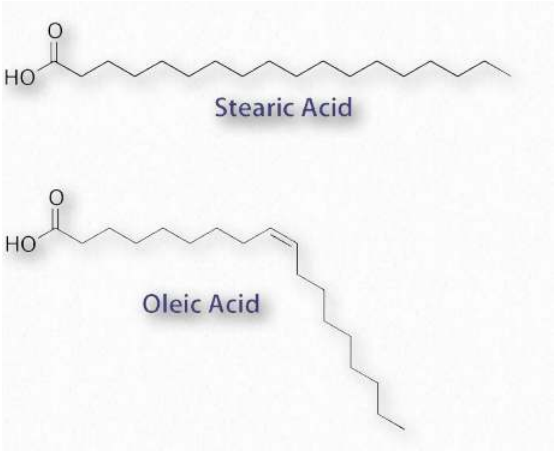


Figure 2.190 - Saturated fatty acid (stearic acid) and unsaturated fatty acid (oleic acid)

The most ubiquitous lipids in cells are the fatty acids. Found in fats, glycerophospholipids, sphingolipids and serving as as membrane anchors for proteins and other biomolecules, fatty acids are important for energy storage, membrane structure, and as precursors of most classes of lipids. Fatty acids, as can be seen from Figure 2.190 are characterized by a polar head group and a long hydrocarbon tail. Fatty acids with hydrocarbon tails that lack any double bonds are described as saturated, while those with one or more double bonds in their tails are known as unsaturated fatty acids. The effect of double bonds on the fatty acid tail is to introduce a kink, or bend, in the tail, as shown for oleic acid.

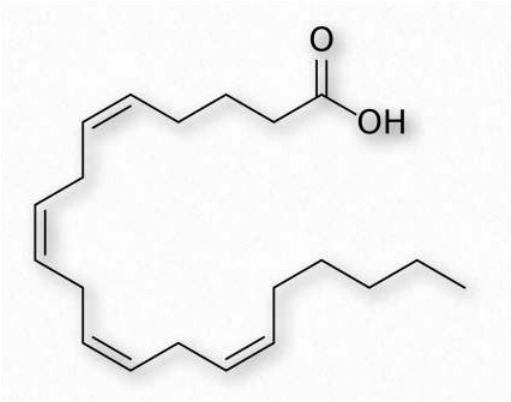


Figure 2.191 - Arachidonic acid - A polyunsaturated fatty acid Wikipedia

Stearic acid, a saturated fatty acid, by contrast has a straight hydrocarbon tail. Figures 2.190-2.194 show the most common saturated and unsaturated fatty acids. Fatty acids with unsaturated tails have a lower melting temperature than those with saturated tails of the same length. Shorter tails also decrease melting temperature. These properties carry over to the fats/oils containing them.

Common name	Chemical structure	
Caprylic acid	$CH_3(CH_2)_6COOH$	8
Capric acid	$CH_3(CH_2)_8COOH$	10
Lauric acid	$CH_3(CH_2)_{10}COOH$	12
Myristic acid	$CH_3(CH_2)_{12}COOH$	14
Palmitic acid	$CH_3(CH_2)_{14}COOH$	16
Stearic acid	$CH_3(CH_2)_{16}COOH$	18
Arachidic acid	$CH_3(CH_2)_{18}COOH$	20
Behenic acid	$CH_3(CH_2)_{20}COOH$	22
Lignoceric acid	$CH_3(CH_2)_{22}COOH$	24
Cerotic acid	$CH_3(CH_2)_{24}COOH$	26

Figure 2.192 - Saturated fatty acids. Number of carbons in right column Wikipedia

Fatty acids with more than one double bond are called polyunsaturated. Plants are excellent sources of unsaturated and polyunsaturated fatty acids. The position of the double bond(s) in fatty acids has important considerations both for their synthesis and for their actions in the body. Biochemically, the double bonds found in fatty acids are predominantly in the cis configuration. So-called trans fats arise as a chemical by-product of partial hydrogenation of vegetable oil.

Name	Double Bond Info	
Myristoleic acid	*cis*-Δ^9	14:1
Palmitoleic acid	*cis*-Δ^9	16:1
Sapienic acid	*cis*-Δ^6	16:1
Oleic acid	*cis*-Δ^9	18:1
Elaidic acid	*trans*-Δ^9	18:1
Vaccenic acid	*trans*-Δ^{11}	18:1
Linoleic acid	*cis,cis*-Δ^9,Δ^{12}	18:2
Linoelaidic acid	*trans,trans*-Δ^9,Δ^{12}	18:2
α-Linolenic acid	*cis,cis,cis*-$\Delta^9,\Delta^{12},\Delta^{15}$	18:3
Arachidonic acid	*cis,cis,cis,cis*-$\Delta^5\Delta^8,\Delta^{11},\Delta^{14}$	20:4
Eicosapentaenoic acid	*cis,cis,cis,cis,cis*-$\Delta^5,\Delta^8,\Delta^{11},\Delta^{14},\Delta^{17}$	20:5
Erucic acid	*cis*-Δ^{13}	22:1
Docosahexaenoic acid	*cis,cis,cis,cis,cis,cis*-$\Delta^4,\Delta^7,\Delta^{10},\Delta^{13},\Delta^{16},\Delta^{19}$	22:6

Figure 2.193 - Unsaturated fatty acids. Right column Indicates number of carbons and double bonds Wikipedia

In humans, consumption of trans fat raises low density lipoprotein (LDL) levels and lowers high density lipoprotein (HDL) levels. Each is thought to contribute to the risk of developing coronary artery disease. The most Figure 2.194 - Fatty acid models. Carboxyl end labeled in red Wikipedia common fatty acids in our body include palmitate, stearate, oleate, linolenate, linoleate, and arachidonate. Two notable shorter fatty acids are nonanoic (9 carbons) and decanoic acid (10 carbons), both of which appear to have anti-seizure effects. Decanoic acid directly inhibits excitatory neurotransmission in the brain and may contribute to the anticonvulsant effect of the ketogenic diet.

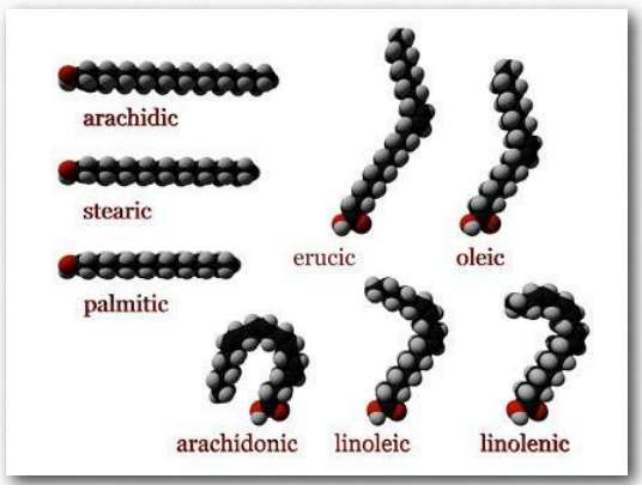


Figure 2.194 - Fatty acid models. Carboxyl end labeled in red Wikipedia

Numbering

Figure 2.195 shows two different systems for locating double bonds in a fatty acid. The ω system counts carbons starting with the methyl end (shown in red) while the Δ system counts from the carboxyl end (shown in blue). For example, an ω-3 (omega 3) fatty acid would have a double bond at the third carbon from the methyl end. In the Δ system, a fatty acid that has a cis double bond at carbon 6, counting from the carboxyl end, would be written as cis-Δ6.

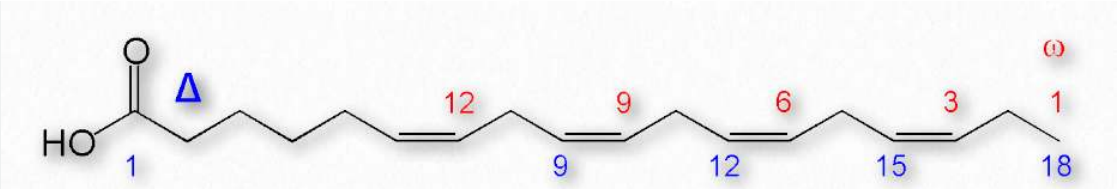


Figure 2.195 - Δ and ω numbering systems for fatty acids Image by Pehr Jacobson

Fatty acids are described as essential fatty acids if they must be in the diet (can't be synthesized by the organism) and nonessential fatty acids if the organism can synthesize them. Humans and other animals lack the desaturase enzymes necessary to make double bonds at positions greater than Δ-9, so fatty acids with double bonds beyond this position must be obtained in the diet. Linoleic acid and linolenic acid, both fall in this category. Related unsaturated fatty acids can be made from these fatty acids, so the presence of linoleic and linolenic acids in the diet eliminates the need to have all unsaturated fatty acids in the diet. Both linoleic and linolenic acid contain 18 carbons, but linoleic acid is an ω-6 fatty acid, whereas linolenic acid is an ω-3 fatty acid. Notably, ω-6 fatty acids tend to be proinflammatory, whereas ω-3 fatty acids are lesser so.

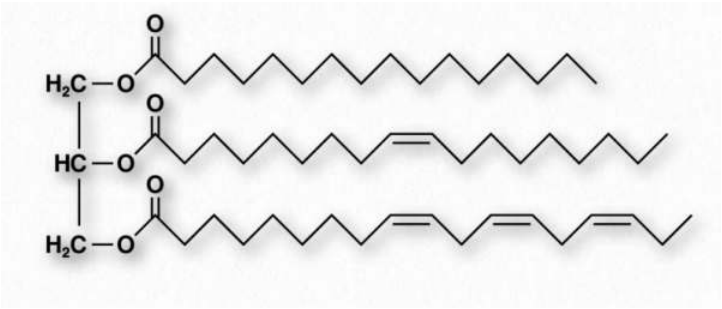


Figure 2.196 - Structure of a fat/oil

Fats/oils

Fats and oils are the primary energy storage forms of animals and are also known as triacylglycerols and triglycerides, since they consist of a glycerol molecule linked via ester bonds to three fatty acids (Figure 2.196). Fats and oils have the same basic structure. We give the name fat to those compounds that are solid at room temperature and the name oil to those that are liquid at room temperature. Note that biological oils are not the same as petroleum oils.

Increasing the number of unsaturated fatty acids (and the amount of unsaturation in a given fatty acid) in a fat decreases the melting temperature of it. Organisms like fish, which live in cool environments, have fats with more unsaturation and this is why fish oil contains polyunsaturated fatty acids.

Adipocytes

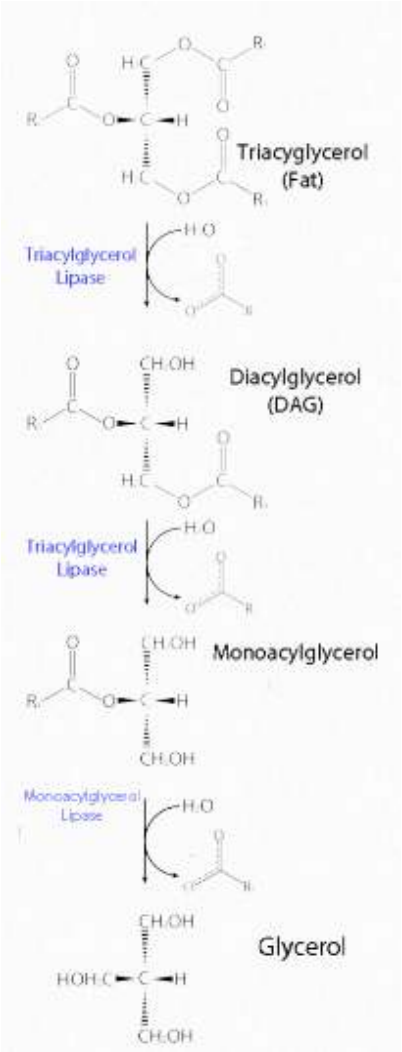


Figure 2.197 - Lipase action on a fat Image by Aleia Kim

Fats are stored in the body in specialized cells known as adipocytes. Enzymes known as lipases release fatty acids from fats by hydrolysis reactions (Figure 2.197). Triacylglycercol lipase (pancreatic - Figure 2.198) is able to cleave the first two fatty acids from the fat. A second enzyme, monoacylglycerol lipase, cleaves the last fatty acid. Fats can be synthesized by replacing the phosphate on phosphatidic acid with a fatty acid.

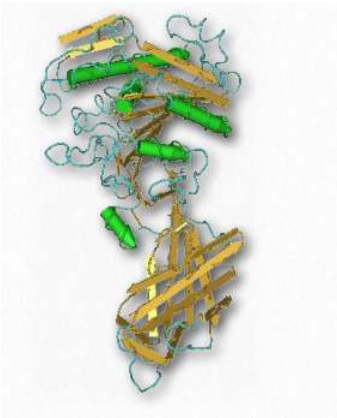


Figure 2.198 - Pancreatic lipase

Glycerophospholipids

Glycerophospholipids (phosphoglycerides) are important components of the lipid bilayer of cellular membranes. Phosphoglycerides are structurally related to fats, as both are derived from phosphatidic acid (Figure 2.199). Phosphatidic acid is a simple glycerophospholipid that is usually converted into phosphatidyl compounds. These are made by esterifying various groups, such as ethanolamine, serine, choline, inositol, and others (Figure 2.200) to the phosphate of phosphatidic acid. All of these compounds form lipid bilayers in aqueous solution , due to their amphiphilic nature.

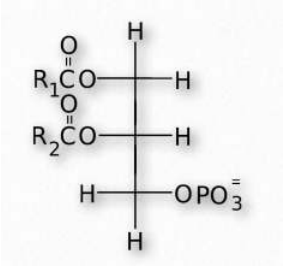


Figure 2.199 - Structure of phosphatidic acid. R1 and R2 are alkyl groups of fatty acids.

Phosphatidylethanolamines

Since all glycerolipids can have a variety of fatty acids at positions 1 and 2 on the glycerol, they all are families of compounds. The phosphatidylethanolamines are found in all living cells and are one of the most common phosphatides, making up about 25% of them. They are common constituents of brain tissue and in the spinal cord, making up as much as 45% of the total phospholipids. Phosphatidylethanolamines are asymmetrically distributed across membranes, being preferentially located on the inner leaflet (closest to the cytoplasm) of the plasma membrane. Metabolically, phosphatidylethanloamines are precursors of phosphatidylcholines. Phosphatidylserines Phosphatidylserines are another group of phosphatidyl compounds that are preferentially distributed across the lipid bilayer of the plasma membrane. Like the phosphatidylethanolamines, phosphatidylserines are preferentially located on the inner leaflet of the plasma membrane. When apoptosis (cell suicide) occurs, the preferential distribution is lost and the phosphatidylserines appear on the outer leaflet where they serve as a signal to macrophages to bind and destroy the cell.

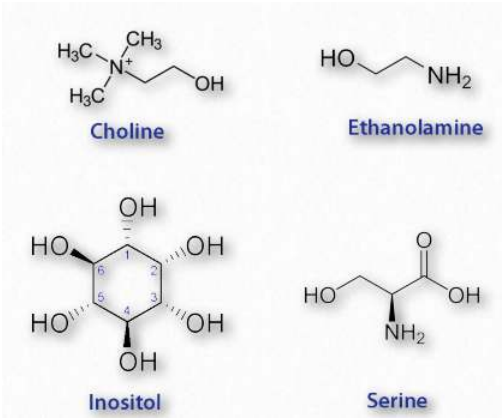


Figure 2.200 - Four common components of phosphatides Wikipedia

Phosphatidylcholines

Phosphatidylcholines (Figure 2.201) are another group of important membrane components. They tend to be found more commonly on the outer leaflet of the plasma membrane. Nutritionally, the compounds are readily obtained from eggs and soybeans. Phosphatidylcholines are moved across membranes by Phosphatidylcholine transfer protein (PCTP). This protein, which is sensitive to the levels of phosphatidylcholines, acts to stimulate the activity of a thioesterase (breaks thioester bonds, such as acyl-CoAs) and activates PAX3 transcription factors.

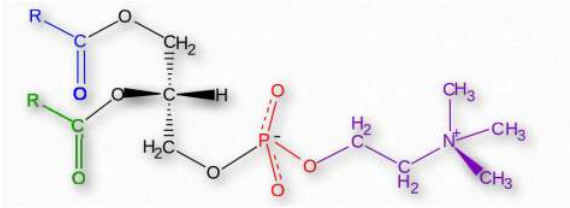


Figure 2.201 - Phosphatidylcholine

Cardiolipins

Cardiolipins are an unusual set of glycerophospholipids in containing two diacylglycerol backbones joined in the middle by a diphosphoglycerol (Figure 2.202). It is an important membrane lipid, constituting about 20% of the inner mitochondrial membrane and is found in organisms from bacteria to humans. In both plants and animals, it is found almost totally in the inner mitochondrial membrane.

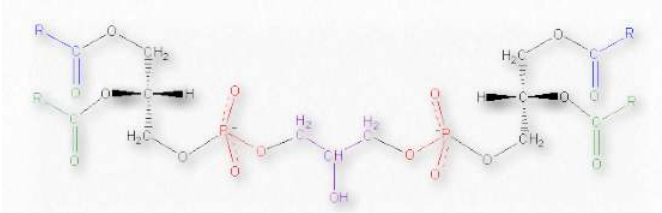


Figure 2.202 - Cardiolipin

The molecules appear to be required for both Complex IV and Complex III of the electron transport chain to maintain its structure. The ATP synthase enzyme (Complex V) of the oxidative phosphorylation system also binds four molecules of cardiolipin. It has been proposed that cardiolipin functions as a proton trap in the process of proton pumping by Complex IV.

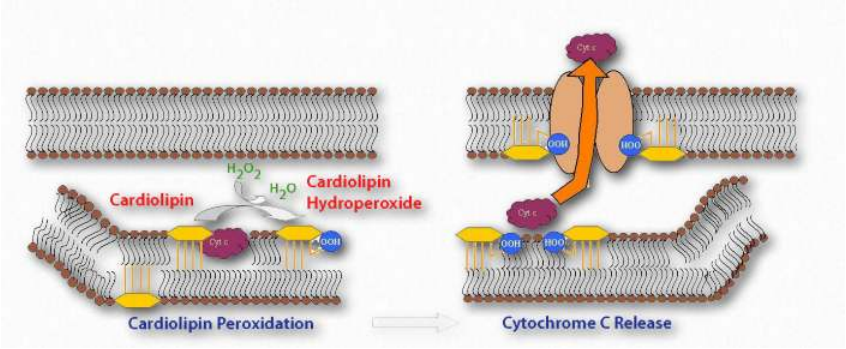


Figure 2.203 - Cardiolipin oxidation and the release of cytochrome C in apoptosis

Cardiolipin also plays a role in apoptosis. As shown in Figure 2.203, oxidation of cardiolipin by a cardiolipin-specific oxygenase causes cardiolipin to move from the inner mitochondrial membrane to the outer one, helping to form a permeable pore and facilitating the transport of cytochrome c out of the intermembrane space and into the cytoplasm - a step in the process of apoptosis.

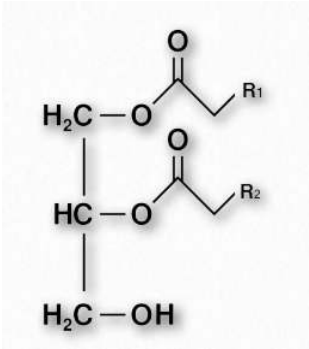


Figure 2.204 - Diacylglycerol

Diacylglycerol

Diacylglycerol (also called diglyceride and DAG - Figure 2.204) is an important intermediate in metabolic pathways. It is produced, for example, in the first step of the hydrolysis of fat and is also produced when membrane lipids, such as PIP2 (phosphatidylinositol-4,5-bisphosphate) are hydrolyzed by phospholipase C in a signaling cascade.

DAG is itself a signaling compound, binding to protein kinase C to activate it to phosphorylate substrates. Synthesis of DAG begins with glycerol-3-phosphate, which gains two fatty acids from two acyl-CoAs to form phosphatidic acid. Dephosphorylation of phosphatidic acid produces DAG. DAG can also be rephosphorylated by DAG kinase to re-make phosphatidic acid or another fatty acid can be added to make fat.

Inositol

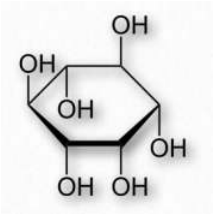


Figure 2.205 - Inositol

Though technically not a lipid itself, inositol is found in many lipids. Inositol is a derivative of cyclohexane containing six hydroxyl groups - one on each carbon (Figure 2.205. It has nine different stereoisomers of which one, cis-1,2,3,5-trans-4,6-cyclohexanehexol (called myo-inositol) is the most common. It has a sweet taste (half that of sucrose).

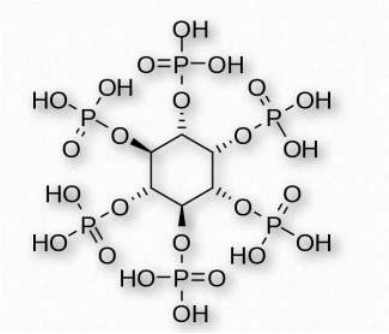


Figure 2.206 - Phytic acid

Numerous phosphorylated forms of the compound exist, from a single phosphate to six (one on each carbon). Phytic acid, for example, in plants, has six phosphates (Figure 2.206) that it uses to store phosphate. Inositol is produced from glucose and was once considered vitamin B8, but is made by the body in adequate amounts, so it is not now considered a vitamin. Phosphorylated forms of inositol are found in phosphoinositides, such as PIP2 and PIP3, both of which are important in signaling processes. Some of these include insulin signaling, fat catabolism, calcium regulation, and assembly of the cytoskeleton.

Phosphoinositides

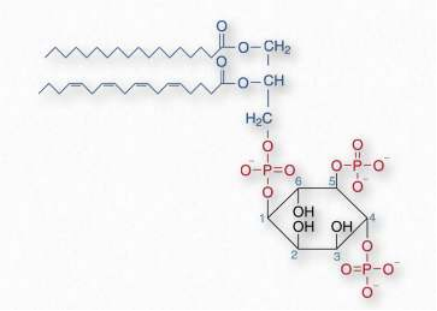


Figure 2.207 - Structure of PIP_2

Compounds based on phosphatidylinositol (PI) are often called phosphoinositides. These compounds have important roles in signaling and membrane trafficking. Hydroxyls on carbons 3,4, and 5 of the inositol ring are targets for phosphorylation by a variety of kinases. Seven different combinations are used. Steric hindrance inhibits phosphorylation of carbons 2 or 6. Naming of these phosphorylated compounds follows generally as PI(#P)P, PI(#P, #P)P, or PI(#P, #P, #P)P where #P refers to the number of the carbon where a phosphate is located. For example, PI(3)P refers to a phosphatidyl compound with a phosphate added to carbons 3 of the inositol ring, whereas PI(3,4,5)P is a phosphatidyl compound with a phosphate added to carbons 3,4,and 5.

Phosphatidylinositol-4,5- bisphosphate

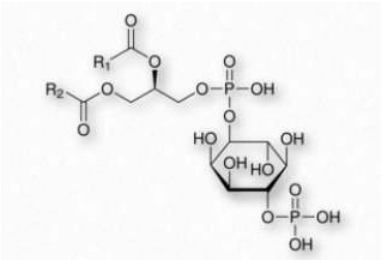


Figure 2.208 - Phosphatidylinositol-4- phosphate

Phosphatidylinositol-4,5-bisphosphate (PIP2 - Figure 2.207) is a phospholipid of plasma membranes that functions in the phospholipase C signaling cascade. In this signaling pathway, hydrolysis catalyzed by phospholipase C releases inositol-1,4,5-trisphosphate (IP3) and diacylglycerol. Synthesis of PIP2 begins with phosphatidylinositol, which is phosphorylated at position 4 followed by phosphorylation at position 5 by specific kinases.

PIP2 can be phosphorylated to form the signaling molecule known as phosphatidylinositol (3,4,5)-trisphosphate (PIP3). Along with PIP3, PIP2 serves as a docking phospholipid for the recruitment of proteins that play roles in signaling cascades. Binding of PIP2 is also required by inwardly directed potassium channels.

Phosphatidylinositol (3,4,5)- trisphosphate

Phosphatidylinositol (3,4,5)-trisphosphate (PIP3) is an important molecule for the activation of signaling proteins, such as AKT, which activates anabolic signaling pathways related to growth and survival. PIP3 can be dephosphorylated by phosphatase PTEN to yield PIP2 and can be synthesized from PIP2 by kinase action of Class I PI 3- kinases. Kinase activity to synthesize PIP3 results in movement of PIP3-binding proteins to the plasma membrane. They include Akt/ PKB, PDK1, Btk1, and ARNO and each is activated by binding to PIP3.

Plasmalogens

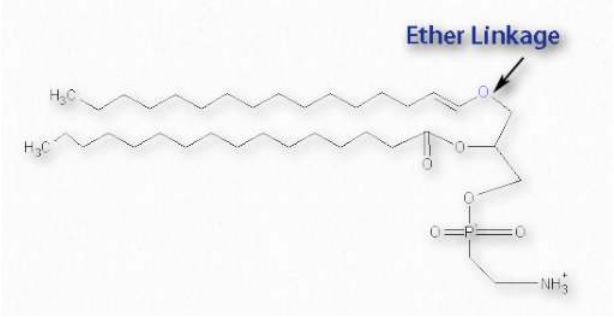


Figure 2.209 - Plasmalogen - A vinyl ether lipid Wikipedia

A special class of the glycerophospholipids are the plasmalogens (Figure 2.209). They differ in containing a vinyl ether linkage at position 1 of glycerol, in contrast to other glycerophopsholipids, which have an ester linkage at this position. Position 2 of each is an ester. The precursor for the ether linkage is typically a 16 or 18 carbon saturated alcohol or an 18 carbon unsaturated alcohol.

At the phosphate tail, the most commonly attached groups are ethanolamine or choline. Plasmalogens are found abundantly in humans in heart (30-40% of choline phospholipids). 30% of the glycerophospholipids in brain are plasmalogens and 70% of the ethanolamine lipids of the myelin sheath of nerve cells are plasmalogens.

Though their function is not understood, it is believed that plasmalogens may provide some protection against reactive oxygen species and have roles in signaling.

Lecithin

Lecithin is a generic term for a combination of lipid substances that includes phosphoric acid, glycerol, glycolipids, triglycerides, and phospholipids. Lecithin is a wetting agent helpful with emulsification and encapsulation and is even used as an anti-sludge additive in motor lubricants. Lecithin is used in candy bars to keep cocoa and cocoa butter from separating. Though considered safe as a food ingredient, lecithin can be converted by gut bacteria to trimethylamine-N-oxide which may contribute to cholesterol deposition and atherosclerosis.

Sphingolipids

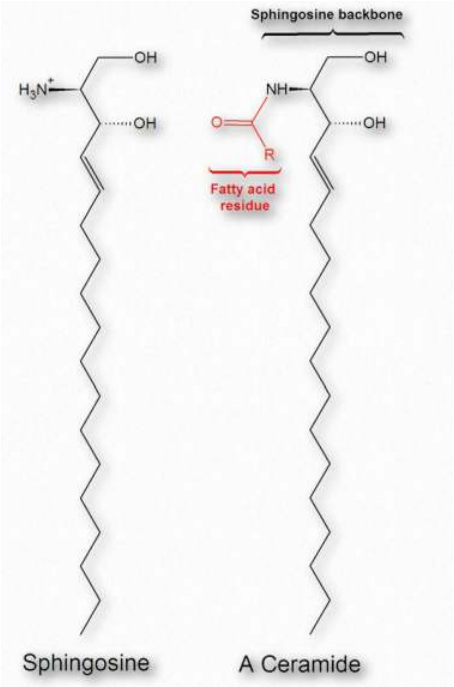


Figure 2.210 - Sphingosine and a ceramide made from it Wikipedia

Fatty acids are also components of a broad class of molecules called sphingolipids. Sphingolipids are structurally similar to glycerophospholipids, though they are synthesized completely independently of them starting with palmitic acid and the amino acid serine. Sphingolipids are named for the amino alcohol known as sphingosine (Figure 2.210), though they are not directly synthesized from it. Figure 2.211 shows the generalized structure of sphingolipids.

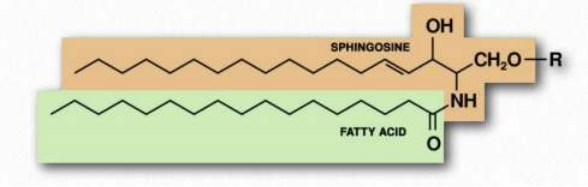


Figure 2.211 Schematic structure of a sphingolipid

If the R-group is a hydrogen, the molecule is called a ceramide. When the R-group is phosphoethanolamine the resulting molecule is sphingomyelin, an important component of the myelin sheath and lipid membranes. If a single, simple sugar is instead added, a cerebroside is created (Figure 2.212). Addition of a complex oligosaccharide creates a ganglioside.

Complex sphingolipids may play roles in cellular recognition and signaling. Sphingolipids are found most abundantly in plasma membrane and are almost completely absent from mitochondrial and endoplasmic reticulum membranes. In animals, dietary sphingolipids have been linked to reduced colon cancer, reductions in LDLs, and increases in HDLs. Like the glycerophospholipids, sphingolipids are amphiphilic. Most sphingolipids except sphingomyelin do not contain phosphate.

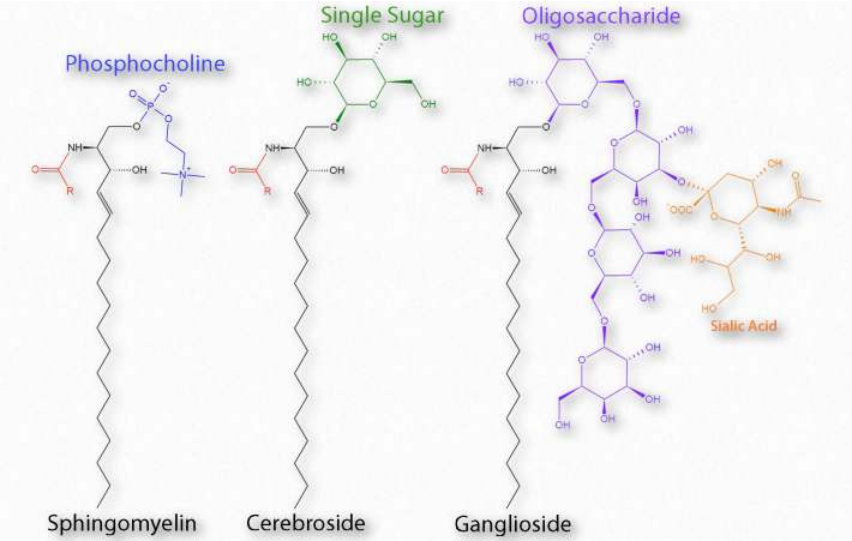


Figure 2.212 - Categories of Sphingolipid Wikipedia

Eicosanoids

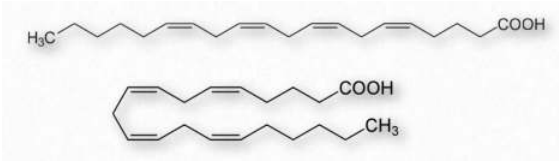


Figure 2.213 - Arachidonic acid drawn as straight (top) and bent (bottom)

Fatty acids made from omega-6 and omega-3 fatty acids include three important fatty acids containing 20 carbons. They include arachidonic acid (an ω-6 fatty acid with four double bonds (Δ-5,8,11,14) - Figure 2.213), eicosapentaenoic acid (an ω-3 fatty acid with five double bonds, and dihomo-γ-linolenic acid (an ω-6 fatty acid with three double bonds). The class of compounds known as eicosanoids is made by oxidation of these compounds. Subclasses include include prostaglandins, prostacyclins, thromboxanes, lipoxins, leukotrienes, and endocannabinoids (Figures 2.214-2.219). Eicosanoids play important roles affecting inflammation, immunity, mood, and behavior.

Prostaglandins

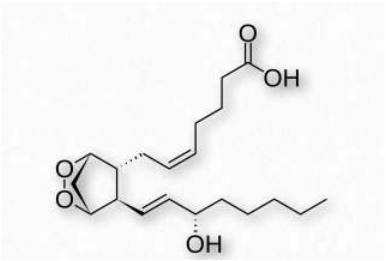


Figure 2.214 - Prostaglandin PGH_2

A collection of molecules acting like hormones, prostaglandins are derived from arachidonic acid and have many differing (even conflicting) physiological effects. These include constriction or dilation of vascular smooth muscle cells, induction of labor, regulation of inflammation, and action on the thermoregulatory center of the hypothalamus to induce fever, among others.

Prostaglandins are grouped with the thromboxanes (below) and prostacyclins (below), as prostanoids. The prostanoids, which all contain 20 carbons are a subclass of the eicosanoids. Prostaglandins are found in most tissues of higher organisms. They are autocrine or paracrine compounds produced from essential fatty acids. The primary precursor of the prostaglandins is the fatty acid known as arachidonic acid and the prostaglandin made from it is known as PGH2 (Figure 2.214), which, in turn is a precursor of other prostaglandins, as well as the prostacyclins and thromboxanes.

Interesting prostaglandins

PGD_2 - inhibits hair follicle growth, vasodilator, causes bronchial constriction, higher in lungs of asthmatics than others.

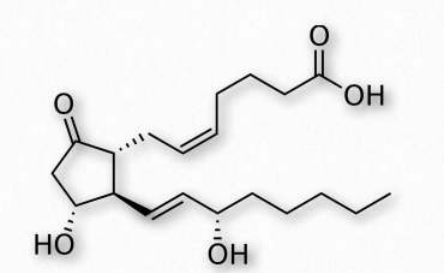


Figure 2.215 Prostaglandin E

PGE_2 (Figure 2.215) - exerts effects in labor (soften cervix, uterine contraction), stimulates bone resorption by osteoclasts, induces fever, suppresses T-cell receptor signaling, vasodilator, inhibits release of noradrenalin from sympathetic nerve terminals. It is a potent activator of the Wnt signaling pathway.

A prostaglandin can have opposite effects, depending on which receptor it binds to. Binding of PGE2 to the EP1 receptor causes bronchoconstriction and smooth muscle contraction, whereas binding of the same molecule to the EP2 receptor causes bronchodilation and smooth muscle relaxation.

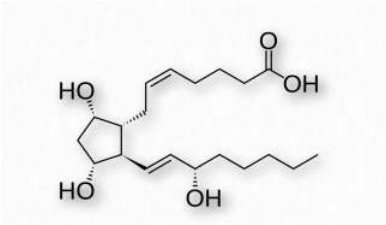


Figure 2.216 - Prostaglandin $F_{2\alpha}$

$PGF_{2\alpha}$ (Figure 2.216)- uterine contractions, induces labor, bronchoconstriction.

PGI_2 - vasodilation, bronchodilation, inhibition of platelet aggregation.

Thromboxanes

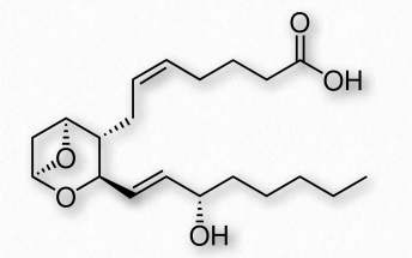


Figure 2.217 Thromboxane $A2_2$

Thromboxanes play roles in clot formation and named for their role in thrombosis. They are potent vasoconstrictors and facilitate platelet aggregation. They are synthesized in platelets, as well. The anti-clotting effects of aspirin have their roots in the inhibition of synthesis of PGH2, which is the precursor of the thromboxanes. The most common thromboxanes are A2 (Figure 2.217) and B2.

Prostacyclin

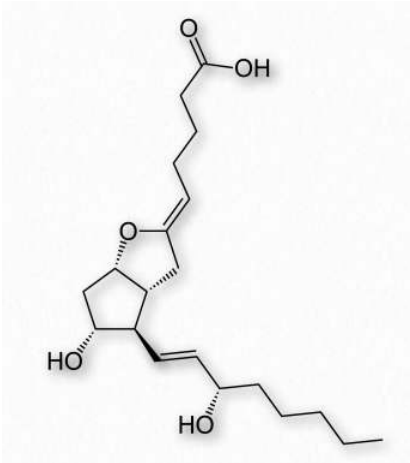


Figure 2.218 - Prostacyclin

Prostacyclin (also known as prostaglandin I2 or PGI2 - Figure 2.218) counters the effects of thromboxanes, inhibiting platelet activation and acting as vasodilators. It is produced from PGH2 by action of the enzyme prostacyclin synthase.

Leukotrienes

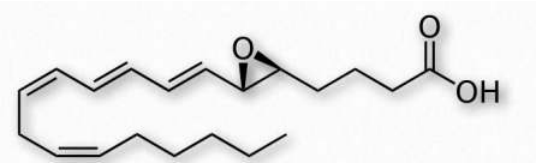


Figure 2.219 - Leukotriene A_4 (LTA_4)

Another group of eicosanoid compounds are the leukotrienes (Figure 2.219). Like prostaglandins, leukotrienes are made from arachidonic acid. The enzyme catalyzing their formation is a dioxygenase known as arachidonate 5-lipoxygenase. Leukotrienes are involved in regulating immune responses. They are found in leukocytes and other immunocompetent cells, such as neutrophils, monocytes, mast cells, eosinophils, and basophils. Leukotrienes are associated with production of histamines and prostaglandins, which act as mediators of inflammation. Leukotrienes also trigger contractions in the smooth muscles of the bronchioles. When overproduced, they may pay a role in asthma and allergic reactions. Some treatments for asthma aim at inhibiting production or action of leukotrienes.

Cholesterol

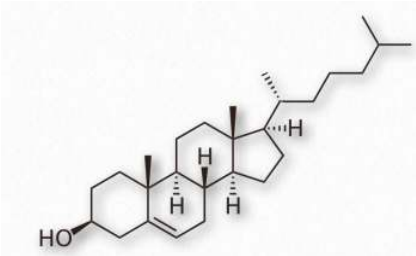


Figure 2.220 - Structure of cholesterol

Arguably, no single biomolecule has generated as much discussion and interest as has cholesterol (Figure 2.220). Certainly, from the perspective of the Nobel Prize committee, no small molecule even comes close, with 13 people having been awarded prizes for work on it. Evidence for cholesterol's importance comes from the study of brain tissue where it comprises 10-15% of the dry mass.

Membrane flexibility

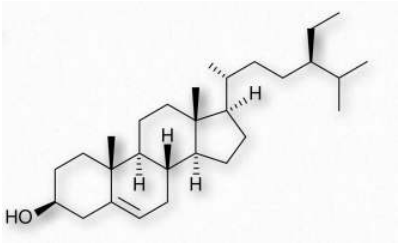


Figure 2.221 - Sitosterol - A phytosterol

In animal cells, cholesterol provides for membrane flexibility that allows for cellular movement that is in contrast to plant and bacterial cells with fixed structures. Cholesterol is made in many cells of the body, with the liver making the greatest amount. The anabolic pathway leading to synthesis of cholesterol is known as the isoprenoid pathway and branches of it lead to other molecules including other fat-soluble vitamins.

Figure 2.222 - Margarine - A common source of trans fat Wikipedia

Cholesterol is only rarely found in prokaryotes (Mycoplasma, which requires it for growth, is an exception) and is found in only trace amounts in plants. Instead, plants synthesize similar compounds called phytosterols (Figure 2.221). On average, the body of a 150 pound adult male makes about 1 gram of cholesterol per day, with a total content of about 35 grams.

Packaging

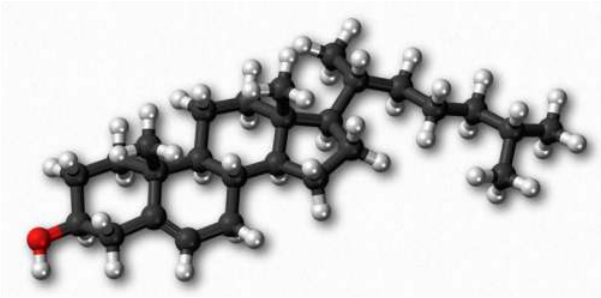


Figure 2.223 - Cholesterol - Ball and stick model

Cholesterol's (and other lipids') hydrophobicity requires special packaging into lipoprotein complexes (called chylomicrons, VLDLs, IDLs, LDLs, and HDLs) for movement in the lymph system and bloodstream. Though cholesterol can be made by cells, they also take it up from the blood supply by absorbing cholesterol-containing LDLs directly in a process called receptor-mediated endocytosis.

Oxidative damage to LDLs can lead to formation of atherosclerotic plaques and this is why cholesterol has gotten such a negative image in the public eye. The liver excretes cholesterol through the bile for elimination into the digestive system, but the compound is recycled there.

Reducing cholesterol levels

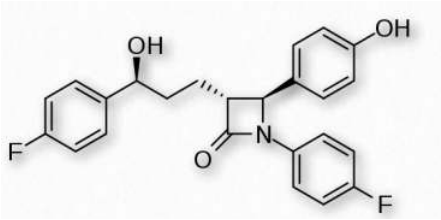


Figure 2.224 - Ezetimibe - An inhibitor of cholesterol absorption

Strategies for reducing cholesterol in the body focus primarily on three areas - reducing consumption, reducing endogenous synthesis, and reducing the recycling. Dietary considerations, such as saturated fat versus unsaturated fat consumption are currently debated. Dietary trans fats, though, correlate with incidence of coronary heart disease. Consumption of vegetables may provide some assistance with reducing levels of cholesterol recycled in the digestive system, because plant phytosterols compete with cholesterol for reabsorption and when this happens, a greater percentage of cholesterol exits the body in the feces. Drugs related to penicillin are also used to inhibit cholesterol recycling. One of these is ezetimibe, shown in Figure 2.224.

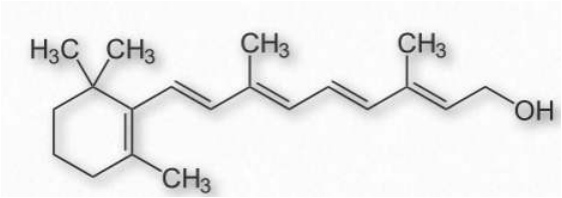


Figure 2.225 - All-trans retinol

Genetic defects in the cholesterol movement system are a cause of the rare disease known as familial hypercholesterolemia in which the blood of afflicted individuals contains dangerously high levels of LDLs. Left untreated, the disease is often fatal in the first 10-20 years of life. While LDLs have received (and deserve) a bad rap, another group of lipoprotein complexes known as the HDLs (high density lipoprotein complexes) are known as “good cholesterol” because their levels correlate with removal of debris (including cholesterol) from arteries and reduce inflammation.

Membrane function

In membranes, cholesterol is important as an insulator for the transmission of signals in nerve tissue and it helps to manage fluidity of membranes over a wide range of temperatures. Stacked in the lipid bilayer, cholesterol decreases a membrane’s fluidity and its permeability to neutral compounds, as well as protons and sodium ions. Cholesterol may play a role in signaling by helping with construction of lipid rafts within the cell membrane.

Vitamin A

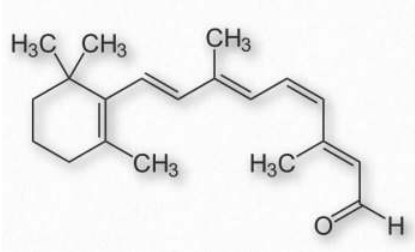


Figure 2.226 - 11-cis retinal

Vitamin A comes in three primary chemical forms, retinol (storage in liver - Figure 2.225), retinal (role in vision - Figure 2.226), and retinoic acid (roles in growth and development). All vitamin A forms are diterpenoids and differ only in the chemical form of the terminal group. Retinol is mostly used as the storage form of the vitamin.

Retinol is commonly esterified to a fatty acid and kept in the liver. In high levels, the compound is toxic. Retinol is obtained in the body by hydrolysis of the ester or by reduction of retinal. Importantly, neither retinal nor retinol can be made from retinoic acid. Retinoic acid is important for healthy skin and teeth, as well as bone growth. It acts in differentiation of stem cells through a specific cellular retinoic acid receptor.

Sources

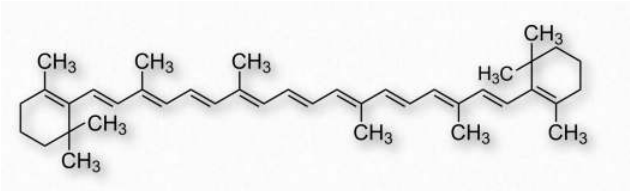


Figure 2.227 - β-Carotene

Good sources of vitamin A are liver and eggs, as well as many plants, including carrots, which have a precursor, β-carotene (Figure 2.227) from which retinol may be made by action of a dioxygenase.

Light sensitivity The conjugated double bond system in the side chain of vitamin A is sensitive to light and can flip between cis and trans forms on exposure to it. It is this response to light that makes it possible for retinal to have a role in vision in the rods and cones of the eyes. Here, the aldehyde form (retinal) is bound to the protein rhodopsin in the membranes of rod and cone cells.

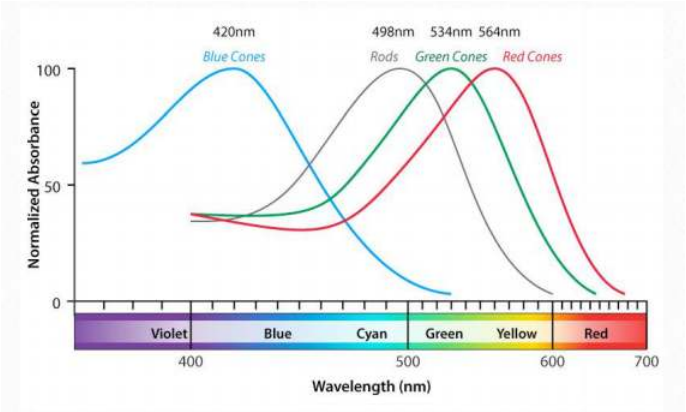


Figure 2.228 - Color sensitivity for cones and rods Image by Aleia Kim

When exposed to light of a particular wavelength, the “tail” of the retinal molecule will flip back and forth from cis to trans at the double bond at position 11 of the molecule. When this happens, a nerve signal is generated that signals the brain of exposure to light. Slightly different forms of rhodopsin have different maximum absorption maxima allowing the brain to perceive red, green and blue specifically and to assemble those into the images we see (Figure 2.228). Cones are the cells responsible for color vision, whereas rods are mostly involved in light detection in low light circumstances.

Deficiency and surplus

Deficiency of vitamin A is common in developing countries and was inspiration for the design and synthesis of the geneticallymodified golden rice, which is used as a source of vitamin A to help prevent blindness in children. Overdose of vitamin A, called hypervitaminosis A is dangerous and can be fatal. Excess vitamin A is also suspected to be linked to osteoporosis. In smokers, excess vitamin A is linked to an increased rate of lung cancer, but non-smokers have a reduced rate.

Vitamin D

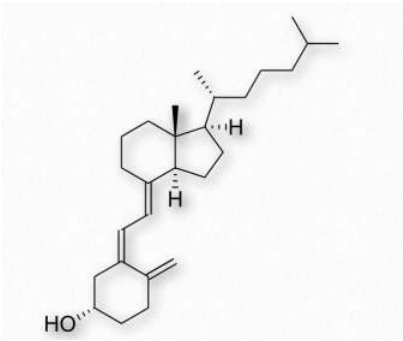


Figure 2.229 - Cholecalciferol - Vitamin D3

The active form of vitamin D plays important roles in the intestinal absorption of calcium and phosphate and thus in healthy bones. Technically, vitamin D isn't even a vitamin, as it is a compound made by the body. Rather, it behaves more like a hormone.

Derived from ultimately from cholesterol, vitamin D can be made in a reaction catalyzed by ultraviolet light. In the reaction, the intermediate 7-dehydrocholesterol is converted to cholecalciferol (vitamin D3) by the uv light (Figure 2.229). The reaction occurs most readily in the bottom two layers of the skin shown in Figure 2.230.

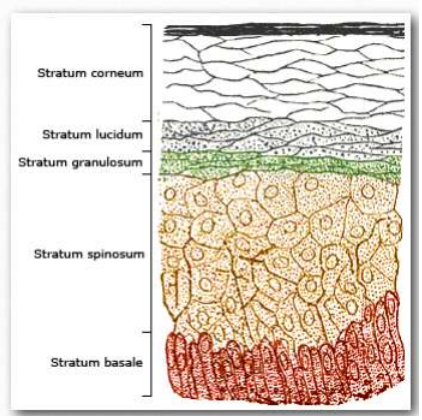


Figure 2.230 - Layers of the skin. Outside is at top.

Forms of vitamin D

Five different compounds are referred to as vitamin D. They are

Vitamin D1 - A mixture of ergocalciferol and lumisterol

Vitamin D2 - Ergocalciferol

Vitamin D3 - Cholecalciferol Vitamin

D4 - 22-Dihydroergocalciferol Vitamin

D5 - Sitocalciferol

Vitamin D3 is the most common form used in vitamin supplements and it and vitamin D2 are commonly obtained in the diet, as well. The active form of vitamin D, calcitriol (Figure 2.231), is made in the body in controlled amounts. This proceeds through two steps from cholecalciferol. First, a hydroxylation in the liver produces calcidiol and a second hydroxylation in the kidney produces calcitriol. Monocyte macrophages can also synthesize vitamin D and they use is as a cytokine to stimulate the innate immune system.

Mechanism of action

Calcitriol moves in the body bound to a vitamin D binding protein, which delivers it to target organs. Calcitriol inside of cells acts by binding a vitamin D receptor (VDR), which results in most of the vitamin's physiological effects. After binding calcitriol, the VDR migrates to the nucleus where it acts as a transcription factor to control levels of expression of calcium transport proteins (for example) in the intestine. Most tissues respond to VDR bound to calcitriol and the result is moderation of calcium and phosphate levels in cells.

Deficiency/excess

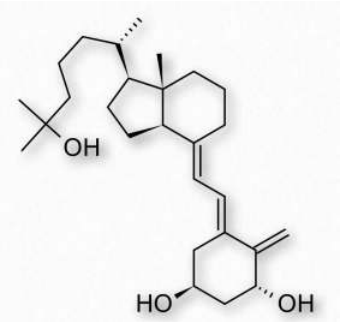


Figure 2.231 - Calcitriol - Active form of vitamin D

Deficiency of vitamin D is a cause of the disease known as rickets, which is characterized by soft, weak bones and most commonly is found in children. It is not common in the developed world, but elsewhere is of increasing concern.

Excess of vitamin D is rare, but has toxic effects, including hypercalcemia, which results in painful calcium deposits in major organs. Indications of vitamin D toxicity are increased urination and thirst. Vitamin D toxicity can lead to mental retardation and many other serious health problems.

Vitamin E

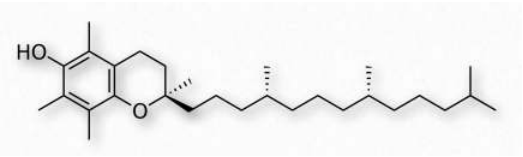


Figure 2.232 α-tocopherol - The most biologically active form of vitamin E Figure

Vitamin E comprises a group of two compounds (tocopherols and tocotrienols - Figure 2.232) and stereoisomers of each. It is commonly found in plant oils. The compounds act in cells as fat-soluble antioxidants. α-tocopherol (Figure 2.233), the most active form of the vitamin, works 1) through the glutathione peroxidase protective system and 2) in membranes to interrupt lipid peroxidation chain reactions. In both actions, vitamin E reduces levels of reactive oxygen species in cells.

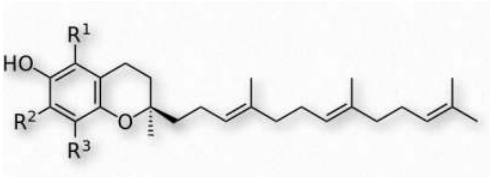


Figure 2.233 - Structure of tocotrienols

Action

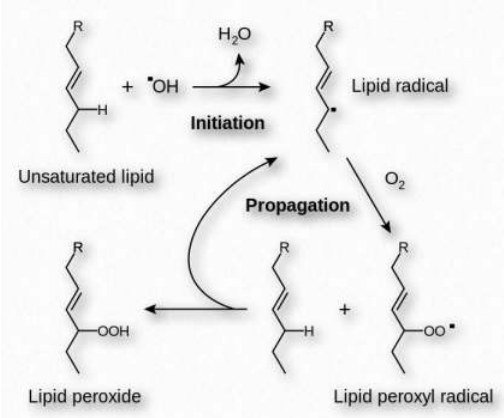


Figure 2.234 - Lipid peroxidation reactions

Vitamin E scavenges oxygen radicals (possessing unpaired electrons) by reacting with them to produce a tocopheryl radical. This vitamin E radical can be converted back to its original form by a hydrogen donor. Vitamin C is one such donor. Acting in this way,

Vitamin E helps reduce oxidation of easily oxidized compounds in the lipid peroxidation reactions (Figure 2.234).

Vitamin E also can affect enzyme activity. The compound can inhibit action of protein kinase C in smooth muscle and simultaneously activate catalysis of protein phosphatase 2A to remove phosphates, stopping smooth muscle growth.

Deficiency/excess

Deficiency of vitamin E can lead to poor conduction of nerve signals and other issues arising from nerve problems. Low levels of the vitamin may be a factor in low birth weights and premature deliveries. Deficiency, however, is rare, and not usually associated with diet.

Excess Vitamin E reduces vitamin K levels, thus reducing the ability to clot blood. Hypervitaminosis of vitamin E in conjunction with aspirin can be life threatening. At lower levels, vitamin E may serve a preventative role with respect to atherosclerosis by reducing oxidation of LDLs, a step in plaque formation.

Vitamin K

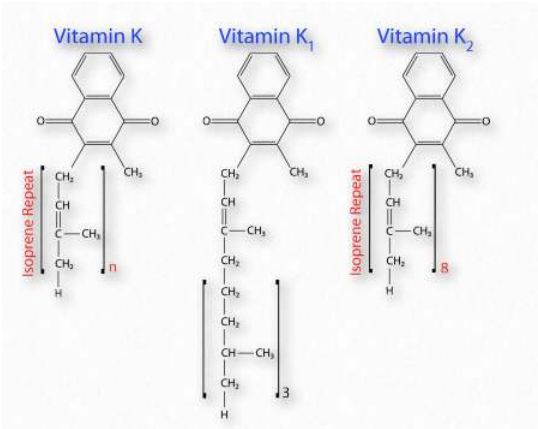


Figure 2.235 - Forms of vitamin K Image by Pehr Jacobson

Like the other fat-soluble vitamins, Vitamin K comes in multiple forms (Figure 2.235) and is stored in fat tissue in the body. There are two primary forms of the vitamin - K1 and K2 and the latter has multiple sub-forms . Vitamins K3, K4, and K5 are made synthetically, not biologically.

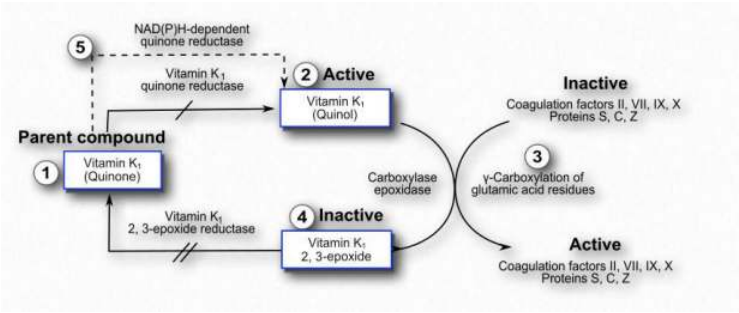


Figure 2.236 - Recycling of vitamin K Wikipedia

Action

Vitamin K is used as a co-factor for enzymes that add carboxyl groups to glutamate side chains of proteins to increase their affinity for calcium. Sixteen such proteins are known in humans. They include proteins involved in blood clotting (prothrombin (called Factor II), Factors VII, IX, and X), bone metabolism (osteocalcin, also called bone Gla protein (BGP), matrix Gla protein (MGP), and periostin) and others.

Modification of prothrombin is an important step in the process of blood clotting (see HERE). Reduced levels of vitamin K result in less blood clotting, a phenomenon sometimes referred to as blood thinning. Drugs that block recycling of vitamin K (Figure 2.236) by inhibiting the vitamin K epoxide reductase, produce lower levels of the vitamin and are employed in treatments for people prone to excessive clotting. Warfarin (coumadin) is one such compound that acts in this way and is used therapeutically.

Individuals respond to the drug differentially, requiring them to periodically be tested for levels of clotting they possess, lest too much or too little occur.

Sources

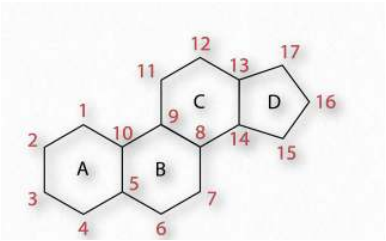


Figure 2.237 - Steroid numbering scheme Image by Pehr Jacobson

Vitamin K1 is a stereoisomer of the plant photosystem I electron receptor known as phylloquinone and is found abundantly in green, leafy vegetables. Phylloquinone is one source of vitamin K, but the compound binds tightly to thylakoid membranes and tends to have low bioavailability. Vitamin K2 is produced by microbes in the gut and is a primary source of the vitamin. Infants in the first few days before they establish their gut flora and people taking broad spectrum antibiotics may have reduced levels, as a result. Dietary deficiency is rare in the absence of damage to the small bowel. Others at risk of deficiency include people with chronic kidney disease and anyone suffering from a vitamin D deficiency. Deficiencies produce symptoms of easy bruising, heavy menstrual bleeding, anemia, and nosebleeds.

Steroids

Steroids, such as cholesterol are found in membranes and act as signaling hormones in traveling through the body.

Steroid hormones are all made from cholesterol and are grouped into five categories - mineralocorticoids (21 carbons), glucocorticoids (21 carbons), progestagens (21 carbons), androgens (19 carbons), and estrogens (18 carbons).

Mineralocorticoids

Mineralocorticoids are steroid hormones that influence water and electrolyte balances. Aldosterone (Figure 2.238) is the primary mineralocorticoid hormone, though other steroid hormones (including progesterone) have some functions like it. Aldosterone stimulates kidneys to reabsorb sodium, secrete potassium, and passively reabsorb water. These actions have the effect of increasing blood pressure and blood volume. Mineralocorticoids are produced by the zona glomerulosa of the cortex of the adrenal gland.

Glucocorticoids

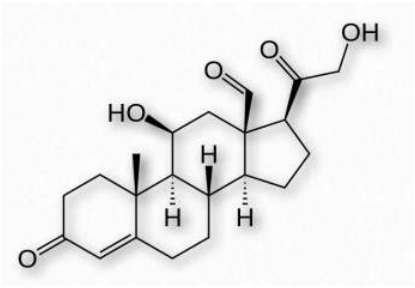


Figure 2.238 - Aldosterone - A mineralocorticoid

Glucocorticoids (GCs) bind to glucocorticoid receptors found in almost every vertebrate animal cell and act in a feedback mechanism in the immune system to reduce its activity. GCs are used to treat diseases associated with overactive immune systems. These include allergies, asthma, and autoimmune dis- Figure 2.237 - Steroid numbering scheme Image by Pehr Jacobson eases. Cortisol (Figure 2.239) is an important glucocorticoid with cardiovascular, metabolic, and immunologic functions. The synthetic glucocorticoid known as dexamethasone has medical applications for treating rheumatoid arthritis, bronchospasms (in asthma), and inflammation due to its increased potency (25-fold) compared to cortisol. Glucocorticoids are produced primarily in the zona fasciculata of the adrenal cortex.

Progestagens

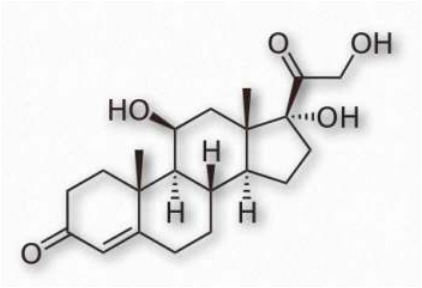


Figure 2.239 - Cortisol - A glucocorticoid

Progestagens (also called gestagens) are steroid hormones that work to activate the progesterone receptor upon binding to it. Synthetic progestagens are referred to as progestins. The most common progestagen is progesterone (also called P4 - Figure 2.240) and it has functions in maintaining pregnancy. Progesterone is produced primarily in the diestrus phase of the estrous cycle by the corpus luteum of mammalian ovaries. In pregnancy, the placenta takes over most progesterone production.

Androgens

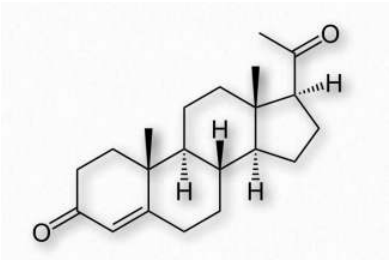


Figure 2.240 Progesterone - A progestagen

Androgens are steroid hormones that act by binding androgen receptors to stimulate development and maintenance of male characteristics in vertebrates. Androgens are precursors of estrogens (see below). The primary androgen is testosterone (Figure 2.241). Other important androgens include dihydrotestosterone (stimulates differentiation of penis, scrotum, and prostate in embryo) and androstenedione (common precursor of male and female hormones).

Estrogens

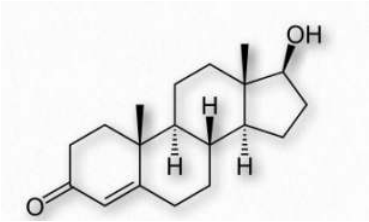


Figure 2.241 - Testosterone - An androgen

The estrogen steroid hormones are a class of compounds with important roles in menstrual and estrous cycles. They are the most important female sex hormones. Estrogens act by activating estrogen receptors inside of cells. These receptors, in turn, affect expression of many genes. The major estrogens in women include estrone (E1), estradiol (E2 - Figure 2.242), and estriol (E3). In the reproductive years, estradiol predominates. During pregnancy, estriol predominates and during menopause, estrone is the major estrogen.

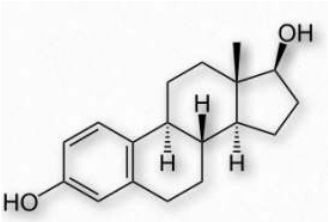


Figure 2.242 - Estradiol - An estrogen

Estrogens are made from the androgen hormones testosterone and androstenedione in a reaction catalyzed by the enzyme known as aromatase. Inhibition of this enzyme with aromatase inhibitors, such as exemestane, is a strategy for stopping estrogen production. This may be part of a chemotherapeutic treatment when estrogenresponsive tumors are present.

Cannabinoids

Cannabinoids are a group of chemicals that bind to and have effects on brain receptors (cannabinoid receptors), repressing neurotransmitter release. Classes of these compounds include endocannabinoids (made in the body), phytocannabinoids (made in plants, such as marijuana), and synthetic cannabinoids (man-made).

Endocannabinoids are natural molecules derived from arachidonic acid. Cannabinoid receptors are very abundant, comprising the largest number of G-protein- 247 Figure 2.243 - Tetrahydrocannabinol - Active ingredient in marijuana coupled receptors found in brain. The best known phytocannabinoid is Δ-9- tetrahydrocannabinol (THC), the primary psychoactive ingredient (of the 85 cannabinoids) of marijuana (Figure 2.243).

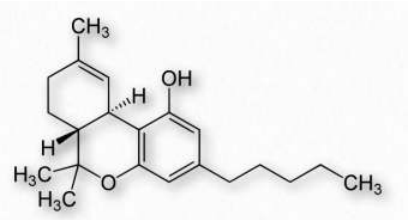


Figure 2.243 - Tetrahydrocannabinol - Active ingredient in marijuana

Anandamide

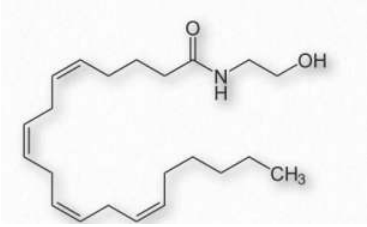


Figure 2.244 - Anandamide - An endocannabinoid

Anandamide (N-arachidonoylethanolamine - Figure 2.244) is an endocannabinoid neurotransmitter derived from arachidonic acid. It exerts its actions primarily through the CB1 and CB2 cannabinoid receptors, the same ones bound by the active ingredient in marijuana, Δ9-tetrahydrocannabinol. Anandamide has roles in stimulating eating/appetite and affecting motivation and pleasure. It has been proposed to play a role in “runners high,” an analgesic effect experienced from exertion, especially among runners. Anandamide appears to impair memory function in rats.

Anandamide has been found in chocolate and two compounds that mimic its effects (N-oleoylethanolamine and Nlinoleoylethanolamine) are present as well. The enzyme fatty acid amide hydrolase (FAAH) breaks down anandamide into free arachidonic acid and ethanolamine.

Lipoxins

Lipoxins (Figure 2.245) are eicosanoid compounds involved in modulating immune responses and they have anti-inflammatory effects. When lipoxins appear in inflammation it begins the end of the process. Lipoxins act to attract macrophages to apoptotic cells at the site of inflammation and they are engulfed. Lipoxins further act to start the resolution phase of the inflammation process.

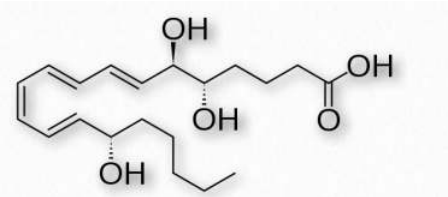


Figure 2.245 - Lipoxin A4

At least one lipoxin (aspirin-triggered LX4) has its synthesis stimulated by aspirin. This occurs as a byproduct of aspirin's acetylation of COX-2. When this occurs, the enzyme's catalytic activity is re-directed to synthesis of 15R-hydroxyeicosatetraenoic acid (HETE) instead of prostaglandins. 15R-HETE is a procursor of 15-epimer lipoxins, including aspirin-triggered LX4.

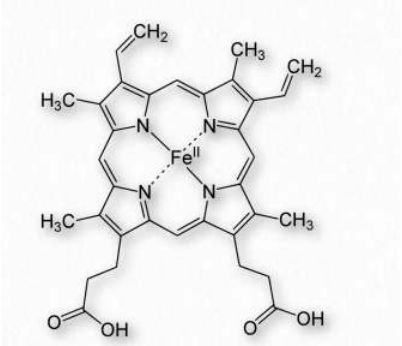


Figure 2.246 - Structure of heme B

Heme

Heme groups are a collection of protein/ enzyme cofactors containing a large heterocyclic aromatic ring known as a porphyrin ring with a ferrous (Fe++) ion in the middle. An example porphyrin ring with an iron (found in Heme B of hemoglobin), is shown in Figure 2.246. When contained in a protein, these are known collectively as hemoproteins (Figure 2.247).

Heme, of course, is a primary component of hemoglobin, but it is also found in other proteins, such as myoglobin, cytochromes, and the enzymes catalase and succinate dehydrogenase. Hemoproteins function in oxygen transport, catalysis, and electron transport. Heme is synthesized in the liver and bone marrow in a pathway that is conserved across a wide range of biology.

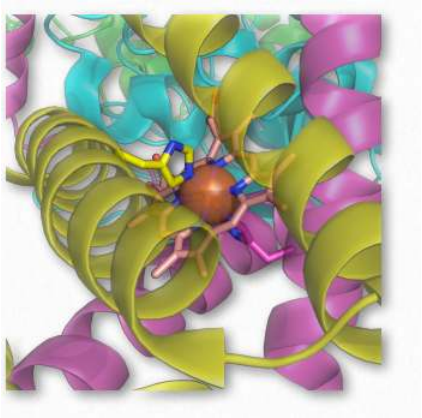


Figure 2.247 - Heme embedded in the succinate dehydrogenase hemoprotein Wikipedia

Porphobilinogen

Porphobilinogen (Figure 2.248) is a pyrrole molecule involved in porphyrin metabolism. It is produced from aminolevulinate by action of the enzyme known as ALA dehydratase. Porphobilinogen is acted upon by the enzyme porphobilinogen deaminase. Deficiency of the latter enzyme (and others in porphyrin metabolism) can result in a condition known as porphyria, which results in accumulation of porphobilinogen in the cytoplasm of cells.

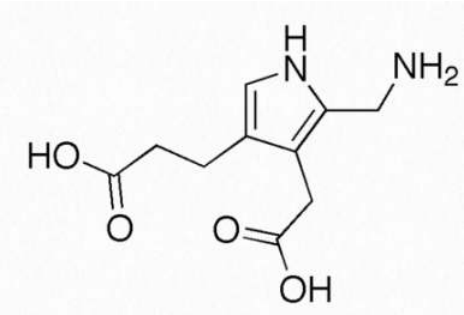


Figure 2.248 - Porphobilinogen

The disease can manifest itself with acute abdominal pain and numerous psychiatric issues. Both Vincent van Gogh and King ` George III are suspected to have suffered from porphyria, perhaps causing the “madness of King George III.” Porphyria is also considered by some to be the impetus for the legend of vampires seeking blood from victims, since the color of the skin in non-acute forms of the disease can be miscolored, leading some to perceive that as a deficiency of hemoglobin and hence the “thirst” for blood imagined for vampires.

Dolichols

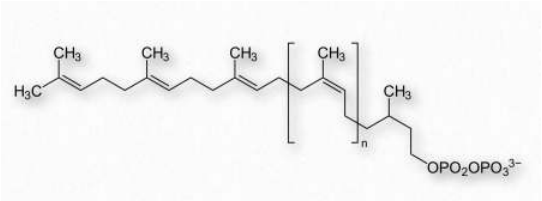


Figure 2.249 - Structure of dolichol pyrophosphate

Dolichol is a name for a group of non-polar molecules made by combining isoprene units together. Phosphorylated forms of dolichols play central roles in the N-glycosylation of proteins. This process, which occurs in the endoplasmic reticulum of eukaryotic cells, begins with a membrane-embedded dolichol pyrophosphate (Figure 2.249) to which an oligosaccharide is attached (also see HERE). This oligosaccharide contains three molecules of glucose, nine molecules of mannose and two molecules of N-acetylglucosamine.

Interestingly, the sugars are attached to the dolichol pyrophosphate with the pyrophosphate pointing outwards (away from) the endoplasmic reticulum, but after attachment, the dolichol complex flips so that the sugar portion is situated on the inside of the endoplasmic reticulum. There, the entire sugar complex is transferred to the amide of an asparagine side chain of a target protein.

The only asparagine side chains to which the attachment can be made are in proteins where the sequences Asn-X-Ser or Asn-X-Thr occur. Sugars can be removed/added after the transfer to the protein. Levels of dolichol in the human brain increase with age, but in neurodegenerative diseases, they decrease.

Terpenes

Figure 2.250 - Pine tree resin - A source of terpenes Wikipedia

Terpenes are members of a class of nonpolar molecules made from isoprene units. Terpenes are produced primarily by plants and by some insects. Terpenoids are a related group of molecules that contain functional groups lacking in terpenes.

Terpenes have a variety of functions. In plants, they often play a defensive role protecting from insects. The name of terpene comes from turpentine, which has an odor like some of the terpenes. Terpenes are common components of plant resins (think pine) and they are widely used in medicines and as fragrances. Hops, for example, gain some of their distinctive aroma and flavor from terpenes. Not all terpenes, however have significant odor.

Synthesis

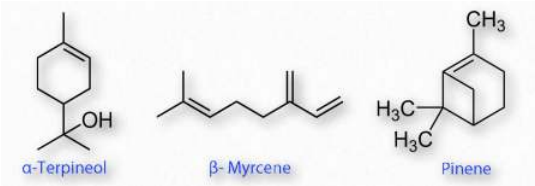


Figure 2.251 - Three monoterpenes

Terpenes, like steroids, are synthesized starting with simple building blocks known as isoprenes. There are two of them - dimethylallyl pyrophosphate and the related isopentenyl pyrophosphate and (Figures 2.252 and 2.253) which combine 1-2 units at a time to make higher order structures. Terpene synthesis overlaps and includes steroid synthesis.

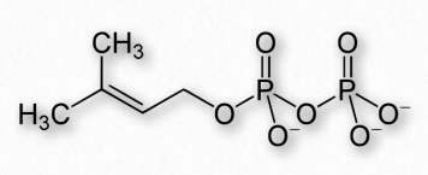


Figure 2.252 - Dimethylallyl pyrophosphate

Terpenes and terpenoids are classified according to how many isoprene units they contain. They include hemiterpenes (one unit), monoterpenes (two units), sesquiterpenes (three units), diterpenes (four units), sesterterpenes (five units), triterpenes (six units), sesquarterpenes (seven units), tetraterpenes (eight units), polyterpenes (many units). Another class of terpene-containing molecules, the norisoterpenoids arise from peroxidase-catalyzed reactions on terpene molecules.

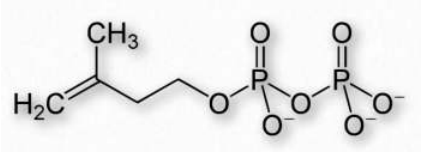


Figure 2.253 - Isopentenyl pyrophosphate

Examples

Common terpenes include monoterpenes of terpineol (lilacs), limonene (citrus), myrcene (hops), linalool (lavender), and pinene (pine). Higher order terpenes include taxadiene (diterpene precursor of taxol), lycopene (tetraterpenes), carotenes (tetraterpenes), and natural rubber (polyterpenes).

Steroid precursors geranyl pyrophosphate (monoterpene derivative), farnesyl pyrophosphate (sesquiterpene derivative), and squalene (triterpene) are all terpenes or derivatives of them. Vitamin A and phytol are derived from diterpenes.

Caffeine

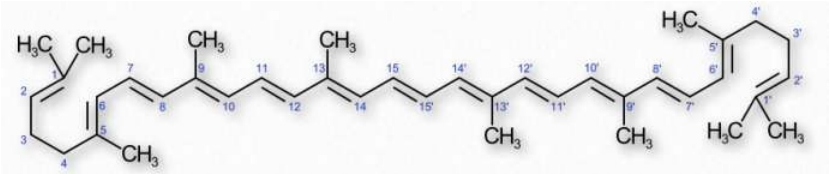


Figure 2.254 - Lycopene - A tetraterpene

Caffeine is the world's most actively consumed psychoactive drug (Figure 2.255). A methylxanthine alkaloid, caffeine is closely related to adenine and guanine and this is responsible for many effects on the body. Caffeine blocks the binding of adenosine on its receptor and consequently prevents the onset of drowsiness induced by adenosine. Caffeine readily crosses the blood-brain barrier and stimulates release of neurotransmitters. Caffeine stimulates portions of the autonomic nervous system and inhibits the activity of phosphodiesterase. The latter has the result of raising cAMP levels in cells, which activates protein kinase A and activates glycogen breakdown, inhibits TNF-α and leukotriene synthesis, which results in reduction of inflammation and innate immunity.

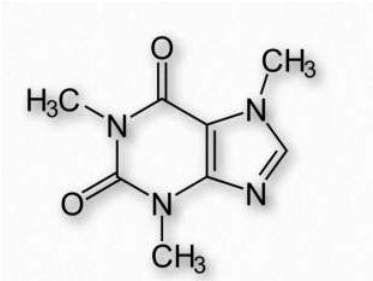


Figure 2.255 - Caffeine

Caffeine also has effects on the cholinergic system (acetylcholinesterase inhibitor), is an inositol triphosphate receptor 1 antagonist, and is a voltage independent activator of ryanodin receptors (a group of calcium channels found in skeletal muscle, smooth muscle, and heart muscle cells).

The half-life of caffeine in the body varies considerably. In healthy adults, it has a half-life of about 3-7 hours. Nicotine decreases the half-life and contraceptives and pregnancy can double it. The liver metabolizes caffeine, so the health of the liver is a factor in the halflife. CYP1A2 of the cytochrome P450 oxidase enzyme is primarily responsible. Caffeine is a natural pesticide in plants, paralyzing predator bugs.

Lipoprotein complexes and lipid movement in the body

Lipoprotein complexes are combinations of apolipoproteins and lipids bound to them that solubilize fats and other non-polar molecules, such as cholesterol, so they can travel in the bloodstream between various tissues of the body. The apolipoproteins provide the emulsification necessary for this. Lipoprotein complexes are formed in tiny "balls" with the water soluble apolipoproteins on the outside and non-polar lipids, such as fats, cholesteryl esters, and fat soluble vitamins on the inside.

They are categorized by their densities. These include (from highest density to the lowest) high density lipoproteins (HDLs), Low Density Lipoproteins (LDLs), Intermediate Density Lipoproteins (IDLs), Very Low Density Lipoproteins (VLDLs) and the chylomicrons. These particles are synthesized in the liver and small intestines.

Apolipoproteins

Name	Lipoprotein Complex(es)	Function
ApoA-I	HDL	Promotes Fat Movement to Liver
ApoA-II	HDL	Inhibit LCAT
ApoA-IV	Chylomicrons / HDL	Activate LCAT
AboB-48	Chylomicrons	Cholesterol Transport
AboB-100	VLDL / LDL	Bind LDL Receptor
ApoC-I	VLDL / LDL	Unknown
ApoC-II	All	Activate Lipoprotein Lipase
ApoC-III	All	Inhibit Lipoprotein Lipase
ApoD	HDL	Unknown
ApoE	VLDL / Chylomicrons / HDL	Clearance of Chylomicrons Remnants and VLDLs

Figure 2.256 - Apolipoproteins

Each lipoprotein complex contain a characteristic set of apolipoproteins, as shown in Figure 2.256. ApoC-II and ApoC-III are notable for their presence in all the lipoprotein complexes and the roles they play in activating (ApoC-II) or inactivating (ApoC-III) lipoprotein lipase. Lipoprotein lipase is a cellular enzyme that catalyzes the breakdown of fat from the complexes. ApoE (see below) is useful for helping the predict the likelihood of the occurrence of Alzheimer's disease in a patient.

Gene editing

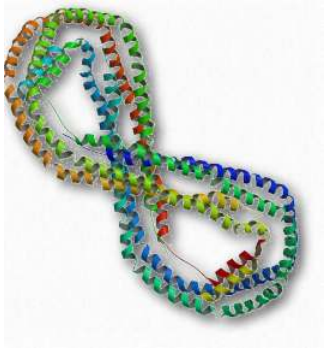


Figure 2.257 - ApoA-I

ApoB-48 and ApoB-100 are interesting in being coded by the same gene, but a unique mRNA sequence editing event occurs that converts one into the other. ApoB-100 is made in the liver, but ApoB-48 is made in the small intestine. The small intestine contains an enzyme that deaminates the cytidine at nucleotide #2153 of the common mRNA. This changes it to a uridine and changes the codon it is in from CAA (codes for glutamine) to UAA (stop codon). The liver does not contain this enzyme and does not make the change in the mRNA. Consequently, a shorter protein is synthesized in the intestine (ApoB-48) than the one that is made in the liver (ApoB-100) using the same gene sequence in the DNA.

Movement

The movement of fats in the body is important because they are not stored in all cells. Only specialized cells called adipocytes store fat. There are three relevant pathways in the body for moving lipids. As described below, they are 1) the exogenous pathway; 2) the endogenous pathway, and 3) the reverse transport pathway.

Exogenous pathway

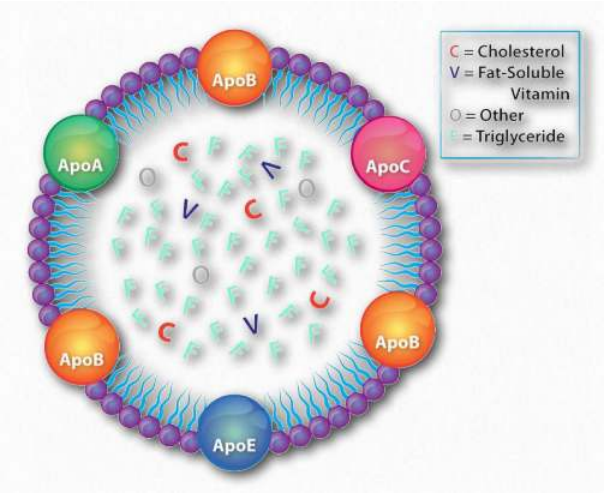


Figure 2.258 - Schematic diagram of a chylomicron Image by Aleia Kim

Dietary fat entering the body from the intestinal system must be transported, as appropriate, to places needing it or storing it. This is the function of the exogenous pathway of lipid movement in the body. All dietary lipids (fats, cholesterol, fat soluble vitamins, and other lipids) are moved by it. In the case of dietary fat, it begins its journey after ingestion first by being solubilized by bile acids in the intestinal tract. After passing through the stomach, pancreatic lipases clip two fatty acids from the fat, leaving a monoacyl glycerol. The fatty acids and monoacyl glycerol are absorbed by intestinal cells (enterocytes) and reassembled back into a fat, and then this is mixed with phospholipids, cholesterol esters, and apolipoprotein B-48 and processed to form chylomicrons (Figures 2.258 & 2.259) in the Golgi apparatus and smooth endoplasmic reticulum.

Exocytosis

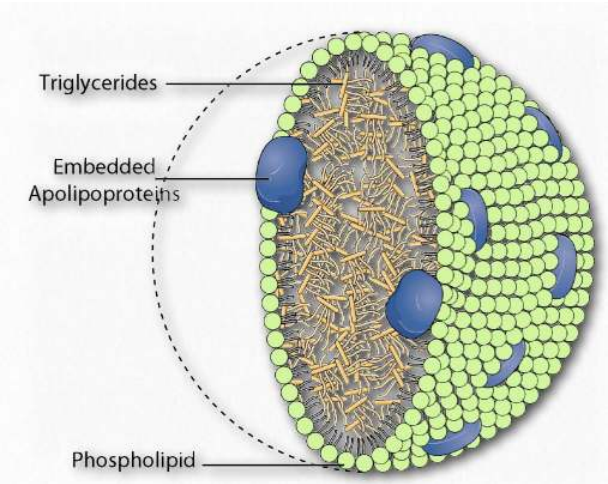


Figure 2.259 - Another perspective of a chylomicron WIkipedia

These are exocytosed from the cell into lymph capillaries called lacteals. The chylomicrons pass through the lacteals and enter the bloodstream via the left subclavian vein. Within the bloodstream, lipoprotein lipase breaks down the fats causing the chylomicron to shrink and become what is known as a chylomicron remnant. It retains its cholesterol and other lipid molecules.

The chylomicron remnants travel to the liver where they are absorbed (Figure 2.260). This is accomplished by receptors in the liver that recognize and bind to the ApoE of the chylomicrons. The bound complexes are then internalized by endocytosis, degraded in the lysosomes, and the cholesterol is disbursed in liver cells.

Endogenous pathway

The liver plays a central role in managing the body's needs for lipids. When lipids are needed by the body or when the capacity of the liver to contain more lipids than is supplied by the diet, the liver packages up fats and cholesteryl esters into Very Low Density Lipoprotein (VLDL) complexes and exports them via the endogenous pathway. VLDL complexes contain ApoB-100, ApoC-I,

ApoC-II, ApoC-III, and ApoE apolipoproteins. VLDLs enter the blood and travel to muscles and adipose tissue where lipoprotein lipase is activated by ApoC-II. In the muscle cells, the released fatty acids are taken up and oxidized. By contrast, in the adipoctyes, the fatty acids are taken up and reassembled back into triacylglycerides (fats) and stored in fat droplets. Removal of fat from the VLDLs causes them to shrink, first to Intermediate Density Lipoprotein (IDL) complexes (also called VLDL remnants) and then to Low Density Lipoprotein (LDL) complexes.

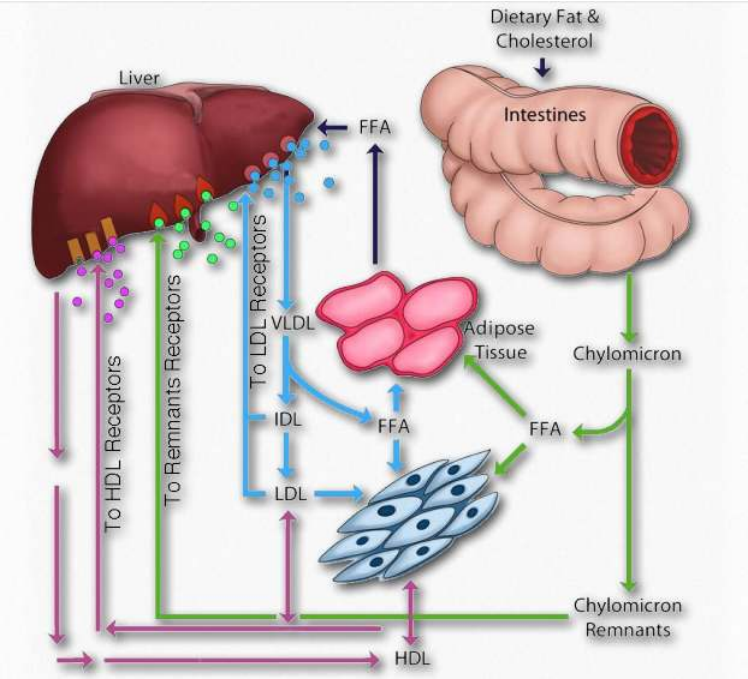


Figure 2.260 - Movement of lipids in the body - Green = exogenous pathway; Blue = endogenous pathway; Purple = reverse transport pathway Image by Aleia Kim

Shrinking of VLDLs is accompanied by loss of apolipoproteins so that LDLs are comprised primarily of ApoB-100. This lipoprotein complex is important because cells have receptors for it to bind and internalize it by receptor-mediated endocytosis (Figure 2.261). Up until this point, cholesterol and cholesteryl esters have traveled in chylomicrons, VLDLs, and IDLs as fat has been stripped stripped away. For cholesterol compounds to get into the cell from the lipoprotein complexes, they must be internalized by cells and that is the job of receptormediated endocytosis.

Reverse transport pathway

Another important consideration of the movement of lipids in the body is the reverse transport pathway (Figure 2.260). It is also called the reverse cholesterol transport pathway, since cholesterol is the primary molecule involved. This pathway involves the last class of lipoprotein complexes known as the High Density Lipoproteins (HDLs). In contrast to the LDLs, which are commonly referred to as “bad cholesterol” (see below also), the HDLs are known as “good cholesterol.”

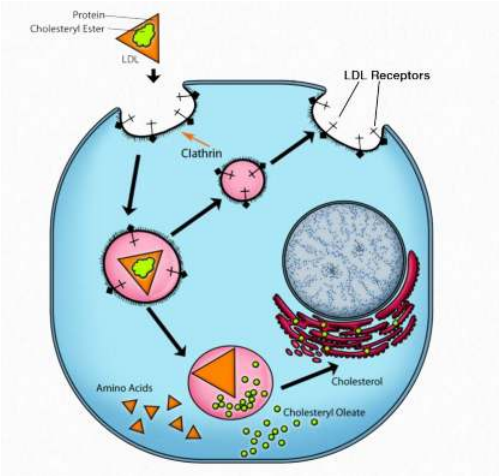


Figure 2.261 - The process of receptor-mediated endocytosis Image by Aleia Kim

HDLs are synthesized in the liver and small intestine. They contain little or no lipid when made (called depleted HDLs), but serve the role of “scavenger” for cholesterol in the blood and from remnants of other (damaged) lipoprotein complexes in the blood. To perform its task, HDLs carry the enzyme known as lecithincholesterol acyl transferase (LCAT), which they use to form cholesteryl esters using fatty acids from lecithin (phosphatidylcholine) and then they internalize them.

The cholesterol used for this purpose comes from the bloodstream, from macrophages, and from foam cells (macrophage-LDL complexes - Figure 2.262). Addition of cholesteryl esters causes the HDL to swell and Figure 2.261 - The process of receptor-mediated endocytosis Image by Aleia Kim when it is mature, it returns its load of cholesterol back to the liver or, alternatively, to LDL molecules for endocytosis. HDLs have the effect of lowering levels of cholesterol and it is for that reason they are described as “good cholesterol.”

Regulation of lipid transfer

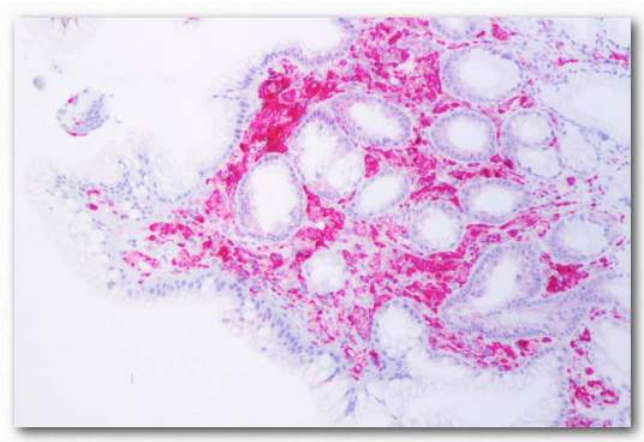


Figure 2.262 - Foam cell aggregate Wikipedia

It is important that cells get food when they need it so some control of the movement of nutrients is critical. The liver, which plays the central role in modulating blood glucose levels, is also important for performing the same role for lipids. It accomplishes this task the use of specialized LDL receptors on its surface. Liver LDL receptors bind LDLs that were not taken up by other cells in their path through the bloodstream. High levels of LDLs are a signal to the liver to reduce the creation of VLDLs for release.

People with the genetic disease known as familial hypercholesterolemia, which manifests with dangerously high levels of LDLs, lack properly functioning LDL receptors on their liver cells.Figure 2.263 - Progression of atherosclerosis Wikipedia Figure 2.263 - Progression of atherosclerosis Wikipedia Figure 2.263 - Progression of atherosclerosis Wikipedia

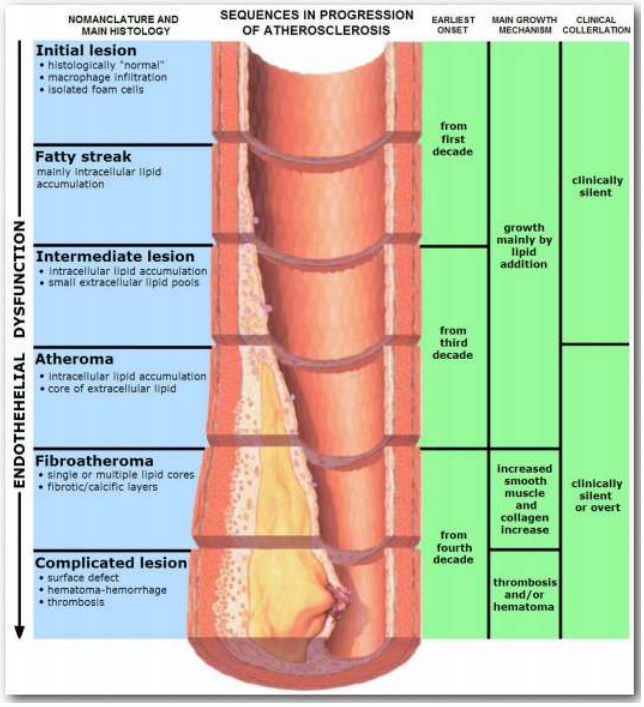


Figure 2.263 - Progression of atherosclerosis Wikipedia

In sufferers of this disease, the liver never gets the signal that the LDL levels are high. In fact, to the liver, it appears that all VLDLs and LDLs are being taken up by peripheral tissues, so it creates more VLDLs to attempt to boost levels. Untreated, the disease used to be fatal early, but newer drugs like the statins have significantly increased life spans of patients. Cellular needs for the contents of LDLs are directly linked to the levels of synthesis of LDL receptors on their membranes. As cells are needing more cholesterol, their synthesis of components for receptors goes up and it decreases as need diminishes.

Good cholesterol / bad cholesterol

It is commonly accepted that “high cholesterol” levels are not healthy. This is due, at least indirectly, to the primary carriers of cholesterol, the LDLs. A primary function of the LDLs is to deliver cholesterol and other lipids directly into cells by receptor mediated endocytosis (Figure 2.237). High levels of LDLs, though, are correlated with formation of atherosclerotic plaques (Figure 2.263 & 2.264) and incidence of atherosclerosis, leading to the description of them as “bad cholesterol.” This is because when LDL levels are very high, plaque formation begins. It is thought that reactive oxygen species (higher in the blood of smokers) causes partial oxidation of fatty acid groups in the LDLs. When levels are high, they tend to accumulate in the extracellular matrix of the epithelial cells on the inside of the arteries. Macrophages of the immune system take up the damaged LDLs (including the cholesterol).

Since macrophages can’t control the amount of cholesterol they take up, cholesterol begins to accumulate in them and they take on appearance that leads to their being described as “foam cells.” With too much cholesterol, the foam cells, however, are doomed to die by the process of programmed cell death (apoptosis). Accumulation of these, along with scar tissue from inflammation result in formation of a plaque. Plaques can grow and block the flow of blood or pieces of them can break loose and plug smaller openings in the blood supply, ultimately leading to heart attack or stroke.

Good cholesterol

On the other hand, high levels of HDL are inversely correlated with atherosclerosis and arterial disease. Depleted HDLs are able to remove cholesterol from foam cells. This occurs as a result of contact between the ApoA-I protein of the HDL and a transport protein on the foam cell (ABC-G1). Another transport protein in the foam cell, ABCA-1 transports extra cholesterol from inside the cell to the plasma membrane where it is taken up into the HDL and returned to the liver or to LDLs by the reverse transport cholesterol pathway.

Figure 2.264 - Actual carotid artery plaque Wikipedia

Deficiency of the ABCA-1 gene leads to Tangier disease. In this condition, HDLs are almost totally absent because they remain empty as a result of not being able to take up cholesterol from foam cells, so they are destroyed by the body.

ApoE and Alzheimer's disease

ApoE is a component of the chylomicrons and is also found in brain, macrophages, kidneys, and the spleen. In humans, it is found in three different alleles, E2, E3, and E4. The E4 allele (present at about 14% of the population) is associated with increased likelihood of contracting Alzheimer's disease. People heterozygous for the allele are 3 times as likely to contract the disease and those homozygous for it are 15 times as likely to do so. It is not known why this gene or allele is linked to the disease. The three alleles differ only slightly in amino acid sequence, but the changes do cause notable structural differences. The E4 allele is associated with increased calcium ion levels and apoptosis after injury. Alzheimer's disease is associated with accumulation of aggregates of the β- amyloid peptide. ApoE does enhance the proteolytic breakdown of it and the E4 isoform is not as efficient in these reactions as the other isoforms.

This page titled 11.3: Structure and Function - Lipids and Membranes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **2.8: Structure and Function - Lipids and Membranes** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **9.1: Structure and Function - Lipids and Membranes** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

CHAPTER OVERVIEW

12: Membranes and Membrane Proteins

12.1: Basic Concepts in Membranes
12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics
12.3: Working with Lipids
12.4: Transport in Membranes
12.5: Problems
12.6: Membranes and Membrane Proteins
12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion
12.8: Diffusion Across a Membrane - Channels
12.9: Diffusion Across a Membrane - Pores
12.10: Active Transport

12: Membranes and Membrane Proteins is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

12.1: Basic Concepts in Membranes

Source: BiochemFFA_3_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Lipid bilayers

The protective membrane around cells contains many components, including cholesterol, proteins, glycolipids, glycerophospholipids, and sphingolipids. The last two of these will, when mixed vigorously with water, spontaneously form what is called a lipid bilayer (Figure 3.1), which serves as a protective boundary for the cell that is largely impermeable to the movement of most materials across it. With the notable exceptions of water, carbon dioxide, carbon monoxide, and oxygen, most polar/ionic require transport proteins to help them to efficiently navigate across the bilayer. The orderly movement of these compounds is critical for the cell to be able to 1) get food for energy; 2) export materials; 3) maintain osmotic balance; 4) create gradients for secondary transport; 5) provide electromotive force for nerve signaling; and 6) store energy in electrochemical gradients for ATP production (oxidative phosphorylation or photosynthesis). In some cases, energy is required to move the substances (active transport).

Facilitated Diffusion

In other cases, no external energy is required and they move by diffusion through specific cellular channels. This is referred to as facilitated diffusion. Before we discuss movement of materials across membranes, it is appropriate we discuss the composition of cellular membranes. Plasma membranes differ from cell walls both in the materials comprising them and in their flexibility. Cell walls will be covered near the end of this chapter.

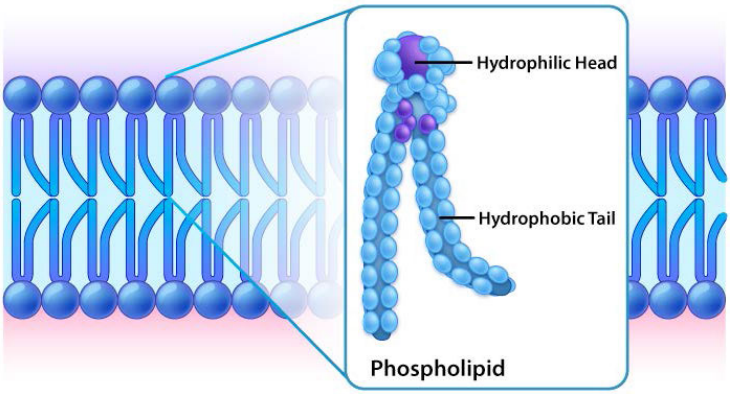


Figure 3.1: Lipid bilayer closeup Image by Aleia Kim

Though some cells do not have cell walls (animal cells) and others do (bacteria, fungi, and plants), there is commonality among cells in that they all possess plasma membranes. There is also commonality in the components of the membranes, though the relative amount of constituents varies. Figures 3.1 and 3.2 illustrate the structure and environments of plasma membranes. All plasma membranes contain a significant amount of amphiphilic substances linked to fatty acids. These include the glycerophospholipids and the sphingolipids. The fatty acid(s) are labeled as hydrophobic tails in the figures.

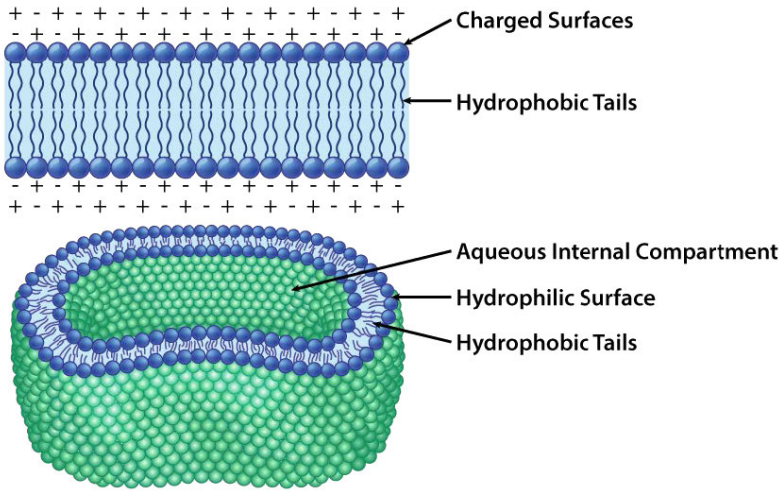


Figure 3.2: Organization of the lipid bilayer Image by Aleia Kim

Hydrophilic heads

The composition of the hydrophilic heads varies considerably. In glycerophospholipids, a phosphate is always present, of course, and it is often esterified to another substance to make a phosphatide (Figure 3.3). Common compounds linked to the phosphate (at the X position) include serine, ethanolamine, and choline. These vary in the their charges so in this way, the charge on the external or internal surface can be controlled. Cells tend to have more negative charges on the exterior half of the lipid bilayer (called the outer leaflet) and more positive charges on the interior half (inner leaflet).

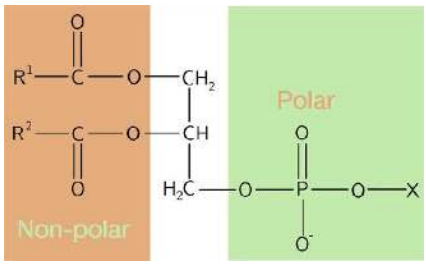


Figure 3.3 - Schematic diagram of a phosphatide

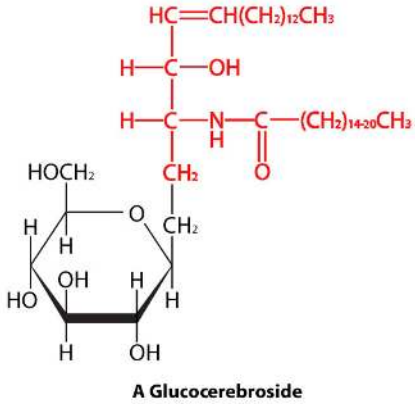


Figure 3.4: A sphingolipid. Polar head in black. Nonpolar tail in red Image by Aleia Kim

Sphingolipids

In sphingolipids (Figure 3.4), the hydrophilic head can contain a phosphate linked to ethanolamine or choline and this describes the structure of sphingomyelin, an important component of neural membranes. Most sphingolipids lack the phosphate and have instead a hydrophilic head of a single sugar (cerebrosides) or a complex oligosaccharide (gangliosides).

Water exclusion

In each case, the glycerophospholipid or sphingolipid has one end that is polar and one end that is non-polar. As we saw in the organization of amino acids with hydrophobic side chains occurring preferentially on the inside of a folded protein to exclude water, so too do the non-polar portions of these amphiphilic molecules arrange themselves so as to exclude water. Remember that the cytoplasm of a cell is mostly water and the exterior of the cell is usually bathed in an aqueous layer. It therefore makes perfect sense that the polar portions of the membrane molecules arrange themselves as they do - polar parts outside interacting with water and non-polar parts in the middle of the bilayer avoiding/excluding water.

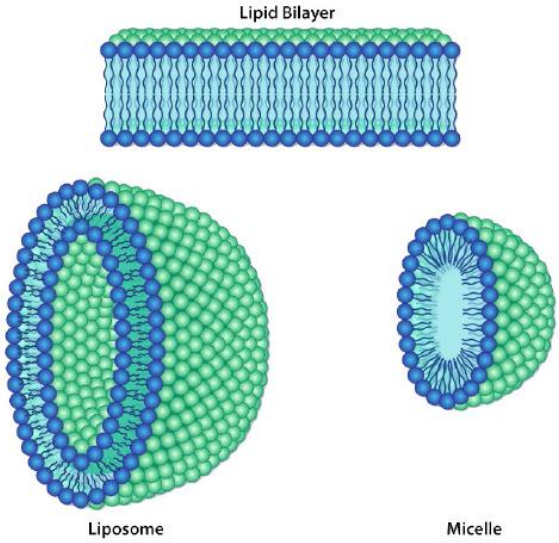


Figure 3.6 - Different lipid bilayer structures Image by Aleia Kim

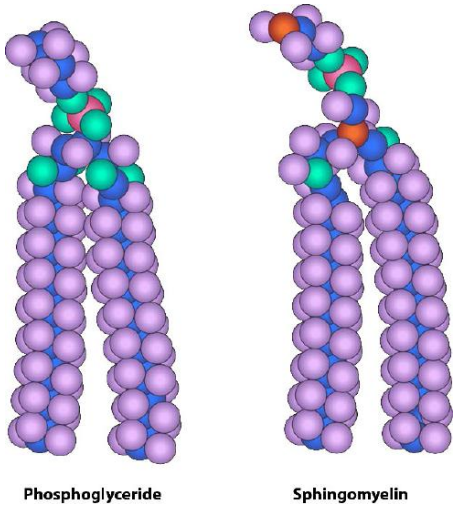


Figure 3.5 - Similarity of form between a phosphoglyceride and sphingomyelin Image by Aleia Kim

Composition Bias

The plasma membrane has distinct biases of composition relative to its inside and the outside (Figure 3.7). First, glycosylation (of lipids and proteins) has the sugar groups located almost exclusively on the outside of the cell, away from the cytoplasm (Figure 3.8). Among the membrane lipids, sphingolipids are much more commonly glycosylated than glycerophospholipids. In addition, some of the glyerophospholipids are found preferentially on one side or the other (Figure 3.7). Phosphatidylserine and phosphatidylethanolamine are found preferentially within the inner leaflet of the plasma membrane, whereas phosphatidylcholine tends to be located on the outer leaflet. In the process of apoptosis, the phosphatidylserines appear on the outer leaflet where they serve as a signal to macrophages to bind and destroy the cell. Sphingolipids are found preferentially in the plasma membrane and are almost completely absent from mitochondrial and endoplasmic reticulum membranes (Figure 3.9).

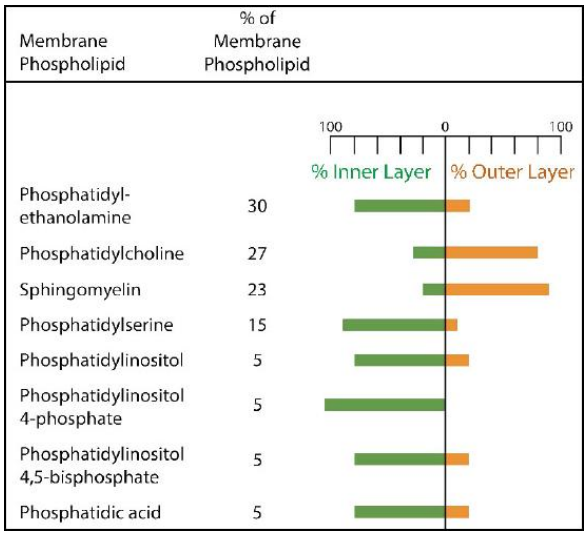


Figure 3.7: Differential distribution of membrane lipids by inner and outer leaflet Image by Pehr Jacobson

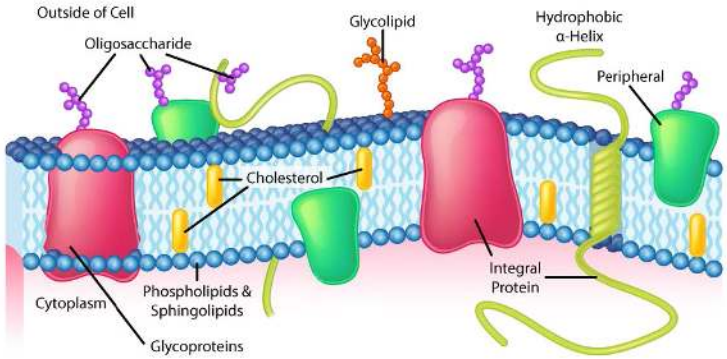


Figure 3.8 - Molecular organization of the lipid bilayer Image by Aleia Kim

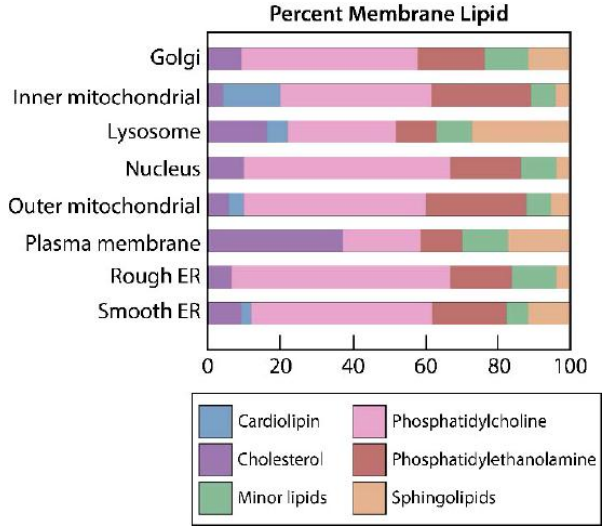


Figure 3.9 - Distribution of lipids in organelle membranes Image by Pehr Jacobson

Organelle membranes

Bias of lipid composition also exists with respect to organelle membranes. The unusual diphosphodiglycerolipid known as cardiolipin, for example, is almost only found in mitochondrial membranes (see HERE) and like phosphatidylserine, its movement is an important step in apoptosis. In signaling, phosphatidylinositols play important roles providing second messengers upon being cleaved (see HERE).

Lateral Diffusion

Movement of lipids within each leaflet of the lipid bilayer occurs readily and rapidly due to membrane fluidity. This type of movement is called lateral diffusion and can be measured by the technique called FRAP (Figure 3.10, see HERE also). In this method, a laser strikes and stains a section of the lipid bilayer of a cell, leaving a spot as shown in B. Over time, the stain diffuses out ultimately across the entire lipid bilayer, much like a drop of ink will diffuse throughout when added to a glass of water. A measurement of the rate of diffusion gives an indication of the fluidity of a membrane.

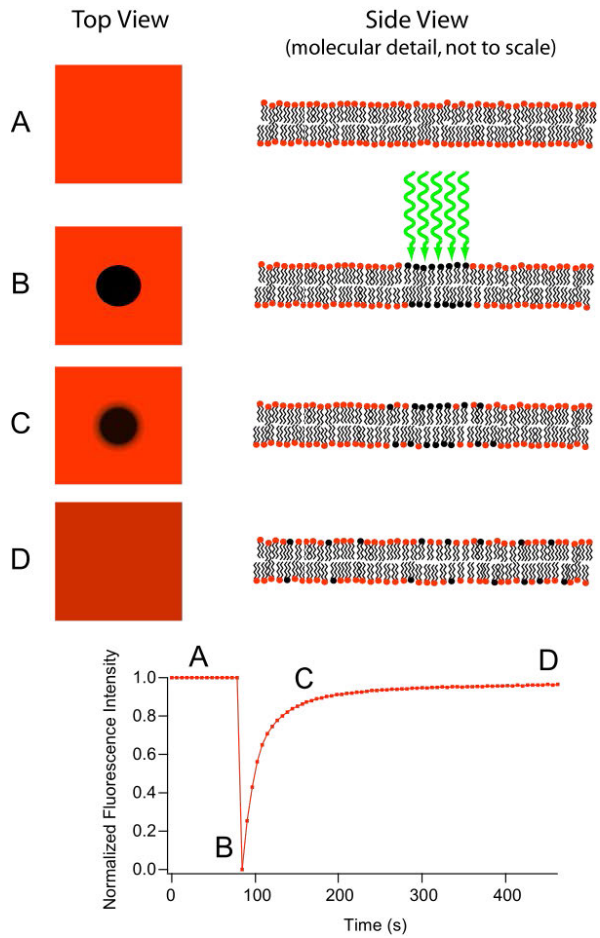


Figure 3.10 - Fluorescence recovery after photobleaching (FRAP)

Transverse Diffusion

While the movement in lateral diffusion occurs rapidly, movement of molecules from one leaflet over to the other leaflet occurs much more slowly. This type of molecular movement is called transverse diffusion and is almost nonexistent in the absence of enzyme action. Remember that there is a bias of distribution of molecules between leaflets of the membrane, which means that something must be moving them.

There are three enzymes that catalyze movement of compounds in transverse diffusion. Flippases move membrane glycerophospholipids/ sphingolipids from outer leaflet to inner leaflet (cytoplasmic side) of cell. Floppases move membrane lipids in the opposite direction. Scramblases move in either direction.

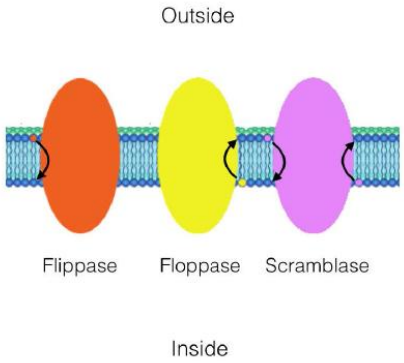


Figure 3.11 - Catalytic action of a flippase, a floppase, and a scramblase

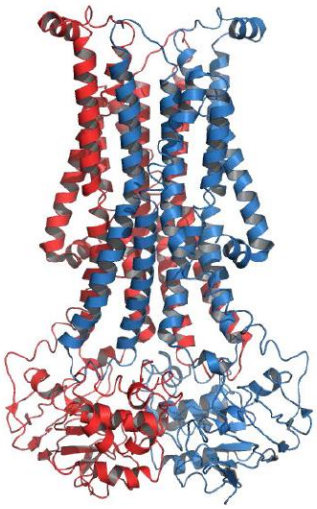


Figure 3.12 - Structure of a flippase Wikipedia

Other components of lipid bilayer

Besides glycerophospholipids and sphingolipids, there are other materials commonly found in lipid bilayers of cellular membranes. Two important prominent ones are cholesterol (Figure 3.13) and proteins. Besides serving as a metabolic precursor of steroid hormones and the bile acids, cholesterol's main role in cells is in the membranes. The flatness and hydrophobicity of the sterol rings allow cholesterol to interact with the nonpolar portions of the lipid bilayer while the hydroxyl group on the end can interact with the hydrophilic part.

Membrane fluidity

Cholesterol's function in the lipid bilayer is complex (Figure 3.13). It influences membrane fluidity. Figure 3.14 shows the phase transition for a membrane as it is heated, moving from a more "frozen" character to that of a more "fluid" one as the temperature rises. The mid-point of this transition, referred to as the Tm, is influenced by the fatty acid composition of the lipid bilayer compounds. Longer and more saturated fatty acids will favor higher Tm values, whereas unsaturation and short fatty acids will favor lower Tm values. It is for this reason that fish, which live in cool environments, have a higher level of unsaturated fatty acids in them - to use to make membrane lipids that will remain fluid at ocean temperatures. Interestingly, cholesterol does not change the Tm value, but instead widens the transition range between frozen and fluid forms of the membrane, allowing it to have a wider range of fluidity.

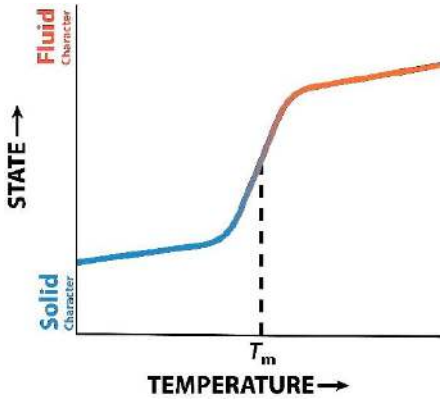


Figure 3.14 - Membrane transition temperature Image by Aleia Kim

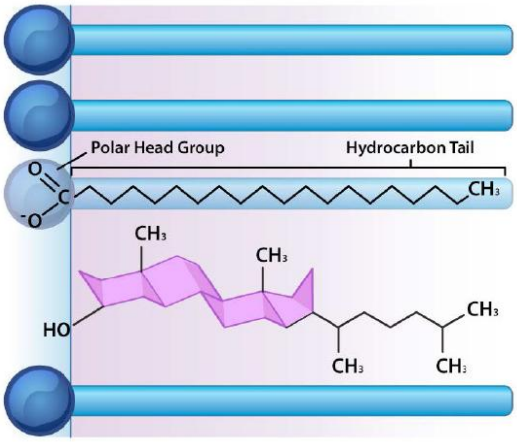


Figure 3.13 - Cholesterol in the lipid bilayer Image by Aleia Kim

Lipid Rafts

Cholesterol is also abundantly found in membrane structures called lipid rafts. Depicted in Figure 3.15, lipid rafts are organized structures within the membrane typically containing signaling molecules and other integral membrane proteins. Lipid rafts affect membrane fluidity, neurotransmission, and trafficking of receptors and membrane proteins.

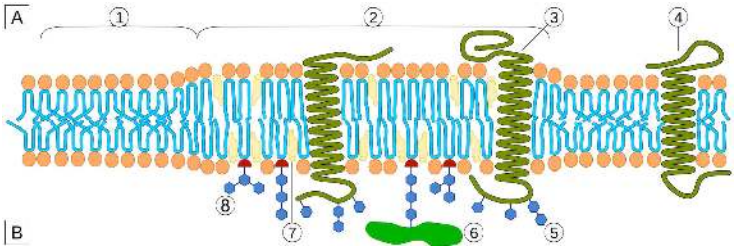


Figure 3.15 - A lipid raft - 1 = Non-raft membrane / 2 = Lipid raft / 3 = Lipid raft associated transmembrane protein / 4 = Non-raft membrane protein / 5 = Glycosylation modifications (on glycoproteins and glycolipids) / 6 = GPIanchored protein / 7 = Cholesterol / 8 = Glycolipid

Features

Distinguishing features of the rafts is that they are more ordered than the bilayers surrounding them, containing more saturated fatty acids (tighter packing and less disorganization, as a result) and up to 5 times as much cholesterol. They also are relatively rich in sphingolipids, with as much as 50% greater quantities of sphingomyelin than surrounding areas of the bilayer. The higher concentration of cholesterol in the rafts may be due to its greater ability to associate with sphingolipids (Figure 3.16). Some groups, such as prenylated proteins, like RAS, may be excluded from lipid rafts.

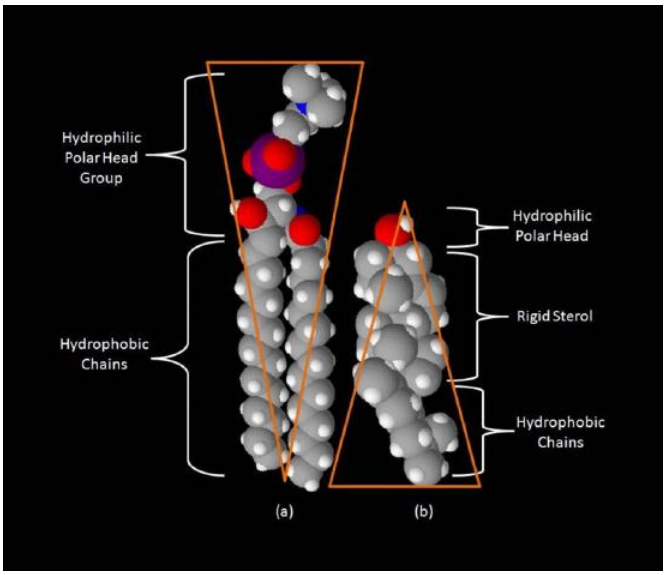


Figure 3.16 - A sphingolipid (left) associating with cholesterol (right)

Lipid rafts may provide concentrating platforms after individual protein receptors bind to ligands in signaling. After receptor activation takes place at a lipid raft, the signaling complex would provide protection from nonraft enzymes that could inactivate the signal. For example, a common feature of signaling systems is phosphorylation, so lipid rafts might provide protection against dephosphorylation by enzymes called phosphatases. Lipid rafts appear to be involved in many signal transduction processes, such as T cell antigen receptor signaling, B cell antigen receptor signaling, EGF receptor signaling, immunoglobulin E signaling, insulin receptor signaling and others. For more on signaling, see HERE.

Barrier

Transport of materials across membranes is essential for a cell to exist. The lipid bilayer is an effective barrier to the entry of most molecules and without a means of allowing food molecules to enter a cell, it would die. The primary molecules that move freely across the lipid bilayer are small, uncharged ones, such as H2O, CO2, CO, and O2, so larger molecules, like glucose, that the cell needs for energy, would be effectively excluded if there were not proteins to facilitate its movement across the membrane.

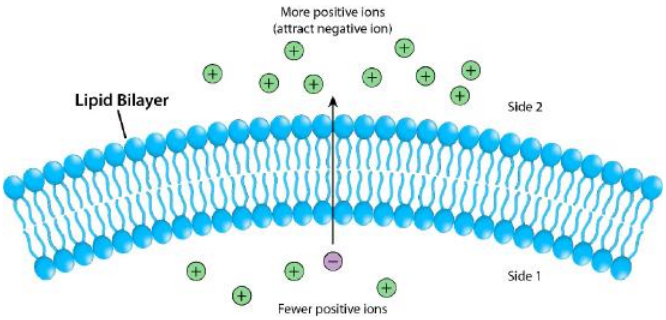


Figure 3.17 - The lipid bilayer as a barrier Image by Pehr Jacobson

Figure 3.17 depicts the barrier that the lipid bilayer provides to movement across it and the pressures (ionic attraction, in this case) that can affect movement. Potential energy from charge and concentration differences are harvested by cells for purposes that include synthesis of ATP, and moving materials against a concentration gradient in a process called active transport.

Membrane proteins

Proteins in a lipid bilayer can vary in quantity enormously, depending on the membrane. Protein content by weight of various membranes typically ranges between 30 and 75% by weight. Some mitochondrial membranes can have up to 90% protein. Proteins linked to and associated with membranes come in several types.

Transmembrane proteins

Transmembrane proteins are integral membrane proteins that completely span from one side of a biological membrane to the other and are firmly embedded in the membrane (Figure 3.18). Transmembrane proteins can function as docking sites for attachment (to the extracellular matrix, for example), as receptors in the cellular signaling system, or facilitate the specific transport of molecules into or out of the cell.

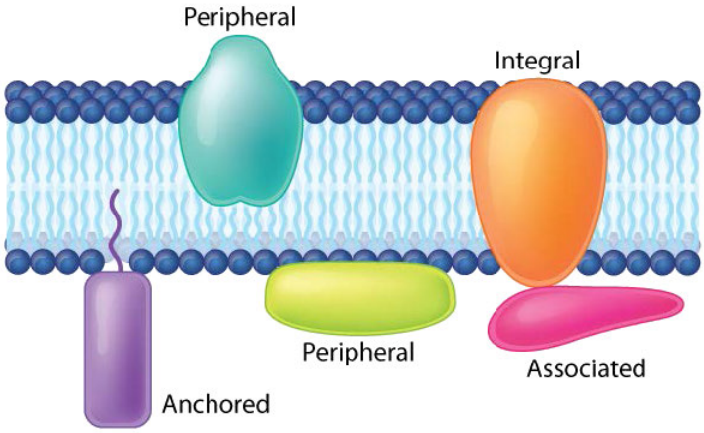


Figure 3.18 - Membrane protein types Image by Aleia Kim

Example of integrated/ transmembrane proteins include those involved in transport (e.g., Na+/K+ ATPase), ion channels (e.g., potassium channel of nerve cells) and signal transduction across the lipid bilayer (e.g., GProtein Coupled Receptors).

Peripheral membrane proteins interact with part of the bilayer (usually does not involve hydrophobic interactions), but do not project through it. A good example is phospholipase A2, which cleaves fatty acids from glycerophospholipids in membranes. Associated membrane proteins typically do not have external hydrophobic regions, so they cannot embed in a portion of the lipid bilayer, but are found near them. Such association may arise as a result of interaction with other proteins or molecules in the lipid bilayer. A good example is ribonuclease.

Anchored membrane proteins

Anchored membrane proteins are not themselves embedded in the lipid bilayer, but instead are attached to a molecule (typically a fatty acid) that is embedded in the membrane (Figure 3.19). The oncogene family of proteins known as ras are good examples. These proteins are anchored to the lipid bilayer by attachment to non-polar farnesyl groups catalyzed by the enzyme farnesyltransferase.

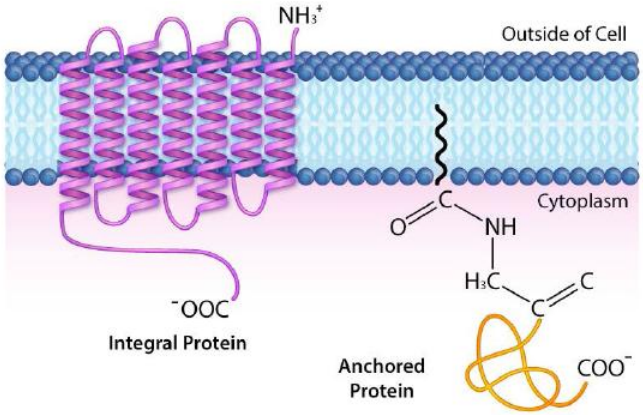


Figure 3.19 - Integral and anchored proteins Image by Aleia Kim

Finer classification

A more detailed classification scheme further categorizes the integral and anchored proteins into six different types (Figure 3.20). Type I and Type II have only one portion of the protein pass through the membrane. They differ in the orientation of the amine and carboxyl end with respect to inside/outside. Type I transmembrane proteins have the amino terminus on the outside and carboxy

terminus on the inside, whereas Type II proteins have this reversed. Type III proteins are a single polypeptide chain that has multiple regions of it cross back and forth across the membrane, often to form a channel. Type IV is a multi-polypeptide protein which has multiple crossings of the membrane. Type V transmembrane proteins do not have a part of them that crosses the membrane, but they are anchored to the membrane by a lipid (such as a fatty acid) embedded in the lipid bilayer. Type VI transmembrane proteins both have a portion of them that crosses the membrane and they are attached to a lipid embedded in the lipid bilayer.

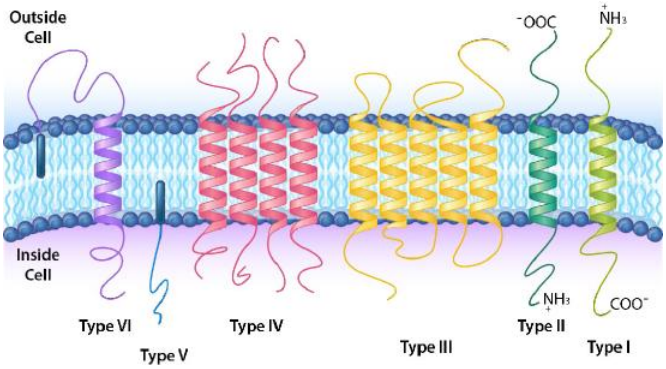


Figure 3.20 Types of integral and anchored membrane proteins Image by Aleia Kim

Blood Types

Cells have hundreds-thousands of membrane proteins and the protein composition of a membrane varies with its function and location. Glycoproteins embedded in membranes play important roles in cellular identification. Blood types, for example, differ from each other in the structure of the carbohydrate chains projecting out from the surface of the glycoprotein in their membranes (Figure 3.21).

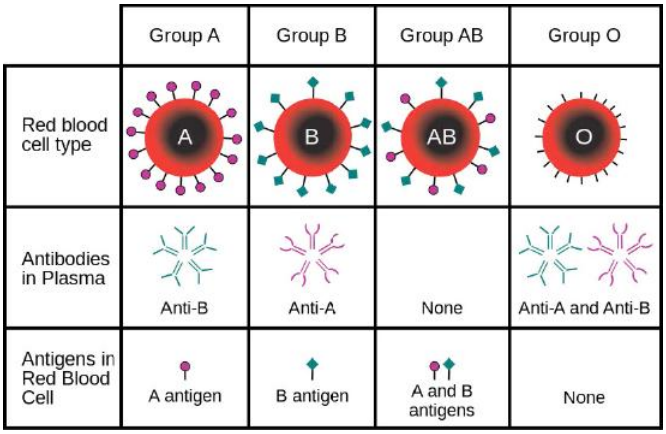


Figure 3.21

Osmotic Pressure

Membranes provide barriers/boundaries for most molecules, but the permeability of water across a lipid bilayer creates a variable that must be considered. The variable here is osmotic pressure. Osmotic pressure (loosely) refers to the tendency of a solution to take in water by the process of osmosis. In Figure 3.22, one can see a visual representation of the concept of the pressure.

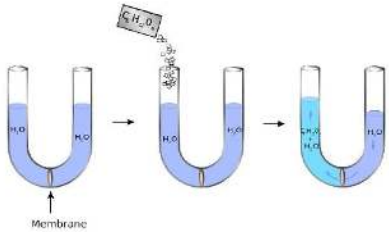


Figure 3.22

A U-shaped tube has at its bottom a semipermeable membrane. Water can pass through the membrane, but sugar molecules (C6H12O6) cannot. On the left side, sugar is added creating a concentration difference between the right and left chambers. Water diffuses across the membrane from right to left in an attempt to equalize the concentrations, causing the level of the right side to decrease and the left side to increase. The pressure resulting from the differences in height is felt at the membrane.

Equalizing concentrations

The liquid on the right does not completely move to the left, though, as might be expected if the only force involved is equalizing the concentration of sugar across the membrane (no sugar on right = no water). Instead, an equilibrium of sorts of water levels is reached even though the concentrations don't equal out. The pressure existing at the membrane then from the differences in level corresponds to the osmotic pressure of the mixture. The osmotic pressure of a solution is the pressure difference needed to halt the flow of solvent across a semipermeable membrane. Osmotic pressure can also be thought of as the pressure required to counter osmosis. The osmotic pres- Figure 3.21 - Blood types arise from cell surface glycoproteins Figure 3.22 - Osmotic pressure. Water diffuses leftwards to try to equalize the solute concentration. The pressure realized at the membrane in the right figure is the osmotic pressure sure of a dilute solution mathematically behaves like the ideal gas law

$$P_{osmotic} = \frac{nRT}{V} \tag{12.1.1}$$

where n is the number of moles, R is the gas constant, T is the temperature in Kelvin, and V is the volume.

It is more convenient in solutions to work with molarity, so

$$P_{osmotic} = MR^{*}T \tag{12.1.2}$$

where M is the molarity of the dissolved molecules, R* is the gas constant expressed in (L atm)/(K mol), and T is the temperature. The Greek letter Π is used to refer to the $P_{osmotic}$term, so

$$\Pi = MR^{*}T \tag{12.1.3}$$

Remember when calculating the molarity to include the molarity of each particle. For example, when one dissolves sucrose in solution, it does not split into smaller particles, so

$$Molarity_{Particles} = Molarity_{Sucrose} \tag{12.1.4}$$

However, for salts, like KOH, which forms two ions in solution (K+ and OH-),

$Molarity_{Particles}$= 2* $Molarity_{KOH}$.

Significant consideration

Osmotic pressure is a significant consideration for cells. Consider the fact that water can move freely across cellular membranes, but most of the contents of the cell, such as proteins, DNA, ions, sugars, etc., cannot. Second, the concentration of these compounds inside the cell is different than the concentration of them outside of the cell. Third, since water can move through the lipid bilayer, it will tend to move in the direction that will tend to equalize solute.

Hypotonic, hypertonic, isotonic

We consider three situations (Figure 3.23). First, if the concentration of solutes is greater inside the cell than outside, water will tend to move into the cell, causing the cell to swell. This circumstance is called hypotonic. Conversely, if the solute concentration is greater outside the cell than inside of it, water will exit the cell and the cell with shrink. This is a hypertonic situation. Last, if the concentrations of solutes into and outside of the cell is equal, this is called an isotonic solution. Here, no movement of water occurs across the cell membrane and the cell retains its size.

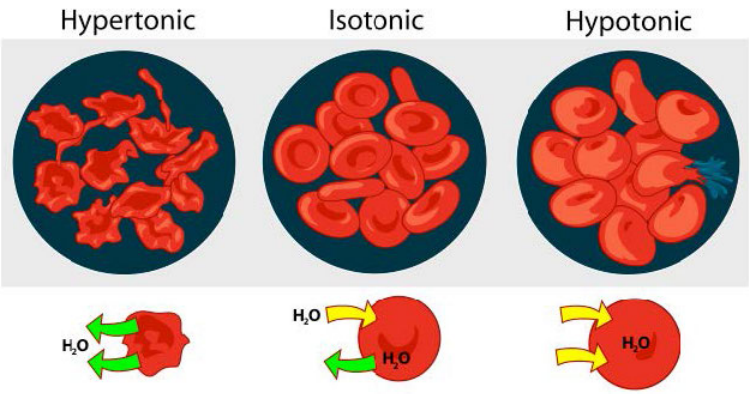
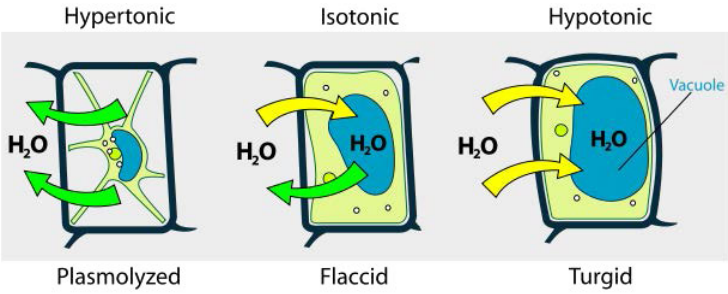


Figure 3.23 - The effect of three different osmotic conditions on red blood cells concentrations on either side of the membrane

If the osmotic pressure is greater than the forces holding together a cellular membrane, the cell will rupture. Because of this, some cells have built in defenses to prevent problems. Plant cells, for example have a fairly rigid cell wall that resists expansion in hypotonic solutions (Figure 3.24). Bacteria also have a cell wall that provides protection.



This page titled 12.1: Basic Concepts in Membranes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Basic Concepts in Membranes** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **9.2: Basic Concepts in Membranes** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics

Learning Goals (ChatGPT o3-mini, 2/2/25)

1. **Describe Membrane Bilayer Composition and Organization:**
   • Explain the structural organization of a membrane bilayer, including the roles of amphiphilic lipids, peripheral and integral membrane proteins, and carbohydrate attachments.
2. **Understand Lipid Distribution in Cells:**
   • Interpret data on lipid composition across various cellular compartments (e.g., nucleus, mitochondria, ER, plasma membrane) and explain the significance of asymmetry between the inner and outer leaflets of membranes.
3. **Explain Lipid Synthesis and Trafficking:**
   • Describe the biosynthetic pathways and cellular locations (primarily the ER) for lipid production and how these pathways contribute to the proper distribution and asymmetric arrangement of lipids in membranes.
4. **Analyze Membrane Dynamics:**
   • Discuss the dynamic nature of membranes by explaining lateral diffusion, flip-flop (transverse) diffusion, and the roles of specific enzymes (flippases, floppases, and scramblases) in maintaining lipid asymmetry.
5. **Interpret Thermodynamic Principles in Membrane Assembly:**
   • Apply the concepts of Gibbs free energy (ΔG), enthalpy, and entropy to explain the spontaneous formation of micelles, bilayers, and other lipid aggregates, including the role of the critical micelle concentration (CMC).
6. **Differentiate Lipid Phases and Phase Transitions:**
   • Compare different phases of lipid bilayers (gel, liquid crystalline, liquid-ordered, and liquid-disordered) and describe how factors such as temperature, lipid saturation, and cholesterol content affect membrane fluidity and permeability.
7. **Evaluate the Role of Cholesterol in Membrane Structure:**
   • Explain how cholesterol modulates membrane fluidity and stability, influences phase behavior, and contributes to the formation of lipid rafts and nanodomains.
8. **Understand Lipid Rafts and Nanodomains:**
   • Describe the concept of lipid rafts, their composition (enriched in sphingolipids, cholesterol, and specific proteins), and their role in organizing membrane domains and facilitating signal transduction.
9. **Interpret Lipid Phase Diagrams:**
   • Analyze phase diagrams for lipid mixtures (e.g., DPPC with cholesterol) to understand how temperature and lipid composition dictate membrane phase behavior and coexistence of different phases.
10. **Explore Membrane Curvature and Nonplanar Structures:**
    • Explain how lipids contribute to nonplanar membrane geometries, such as vesicles, tubular structures, and hexagonal phases, and discuss their biological relevance in processes like fusion, fission, and organelle biogenesis.
11. **Examine Dynamic Membrane Events:**
    • Discuss dynamic events such as membrane trafficking, vesicle fusion/fission, and the insertion of membrane proteins, and explain how these processes are driven by both physical forces and protein-mediated mechanisms.
12. **Introduce Lipid Signaling Pathways:**
    • Describe how enzymatic cleavage of specific membrane lipids produces secondary messengers that participate in intracellular signaling, linking membrane composition to cellular responses.
13. **Differentiate Lipid Droplets and Lipoproteins:**
    • Compare the structures and functions of lipid droplets and lipoproteins, including their formation (e.g., budding from the ER), composition (monolayer vs. bilayer), and roles in energy storage and lipid transport.

These learning goals provide a roadmap for understanding how lipid structure and dynamics underpin membrane function and cellular processes, preparing you for more detailed studies of membrane biochemistry and lipid signaling in advanced topics.

12.2.1: An overview of lipid bilayers

A membrane bilayer consists of more than just two leaflets of amphiphilic molecules. It also contains membrane proteins (which we will discuss in the next chapter), which can also be attached to carbohydrates. Most assuredly, you have seen various representations of a bilayer before. Before we proceed with a more detailed description of the lipids in the bilayer and their associated properties, we present Figure **12.2.1** to focus our discussion.

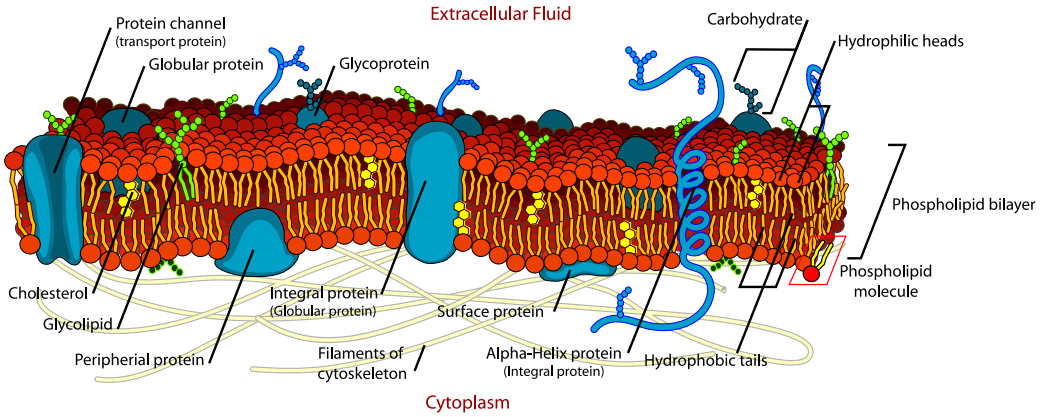


Figure **12.2.1**: A representation of a membrane bilayer with peripheral and integral membrane proteins. https://cnx.org/contents/FPtK1zmh@8.108:q2X995E3@12/The-Cell-Membrane. Creative Commons Attribution 4.0 International license.

To understand the movement of lipids in an actual cell, a better understanding of lipid synthesis and trafficking in cells is important. Table **12.2.1** below shows the distribution of four classes of lipids in a macrophage, a type of immune cell (Andreyev, A.Y. et al), while the following figure shows the lipid composition of membrane organelles.

Table **12.2.1**: Distribution of Lipids in Resting Macrophage

Lipid Categories	Nucleus	Mitochondria	ER	Plasma Memb	microsome	cytosol	Whole cell
Glycero-phospholipids	149	152	150	151	142	109	155
Prenol lipids	5	5	5	5	5	5	5
Sphingolipids	48	47	48	48	48	47	48
Sterol lipids	13	12	12	13	11	5	12
Total	215	216	215	217	206	166	220

Lipids in membranes are often distributed asymmetrically. A biological membrane's inner and outer leaflets usually have different PL compositions. For example, the outer leaflet in red blood cell membranes is composed mostly of sphingomyelin (SM) and PC. In contrast, the inner leaflet is composed mostly of PE and phosphatidyl serine (PS). This phospholipid contains the amino acid serine linked through its side chain (-CH2OH) to phosphate in position 3 of diacylglycerol. With a negative charge on the phosphate and carboxylate and a positive charge on the amine of PS, this phospholipid is acidic with a net negative charge. All the PS is located in the inner leaflet! This observation will become important later on when we discuss programmed cell death. A dying cell will expose PS in the outer leaflet. This is one of the markers of a dying cell. Table **12.2.2** below shows the membrane lipid composition in an average mammalian cell.

Table **12.2.2**: Membrane lipid composition in an average mammalian cell.

Lipid	%
PC	45-55
PE	15-25
PI	10-15
PS	5-10
PA	1-2
SM	5-10
cardiolipin (bis-PG)	2-5
cholesterol	10-20

Lipid membranes also surround a variety of intracellular organelles found in eukaryotic cells. As a refresher, Figure 12.2.2 shows the anatomy of a typical eukaryotic cell with its variety of intracellular organelles.

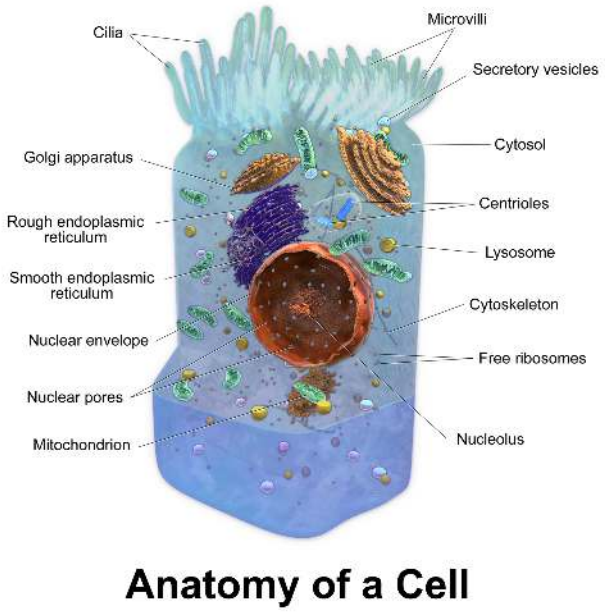


Figure 12.2.2: The Anatomy of a eukaryotic cell. https://opm.phar.umich.edu/biological_membranes https://creativecommons.org/licenses/by/3.0/

Figure 12.2.3 shows the average distribution of membrane lipids in different eukaryotic organelles.

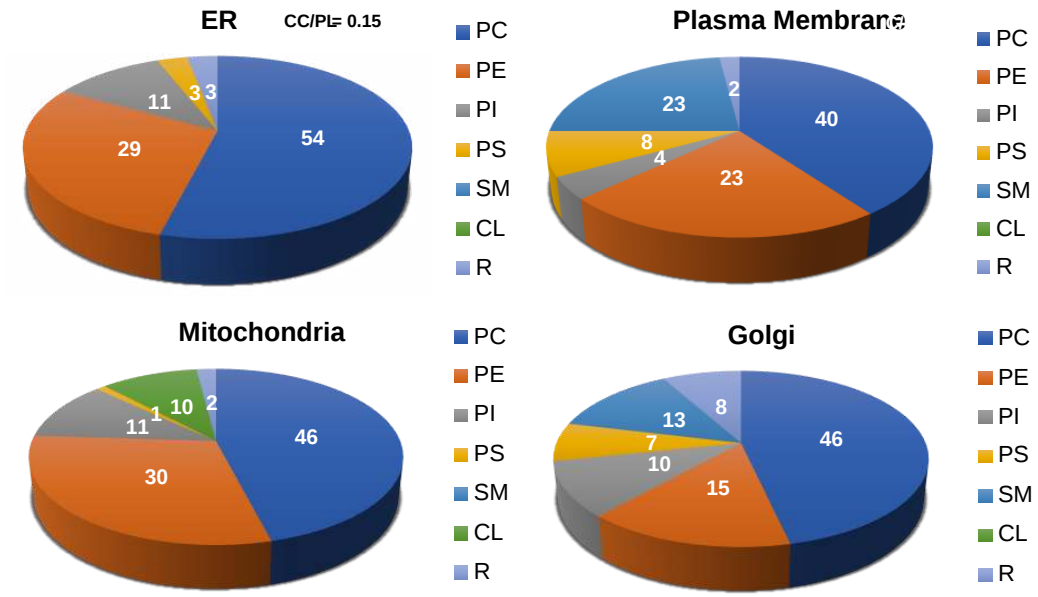


Figure 12.2.3: Average distribution of membrane lipids in different eukaryotic organelles

Figure 12.2.4, our last overview, shows the location of lipid synthesis and the resulting distribution of lipids in each leaflet. Note that most lipids are synthesized in the endoplasmic reticulum (ER).

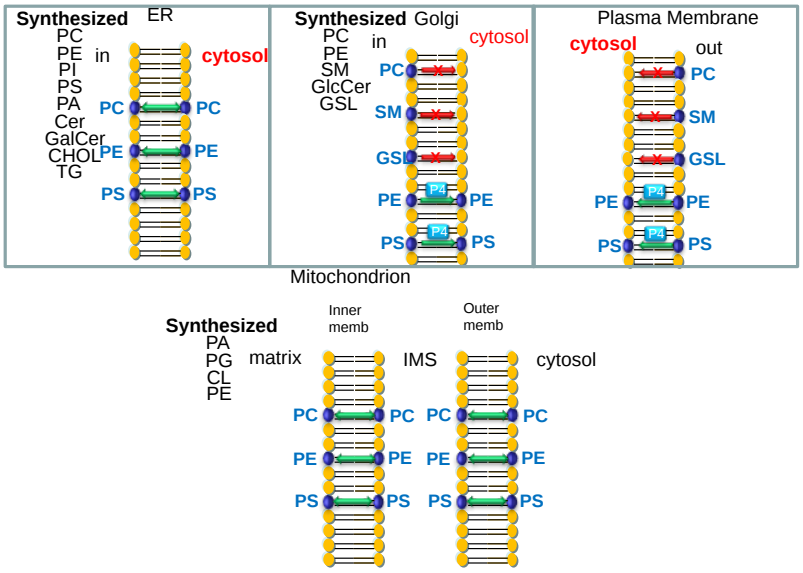


Figure 12.2.4: Location of lipid synthesis and the resulting distribution of lipids in each leaflet.

12.2.2: Dynamics of Membrane Bilayers

Molecules are not static but rather are dynamic. This also applies to molecular aggregates. In the first part of the section, we will discuss the rigid movement of whole lipid molecules in a bilayer, within a leaflet, and between leaflets. In the second part and the following supplement, we will consider the movement of atoms within a molecule. The movements include motions like bond bending, bond stretching, and torsion angle changes, as we saw in the previous chapter on the conformations of n-butane. The position of all atoms within a molecule can be simulated as a function of time - a molecular dynamics simulation. Such motions affect the molecule's energy, which can be calculated using classical molecular mechanics and electrostatics for given atom positions.

Liposomes and bilayers in general must be somewhat dynamic; otherwise, they would be impenetrable barriers across which nothing could pass. Cell membranes must separate the outside of a cell from the inside, but they must also allow the passage of

molecules and ions across the membrane. What is the evidence that membranes are dynamic?

First, lipids can diffuse laterally in the membrane. This can be shown as follows. Make a liposome from phosphatidylethanolamine, PE, which has been labeled with TNBS (trinitrobenzene sulfonate). The NH_2 on the head group of PE can attach the TNBS, which undergoes nucleophilic aromatic substitution with the expulsion of the SO_3^{2-}. The TNB group attached to the PE head group absorbs UV light and emits light of a higher wavelength in a process called fluorescence. Next, the fluorescence intensity of a surface region can be recorded using a fluorescent microscope. Then, a laser shines on a small area of the surface, which can photobleach the fluorescence in the area. Over time, fluorescence can be detected from the region again. The rate at which it returns measures the lateral diffusion of the labeled lipids into the region. Lipids can undergo lateral diffusion at a rate of about 2 mm/s. This implies that the lipids can transit the surface of a bacterium in 1 sec.

Transverse or flip-flop diffusion (movement of a phospholipid from one leaflet to the other, not within a given leaflet) should be more difficult. Experimentally, this is investigated as shown in the diagrams below.

Flip-Flop Diffusion in Liposomes: To test flip-flop diffusion in an artificial membrane, liposomes are made with a mixture of PC and a PC derivative with a nitroxide spin label (has a single unpaired electron) as shown in Figure 12.2.5.

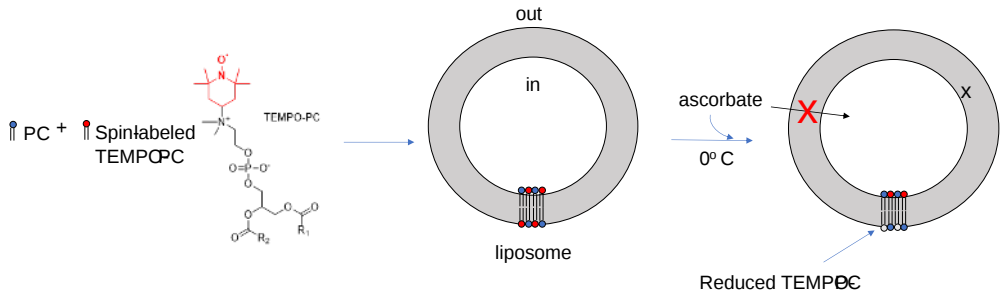


Figure 12.2.5: Flip-flop diffusion in vesicles

Both the inner and outer leaflets of the membrane have the labeled PC. Like a proton in NMR spectroscopy, a single electron has a spin that can give rise to an electron-spin resonance (ESR) signal (as a proton gives rise to a nuclear magnetic resonance signal) when irradiated with the appropriate frequency electromagnetic radiation (microwave frequency for ESR, radio frequency for NMR) in the presence of a magnetic field. The ESR signal is determined. Ascorbic acid, a water-soluble vitamin and antioxidant/reducing agent, is added to the liposomes. This reduces the spin-labeled PC in the outer leaflet but not the inner leaflet of the bilayer since ascorbic acid can not enter the liposome or otherwise interact with it. This reduces the ESR signal to a lower, constant value.

The sample is divided into two. One sample is left at 0 ^{o}C, while the other is raised to 30 ^{o}C. The ESR signal is recorded as a function of time. The 0 ^{o}C prep shows no change in ESR with time, while the 30 ^{o}C prep ESR signal decreases with time. This decrease results from flip-flop diffusion of labeled PC from the inner leaflet to the outer and subsequent reduction by ascorbic acid. These experiments in experimental bilayer systems like liposomes show that flip-flop diffusion is orders of magnitude slower than lateral diffusion.

Flip-Flop Diffusion in Bacterial Cells

An analogous experiment can be done with bacteria. Radiolabeled $^{32}PO_4^-$ is added to cells for one minute, which leads to the labeling of newly synthesized phospholipid (PL) located in the inner leaflet. The cells are then split into two samples. One sample is reacted immediately with TNBS, which will label only PE in the outer leaflet. The other sample is incubated for 3 minutes (to allow PL synthesis) and then reacted with TNBS. This is shown in Figure 12.2.6.

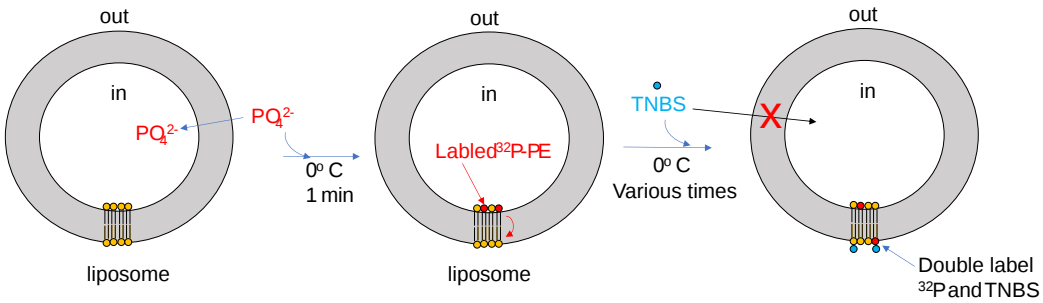


Figure 12.2.6: Flip-Flop Diffusion in Bacterial Cells

After a short labeling period, the cells are destroyed by adding organic solvents, which prevents the biosynthesis of new lipids. The lipids are extracted into the solvent and then subjected to TLC.

The lipids can be labeled in three ways. Some will be labeled with ^{32}P alone, some with TNBS alone, and some with both ^{32}P and TNBS. TLC (or other techniques such as HPLC or GC) can easily separate PC and TNBS-labeled PC since they have different structures and will migrate to different places on a TLC plate. No chromatographic technique could separate PC and ^{32}P-PC since their molecular structure is the same, the only difference being in the P nuclei (different number of neutrons).

Those lipids with double labels (TNB and ^{32}P) must have flipped from the inner leaflet to the outer leaflet, where they could be labeled with TNBS. The cells incubated for 3 minutes before the addition of TNBS have a much higher level of doubly labeled PLs. Quantifying these data as a function of differing times of incubation at elevated temperatures shows that the rate of flip-flop diffusion is much higher in cells than in liposomes, which suggests that the process is catalyzed, presumably by a protein transporter (flippase or Transbilayer amphipath transporter - TAT) in cells.

Different enzymes catalyze the movement of lipids in bilayers, including flippases, floppases, and scramblases, as illustrated in Figure 12.2.7. Most require ATP hydrolysis for the physical movement of the lipid across leaflets.

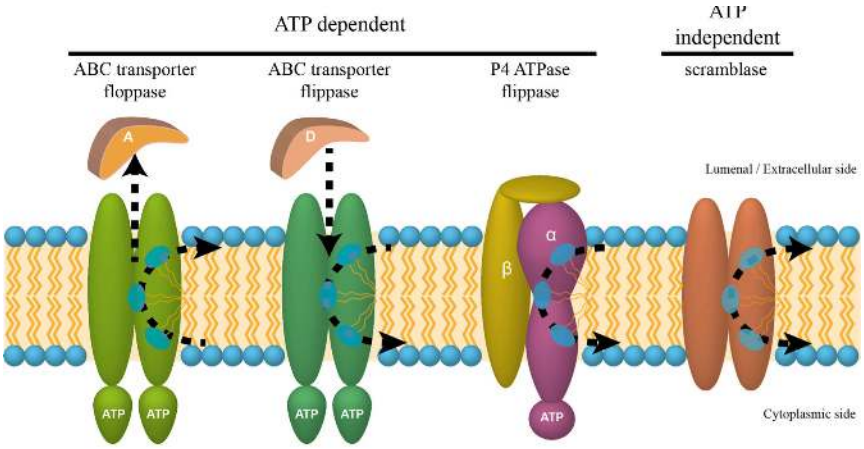


Figure 12.2.7: Enzymes that move lipids between leaflets. Juliana Rizzo et al. https://doi.org/10.1016/j.csbj.2019.09.001. Creative Commons: https://creativecommons.org/licenses/by/4.0/

Figure 12.2.8 shows an interactive iCn3D model of a human plasma membrane phospholipid flippase with bound phosphatidylserine (PS) shown in spacefill.

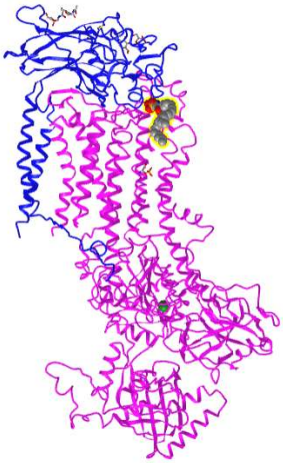


Figure 12.2.8: Human plasma membrane phospholipid flippase with bound phosphatidylserine (6lkn) (Copyright; author via source).

Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...kcJeKtZGDa53W6

The iCn3D model below shows the structure of a human plasma membrane phospholipid flippase with bound phosphatidylserine (PS) shown in spacefill. This protein moves PS from the outer to the inner plasma membrane leaflet, maintaining its asymmetric distribution. The other common aminophospholipid, PE, is predominantly found in the inner leaflet. In contrast, PC and SM are found predominantly in the outer leaflet. The movement of PS is from a low to a high concentration and requires ATP. A build-up of PS in the outer leaflet is one signal that initiates programmed cell death (apoptosis) in the cell. Clotting is also initiated when cellular damage leads to exposed PS.

Flippases are proteins that move lipids from the outer to the inner leaflet, while floppases move them from the inner to the outer leaflet. Most are also ATPases. Both promote lipid asymmetry in the membrane, and floppases also help move lipids out of the cell. Scramblases move lipids in either direction and break the asymmetry of the lipid distribution. They are important in signaling. For example, they are used to expose phosphatidylserine to the outer leaflet, which promotes programmed cell death.

Here are links to iCn3D models of

- a floppase MsbA (6BPP)
- scramblase (6P49 AND 7JLP)

12.2.3: Conformational Transitions in Bilayers

Suppose a vesicle preparation is placed in a sensitive calorimeter and the temperature is increased. The vesicle preparation absorbs a significant amount of heat at a temperature characteristic of the phospholipids that compose the vesicle. This is analogous to what would happen if solid water were heated. At the melting point of water, an increment of heat is required, the heat of fusion, to break H-bonds and cause melting. Likewise, the heat of vaporization is measured when H-bonds are broken on the liquid-gas transition. These transitions are associated with non-covalent processes, namely, breaking H-bonds. Graphs of heat absorbed (Q) as a function of temperature, or heat absorbed/T (i.e., the heat capacity) vs temperature for the melting and evaporation of water, are shown in Figure 12.2.9. These transitions occur at the melting point (T_M) and the boiling point.

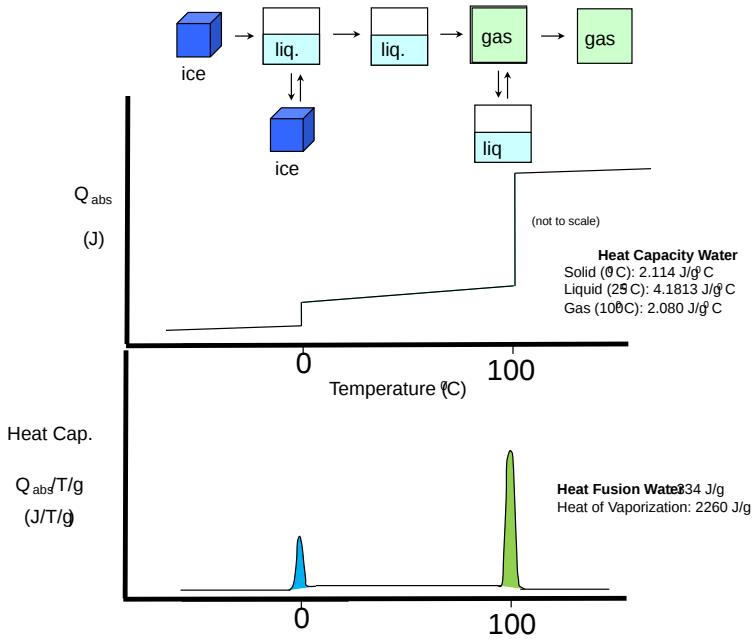


Figure 12.2.9: heat absorbed (Q) vs temperature (top) and heat absorbed/T (heat capacity) vs temperature for the melting and evaporation of water

The bottom heat capacity graph is nothing more than the derivative curve (or slope at each point) of the Q_{abs} curve!

Likewise, lipid vesicles undergo phase transitions comparable to the melting of water. One significant phase transition at a "melting point" (T_M) = 42 ^{0}C can be seen in the graph of heat capacity vs temperature for vesicles made of DPPC, shown in Figure 12.2.10.

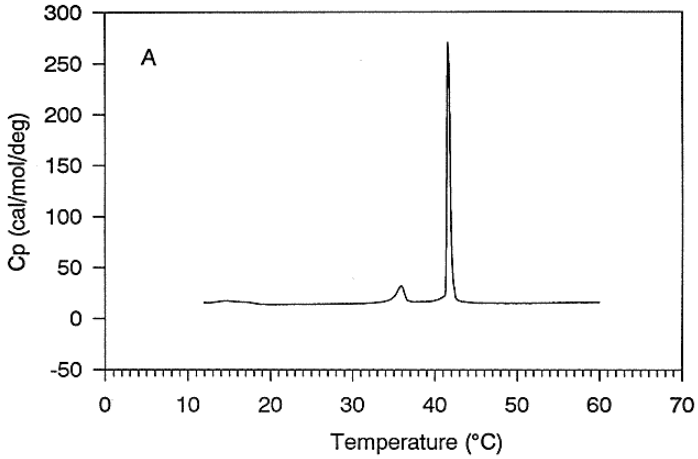


Figure 12.2.10: Heat capacity vs temperature for vesicles made of DPPC

This transition is caused by conformational chains in the packing of the acyl chains of the phospholipids as the acyl chains change from trans to gauche conformations. These changes involve not the simple translation of lipid molecules within and between bilayers but rather the movement of atoms within the molecules. These kinds of motions can be modeled using molecular dynamics simulations. Before the transition, the acyl chains are more tightly packed in the **gel phase**, and after the transition, they are less tightly packed in the **liquid crystalline phase** since many chains are in the gauche conformation. A minor transition is also noted at around 36 ^{0}C. This is associated with changes in the orientation of head groups.

As with water going from ice to liquid, the vesicles remain intact after the phase transition. It's not like the transition of liquid water to the gas phase. Vesicles in the liquid crystalline phase are more fluid, dynamic, and permeable. Note that the liposomes have not been destroyed but have undergone a phase change, similar to ice melting into liquid water.

The lipid vesicle phases are named gel and liquid crystalline to reflect the rigidity of the bilayer.

- **Gel phase** (L_β): In the gel phase, which is found at temperatures < T_M, the lipids are ordered with maximal packing. The acyl chains in both leaflets can be tilted so that they align in a parallel fashion (as shown in the figure below) or in a cross-tilted fashion in which they tilt toward each other. In the gel phase, the lipids diffuse slowly. This phase is sometimes called the solid phase. The gel phase is favored by low temperature and high saturation of esterified fatty acids. Saturated PC bilayers give a gel phase in the lab,
- **Liquid crystalline phase** (L_α): In this liquid crystalline phase, which is found at temperatures> T_M, some saturated acyl chains have undergone all-trans to gauche conformational changes. These introduce kinks into the chains, which reduce packing. The notation L_α is used for bilayers of pure lipids.
- **Liquid crystalline ordered** (L_0) and **Liquid crystalline disordered** (L_d): These phases typically occur with the addition of relatively high amounts of cholesterol. Cholesterol modulates the fluidity of membranes, as we will see in a bit, and affects bilayer properties at temperatures both < and > TM. The L_0 phase is often enriched in saturated (sphingo)lipids and cholesterol, while the L_d phase is often enriched in unsaturated glycerophospholipids. The **liquid crystalline disordered** (L_d) has fast translational diffusion and lower order, while the **Liquid crystalline ordered** (L_0) has fast diffusion with higher order.

Most membrane lipids *in vivo* contain unsaturated fatty acids, and specific lipids are used in given environments to avoid the gel phase. Figure 12.2.11 shows some of these phases

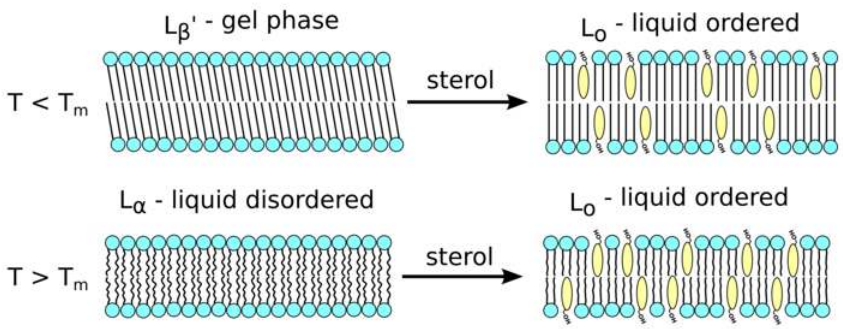


Figure 12.2.11: Common phases in membrane bilayers https://phys.libretexts.org/Courses/...se_Transitions

Figure 12.2.12 shows a snapshot of a molecular dynamics simulation of a bilayer in a gel (A) and liquid crystalline (B) phase. Note that the width of the liquid crystalline phase is smaller.

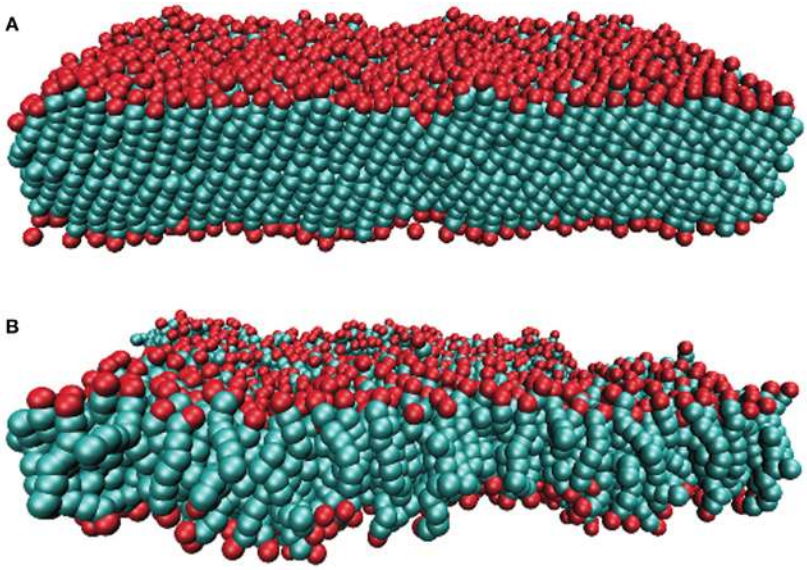


Figure 12.2.12: Snapshot of a molecular dynamics simulation of a bilayer in a gel (A) and liquid crystalline (B) phase. Tayebeh Jadidi et al. Bioeng. Biotechnol., 22 April 2014 | https://doi.org/10.3389/fbioe.2014.00008. **Creative Commons Attribution License (CC BY)**.

Vesicles of different PL have different T_M as shown in Table 12.2.3 below.

Table 12.2.3: Melting point (TM) of vesicles made with different phospholipids

Lipid	T_M	Lipid	T_M
12:0 PC	-1	12:0 PA	31
14:0 PC	23	14:0 PA	50
16:0 PC	41	16:0 PA	67
18:0 PC	55	18:0 PA	76
18:1 PC	-20	18:1 PA	-8
18:2 PC	-53	-	-
18:3 PC	-60	-	-

Vesicles made from phospholipids with bigger head groups have a lower TM since they are less "stable". For example, the Tm for vesicles of di-16:0 versions of PA, PE, and PC have T_Ms of 67, 63, and 41 degrees C, respectively, as shown in Figure **12.2.13**.

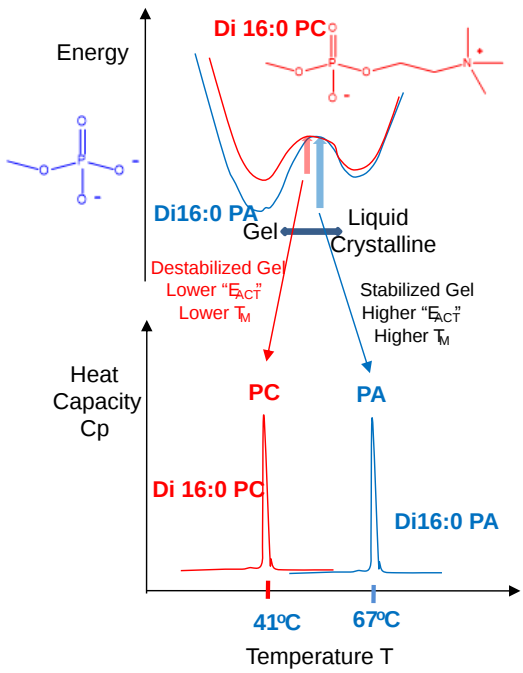


Figure **12.2.13**: Melting point (T_M) of 16:0 vesicles with different head groups

12.2.4: Cholesterol and Membrane Fluidity

Cholesterol is also a ubiquitous component of animal cell membranes. Its size will allow it to fit into either leaflet with its polar OH pointed to the outside. One function of cholesterol in membranes is to keep the membrane fluid at any reasonable temperature. When a membrane is at a temperature less than the T_M, it is ordinarily in a gel, not a liquid crystalline phase. The cholesterol helps prevent the ordered packing of the acyl chains of the PLs, which increases their freedom of motion. Hence, the fluidity and permeability of the membrane are increased. At temperatures greater than the T_M, the rigid ring of cholesterol reduces the freedom of the acyl chains to rotate. Hence, it decreases the number of chains in the gauche conformation. This decreases fluidity and permeability. Cholesterol affects membrane structure at temperatures both below and above the T_M as

Figure **12.2.14** shows the results of a molecular dynamics simulation depicting the relative order in a DMPC membrane with and without cholesterol

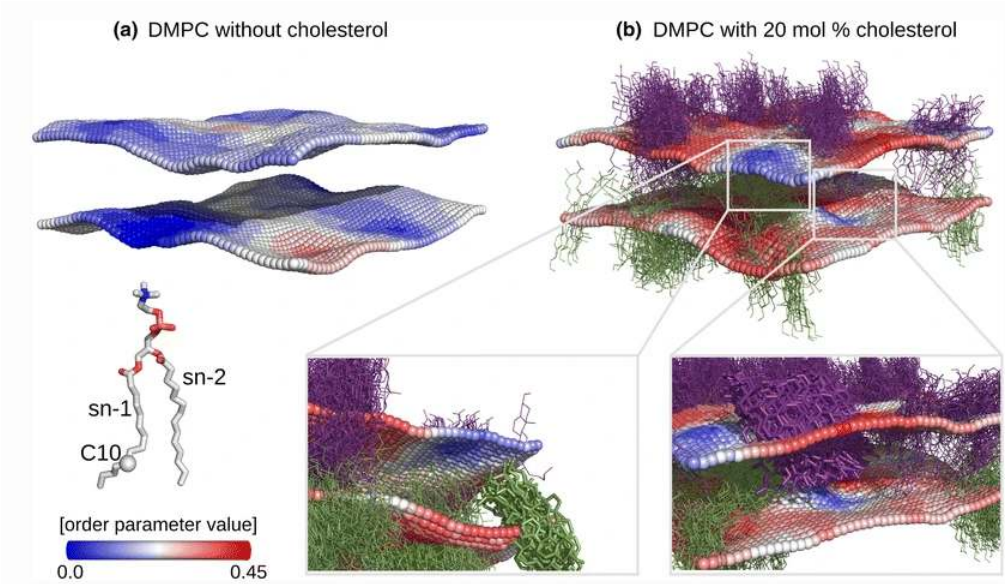


Figure 12.2.14: "S_{CD} order parameter analysis of cholesterol-enriched DMPC bilayer. Order parameter values for the C10 carbon of the acyl chain sn-1 are shown. The box area was 5.6 × 5.6 nm^2 and 50 bins were used along the *x* and *y* axes. **a** Local deuterium order parameters of a pure DMPC membrane averaged over a 100 ns MD trajectory. **b** Local S_{CD} order parameters for a 20 mol% DMPC bilayer averaged over a 10 ns MD trajectory excerpt. The cholesterol molecules represented as *lines* are overlayed for all the frames. The enlarged view figures emphasize the areas where cholesterol (visualized as sticks) decreases ordering of the lipid chains on the opposite leaflet" (from below ref)

12.2.5: Lipid Rafts and Nanodomains

Not only are lipids asymmetrically distributed between leaflets of a bilayer, but they are also distributed asymmetrically within a single leaflet. Certain lipids often cluster within a leaflet to form lipid "rafts," which can be considered to result from lateral phase separation of the lipids within one leaflet of the bilayer. Divalent cations like calcium, which can bind to negatively charged PLs like PS, can cause "rafts" of PS to form, giving rise to lateral asymmetry within a bilayer leaflet. Rafts also appear to be enriched in cholesterol and lipids with saturated fatty acids, especially sphingolipids, leading to regions of enhanced packing and reduced fluidity. Cholesterol would stabilize packing in spaces created with lipids with large head groups. You can think of these rafts as nanodomains, analogous to the domains we observed in protein structure.

Cholesterol appears to be a key player in the formation of lipid rafts. It is planar and inflexible and would pack better with saturated fatty acid chains. It could also induce them to elongate to form lower energy zig-zag structures in which all the methylene groups are anti. Lipid rafts represent a more ordered lipid phase (Lo) than the more disordered surrounding phase (Ld). Also, compared to the structure of glycerophospholipids, the atoms in the region linking the head group and the nonpolar fatty acid chains in sphingolipids have greater potential for H-bond interactions with cholesterol and other sphingolipids, as shown in Figure 12.2.15.

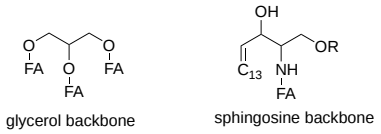


Figure 12.2.15: Comparison of glycerol and sphingosine headgroup structure. FAs are fatty acids esterified (glycerol) or amide-linked (sphingosine)

Rafts probably bind or exclude the binding of other biological molecules, like proteins. Some proteins are chemically modified with a glycosylphosphoinositol (GPI) group at the carboxy terminus. The PI group can insert into the membrane, anchoring the protein to the bilayer. Proteins also appear to induce raft formation. Lipid rafts appear to be enriched in glycosylphosphoinositol (GPI)-anchored proteins, as seen in Chapter 11.1. Recent studies have shown that the Ebola virus interacts with lipid rafts while entering and exiting the infected cell. Rafts are also involved in how cells sense and respond to their environment. Signaling molecules outside the cell can bind receptor proteins in the membrane. As we will see later, conformational changes in the receptor protein signal the inside of the cells that the receptor is bound with a ligand. Once bound, the receptor can move in the membrane and often cluster in outer leaflet rafts that contain cholesterol and sphingolipids. Inner leaflet rafts have also been observed. Figure

12.2.16 shows two versions of an animated lipid raft. The large shapes represent membrane proteins selectively found in the rafts. The most modern definition of a lipid raft is a nanoscale assembly of sphingolipids, cholesterol, and proteins that can be stabilized into platforms.

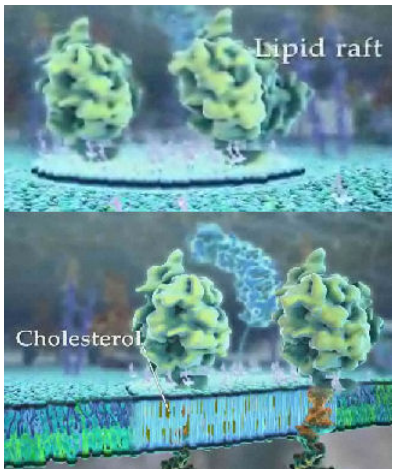


Figure 12.2.16: Lipid Rafts enriched in SM and Cholesterol. (screen capture from: http://multimedia.mcb.harvard.edu/anim_innerlife.html)

Figure 12.2.17 shows a simplified lipid raft model.

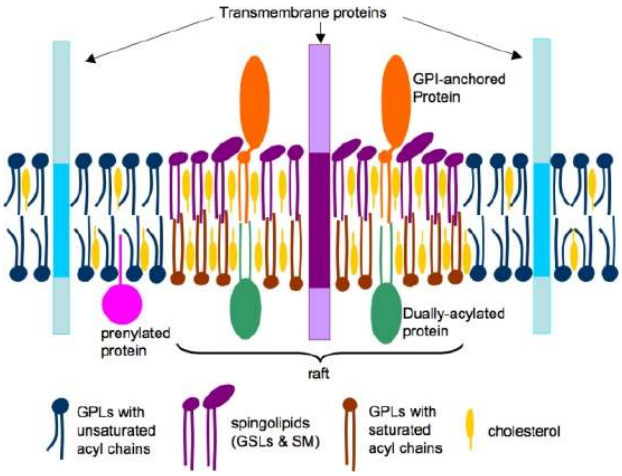


Figure 12.2.17: A simplified model of lipid rafts in cell membranes. https://www.researchgate.net/figure/..._fig4_45952334. license: https://creativecommons.org/licenses/by/3.0/

"The phospholipids (blue and brown) and cholesterol (yellow) are distributed in both the leaflets, whereas sphingolipids (violet) are enriched in the outer leaflet of the bilayer. The acyl chains of raft lipids are generally long and saturated (violet and brown). In contrast, those in non-raft domains are shorter and contain singly or multiply unsaturated acyl chains (blue). Raft domains contain concentrations of dually-acylated (green) and GPI-anchored (brown) proteins, whereas transmembrane (blue) and prenylated (green) proteins are usually non-raft associated."

In contrast to single proteins, for example, lipid bilayers are physical mixtures. In a thermodynamic sense, entropy would disfavor raft formation as random mixing is favored. Hence enthalpy must drive the interaction between neighboring molecules to produce nanodomains and rafts.

12.2.6: Lipid Phase Diagrams

You are familiar with the phase diagrams of water from introductory chemistry classes. Phase diagrams show the phases accessible under different conditions, such as temperature and pressure. Figure 12.2.18 shows a traditional phase diagram for water.

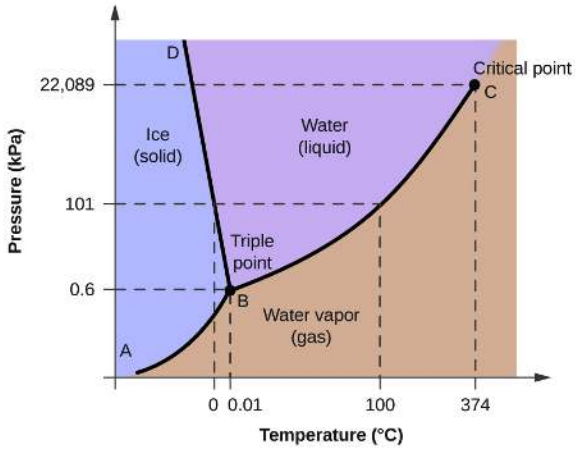


Figure 12.2.18: A traditional phase diagram for water. https://opentextbc.ca/chemistry/chap...hase-diagrams/. Chemistry by Rice University is licensed under a Creative Commons Attribution 4.0 International License.

The horizontal dotted line at 101 kPa shows the states of water as a function of temperature at 101 kPa = 1 atm pressure. The phase transition of solid to liquid water occurs at 0 ^{0}C (freezing/melting point of water), while the liquid to gas transition occurs at 100 ^{0}C (boiling point of water). At a reduced pressure (0.6 atm), all three water phases can exist at 0.01 ^{0}C, the triple point of water.

Analogously, lipid bilayers have phase transition diagrams as well. Instead of showing phases as a function of temperature and pressure, they are usually shown as a function of temperature and concentration of a specific lipid component such as cholesterol, which, as described above, affects T_M, fluidity, and raft formation. An example of a **theoretical** phase diagram for membranes composed of saturated dipalmitoylphosphatidylcholine (16:0), the most common saturated fatty acid in animals, plants, and microorganisms, vs. cholesterol content is shown in Figure 12.2.19. This is a binary DPPC–cholesterol bilayer system.

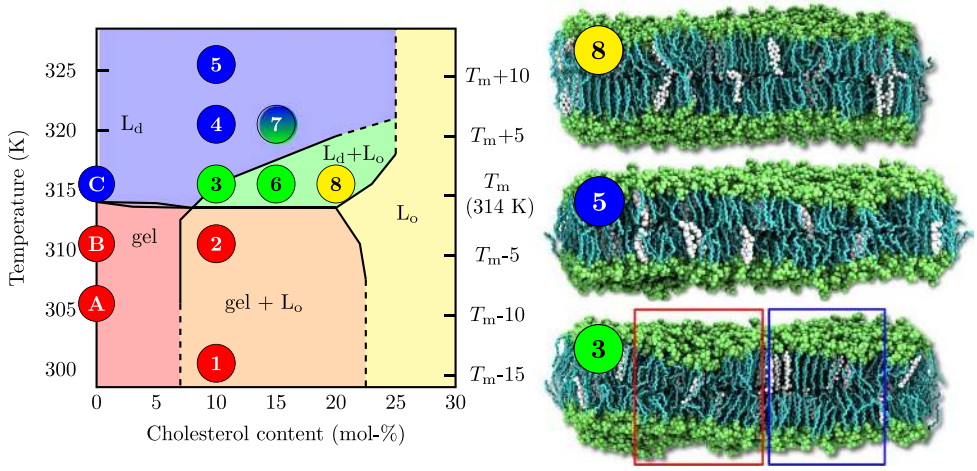


Figure 12.2.19: Phase diagram with phase boundary lines. Javanainen, M., Martinez-Seara, H. & Vattulainen, I. Nanoscale Membrane Domain Formation Driven by Cholesterol. *Sci Rep* **7,** 1143 (2017). https://doi.org/10.1038/s41598-017-01247-9. **https://rdcu.be/chbtH.** http://creativecommons.org/licenses/by/4.0/.

Left Panel: Liquid–liquid coexistence is expected in the L_d + L_o regime (green). The L_o, L_d, gel, and gel + L_o regions are colored in yellow, blue, red, and orange, respectively. The locations of the simulated systems are shown by numbers (DPPC–cholesterol) and letters (pure DPPC). Their colors indicate their phases. Systems 3, 6, and 7 show heterogeneous behavior.

Right Panel: Snapshots of selected systems labeled by the points in the phase diagram: (8) Chol20 316 (L_o phase), Chol10 326 (L_d phase), Chol10 316 (ordered/disordered/hexagonal), DPPC is shown in cyan and lime (chains and other parts), and cholesterol in white. Water, ions, and lipid chain hydrogens are omitted for clarity. Red and blue boxes highlight disordered and ordered regions, respectively.

DPPC vesicles have a T_M = 41 ^{0}C or 314 K. The phase diagrams show that when cholesterol is added to DPPC bilayers at temperatures above T_M, the bilayers change from the liquid-disordered (Ld) phase to the uniform liquid-ordered (Lo) phase at around 20 mol% cholesterol. At cholesterol concentrations between around 10-20 mol% and temperatures just above the T_M, the L_0

and L_d phases coexist. The coexistence of two phases mimics a raft. At low cholesterol levels, the bilayer changes from the gel (or L_β) to the L_d phase. At really high cholesterol, only the L_0 phase exists. This makes sense as the fatty acids are all saturated. Cholesterol's rigid rings reduce the freedom of the acyl chains to rotation and hence decrease the number of chains in the gauche conformation. Yet, as an "impurity" (cholesterol) has been added to the system, the system is less rigid and more fluid-like. From 0-7 mol% cholesterol, two phases exist: the gel (or L_β) and the L_d phase. Between 7-23 mol% cholesterol, a combination of phases is seen.

As mentioned above, the outer leaflet of mammalian plasma membranes is composed mostly of sphingomyelin (SM) and PC. In contrast, the inner leaflet is composed mostly of phosphatidyl ethanolamine (PE), phosphatidyl serine (PS), and cholesterol. These don't appear to separate into L_d and L_0 phases and appear not to form nanodomains or rafts. In the lab, outer membrane lipids easily form vesicles, but the inner leaflet polyunsaturated PEs form other phases (hexagonal or cubic). It appears that asymmetric lipid distribution is critical for biological bilayer formation.

Since nanodomains are challenging to observe and study experimentally, molecular dynamics simulations provide insight into their structure and properties. The right-hand images show snapshots of the membrane from molecular dynamics simulations. Cholesterol is shown in white in Figure 12.2.19. **8** shows the membrane in the Lo phase, **5** shows it in the Ld phase (note the reduced membrane width), and **3** in an ordered (blue box) and disordered (red box) phase. The simulations support the theoretical phase diagram showing the coexistence of the L_d and L_0 phases.

Also, a hexagonal-closest packed (also cholesterol-poor) region occurs in phase diagram regions showing the coexistence of both L_d and L_0 phases. Cholesterol is excluded from most ordered regions. This is shown in Figure 12.2.20, showing a top-down view of the membrane. The dark blue lipid headgroups are hexagonally closest packed, an ideal you will remember from introductory chemistry courses. Cholesterol (yellow spheres) seems to be most prevalent in boundary regions.

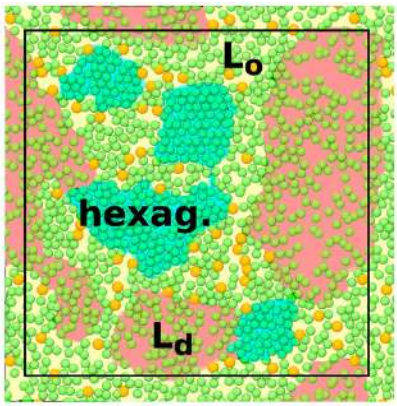


Figure 12.2.20: Top-down view showing membrane phase observed in a molecular dynamics simulation. Javanainen, M. et al. ibid. DPPC chains are shown in green/blue spheres and cholesterol in orange. The estimated regions of lipids displaying L_o-like, L_d-like, and hexagonally packed regions are shaded with yellow, red, and blue, respectively.

Imagine a phase diagram for a three-component system (DOPC, sphingomyelin, and cholesterol) or even higher component systems! We won't show any, but from the "simple" two-component system described above, it should be evident that we have a long way to go before understanding the complexity of membrane bilayers.

The actual biological membrane must be able to adopt very nonplanar shapes with positive and negative curvature. Membranes must also be able to fuse (for example, the egg and sperm). Another lipid phase, which we have not yet discussed, may be involved. New lipid phases of a single membrane lipid can form based on the relative percentages of lipid and water. These include hexagonal phases. Figure 12.2.21 shows an image of a hexagonal phase of phosphatidylethanolamine with 16:0 fatty acids. Note the water inside the middle ring of PE molecules.

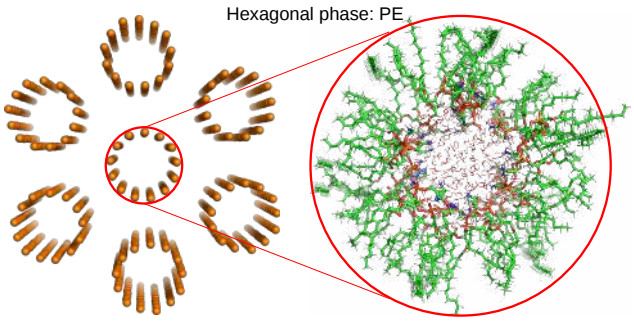
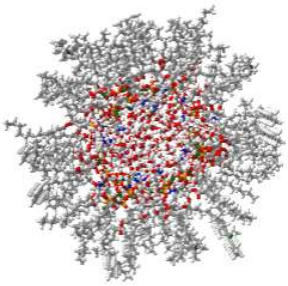


Figure 12.2.21: Hexagonal phase of phosphatidylethanolamine with 16:0 fatty acids

Figure 12.2.22 shows an interactive iCn3D model of the inner ring outlined in the red circle. Note that water surrounds each of the closest-packed lipid tubules, which extend back into the figure in 3D. This phase creates an aqueous channel through the interior of each tubule.



NCBI iCn3D Figure 12.2.22: Hexagonal phase of phosphatidylethanolamine with 16:0 fatty acids (Copyright; author via source). Click the image for a popup or use this external link: not available

Figure 12.2.23 shows variants of the hexagonal phase and a cubic phase of lipids.

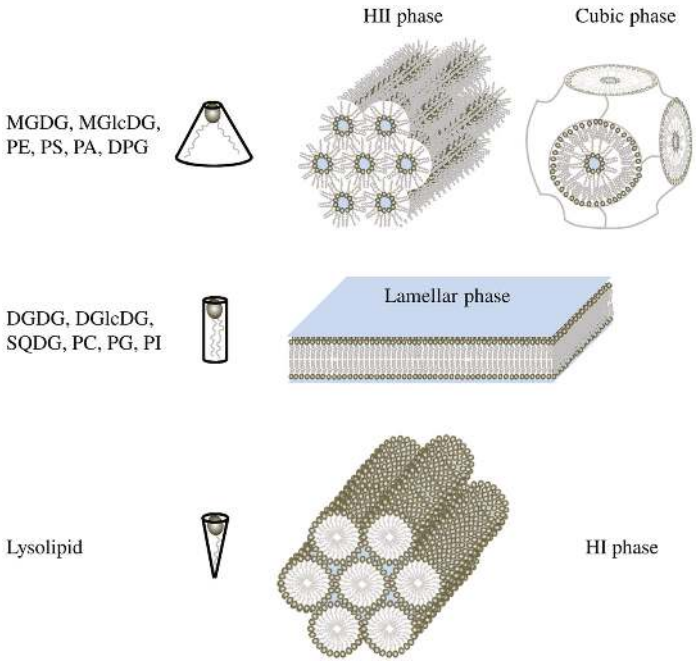


Figure 12.2.23: Hexagonal and cubic phases of lipids. Juliette Jouhet. Importance of the hexagonal lipid phase in biological membrane organization. Frontiers in Plant Science, Frontiers, 2013, 4, pp.494. ff10.3389/fpls.2013.00494ff. ffhal-00942927f. Creative Commons.

Double-chain amphiphiles with a small polar head (like PE and PA) are more likely to form a hexagonal II phase with elongated tubules than are more cylindrical lipids like PC. Note that if a single fatty acid is removed from a double-chain amphiphile like PE, a narrow conical shape single-chain amphiphile arises, forming either a micelle (not shown) or a hexagonal I phase. *In vitro*, equimolar amounts of PC (which forms the lamellar phase) and PE (which can form an HII phase) can either form the HII phase (at low aqueous pressure) or a lamellar phase (at high aqueous pressure).

Are hexagonal lipid phases found in biological membranes? The answer is likely yes in the formation of unusual cellular structures. The plant plasmodesmata, shown in Figure **12.2.24**, are one such structure. Just focus on the two bilayers connected by the membrane-lined membranes of the channel.

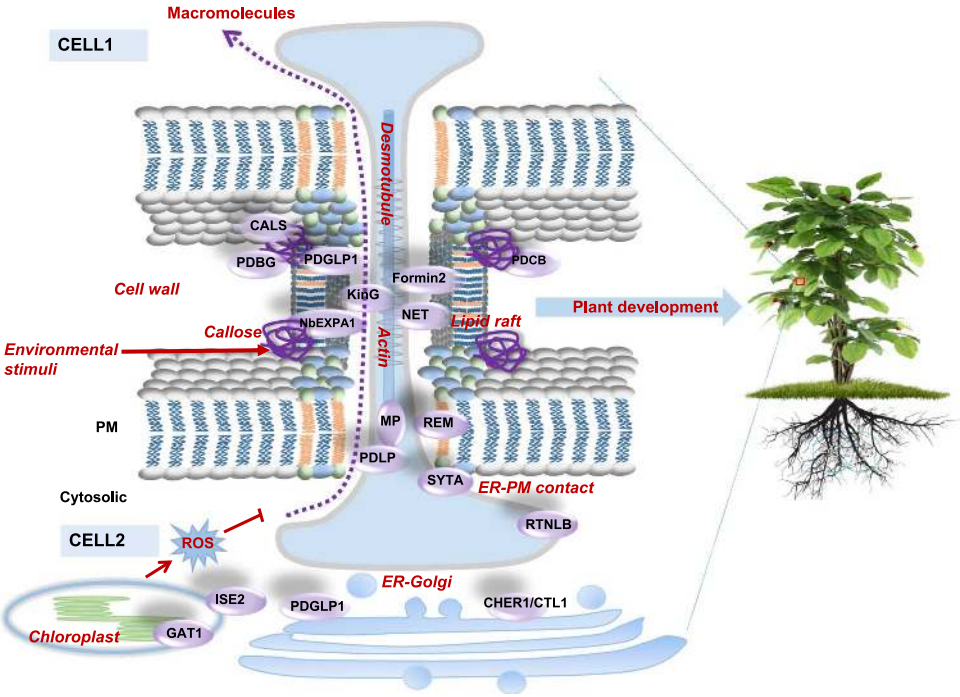


Figure **12.2.24**: Plant plasmademata. Yanbiao Sun et al. https://www.nature.com/articles/s41438-019-0129-3. Creative Commons license. http://creativecommons.org/licenses/by/4.0/.

Plasmodesmata are membrane-lined channels that cut across the plant cell wall and directly connect cells, allowing the flow of water and nutrients between the cells. HII phases have been seen in the endoplasmic reticulum membrane. These membranes, along with the mitochondrial inner membrane and the inner membrane of chloroplasts, are highly curved and contain higher concentrations of lipids that allow for curvature and the formation of H II phases. The function of some membrane enzymes and processes, such as membrane fusion and fission, is enhanced by HII-forming lipids.

12.2.7: Piecing it all together

Our emerging understanding of lipid structure has taken us from micelles and vesicles to the complexity of actual biological membranes with different phases and nanodomain structures. This complexity is needed as membranes must be dynamic in ways that the proteins, for example, aren't. They must be able to pinch off either as extracellular or intracellular vesicles, for example, in exocytosis and endocytosis. Both positive and negative curvatures of the membrane must be enabled. The incredible complexity of the "Lego-like" lipid monomers that assemble and rearrange into every fluctuating membrane is yet another exquisite illustration of our repeated mantra that structures and shapes mediate all functions. The "Lego-like" membrane monomers in various phases and regions of membrane curvature are illustrated in Figure **12.2.25**.

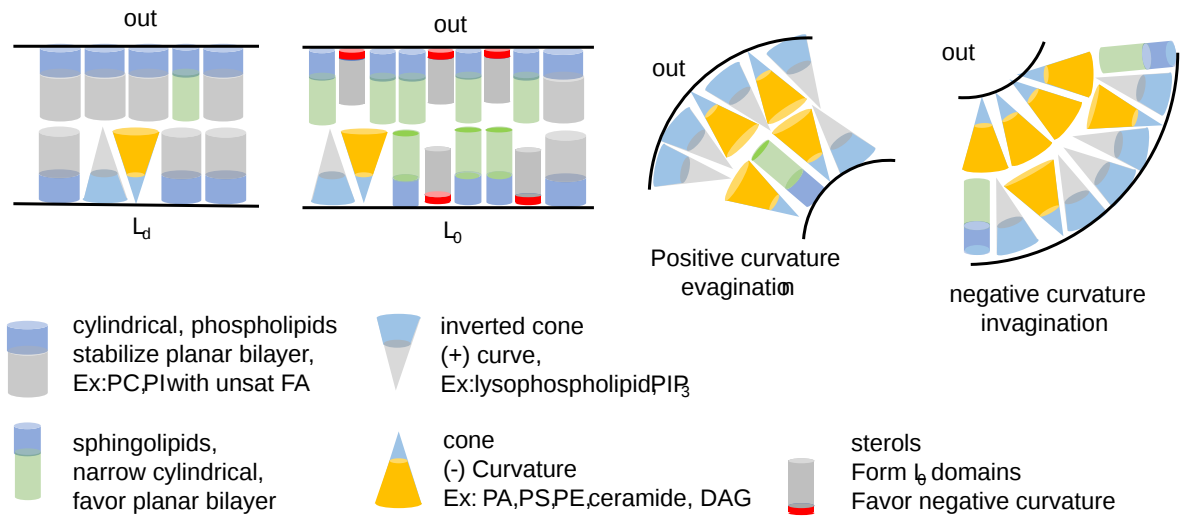


Figure 12.2.25: "Lego-like" membrane monomers in various phases and regions of membrane curvature

Given their shape, some glycerolipids don't even appear to form bilayers by themselves spontaneously. The shape and type of fatty acids found in the membrane lipids will determine the local properties of the membrane, its phases, and the presence of nanodomains. The activity of membrane enzymes and membrane fission and fusion events will also depend on the local properties of membranes. The figure above doesn't even account for the presence of peripheral and integral membrane proteins, which we will discuss in the next section.

As we learned with proteins, we can be misled by looking at beautiful but static images of proteins. Their dynamic motion is critical to their function. In addition to the dynamic membrane events discussed above, many other events take place at and in membranes. Here are a few.

12.2.8: Other dynamic events in bilayers

Membranes are not static. They are synthesized, their contents shuffled, fuse with other membranes, and large proteins are inserted into them. Let's explore some of these.

12.2.8.1: **Membrane Trafficking**

Movement of key "cargo" molecules into (endocytosis) and out of (exocytosis) the cells occurs mostly through membrane-encapsulated vesicles. Vesicles contain all types of biological molecules, including synthetic and absorbed lipids. Part of the differences in lipid composition between membrane layers and different organelles are derived from this highly orchestrated and controlled movement of vesicles. Details are shown in Figure 12.2.26below.

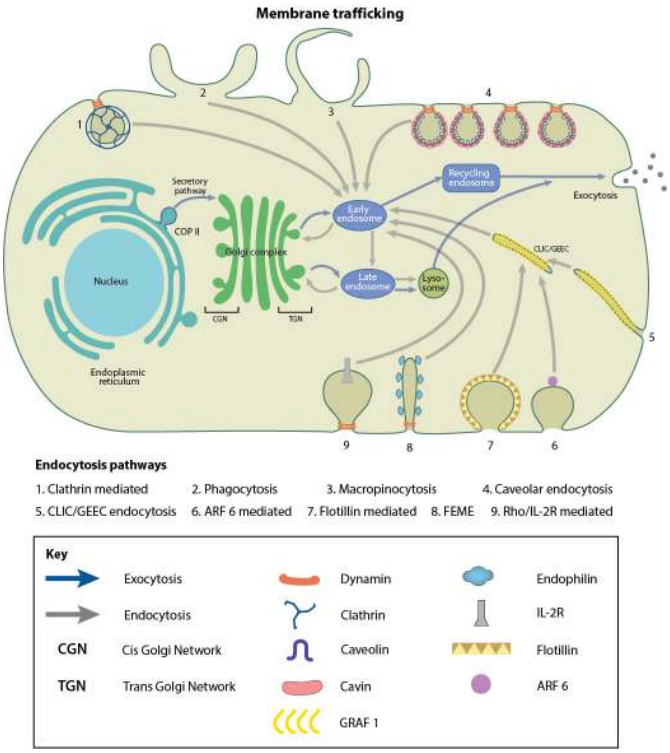


Figure 12.2.26: Players in membrane trafficking. https://www.mechanobio.info/what-is-...e-trafficking/. Creative Commons Attribution-NonCommercial 4.0 International License

In eukaryotes, the biosynthetic secretory pathways move molecules from the endoplasmic reticulum to the cis Golgi (CGN) to the trans-Golgi (TGN) and the plasma membrane (for integral membrane proteins) or for secretion. Since most lipids are synthesized in the endoplasmic reticulum (ER), their distribution to different cell locations is critical in maintaining the asymmetric distribution of lipids found in cells.

12.2.8.2: Fusion of membranes: Fusion Peptides

Another dynamic event in membranes is the fusion of two bilayers from two different cell vesicles or of a vesicle and a cell membrane. Fusion peptides facilitate these events. Figure (\PageIndex{27}\) below shows a molecular dynamics simulation snapshot showing how a fusion peptide in a single DMPC bilayer causes a constriction of the bilayer with the two leaflets approaching each other.

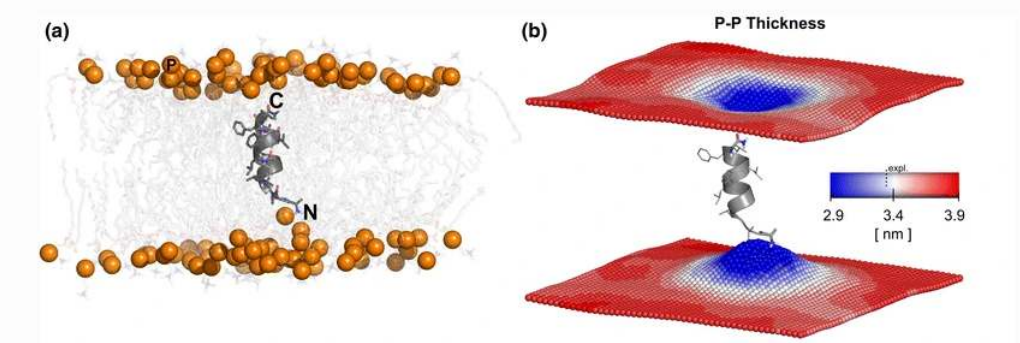


Figure (\PageIndex{27}\): A molecular dynamics simulation snapshot showing how a fusion peptide resides in a single bilayer. The first 12 amino acids of a GP41 peptide are inserted in a DMPC membrane patch.

Figure (\PageIndex{28}\) below shows the area per lipid (APL).

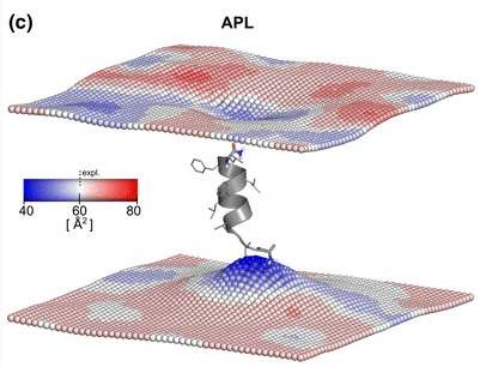


Figure (\PageIndex{28}\): Area per lipid (APL) of a DMPC membrane with an inserted fusion peptide.

12.2.8.3: Insertion of Membrane Proteins

Another dynamic event is the insertion of a membrane protein. Now let's look at changes in the bilayer on insertion of the voltage-dependent anion channel (VDAC) membrane protein, as shown in Figure (\PageIndex{29}\) below.

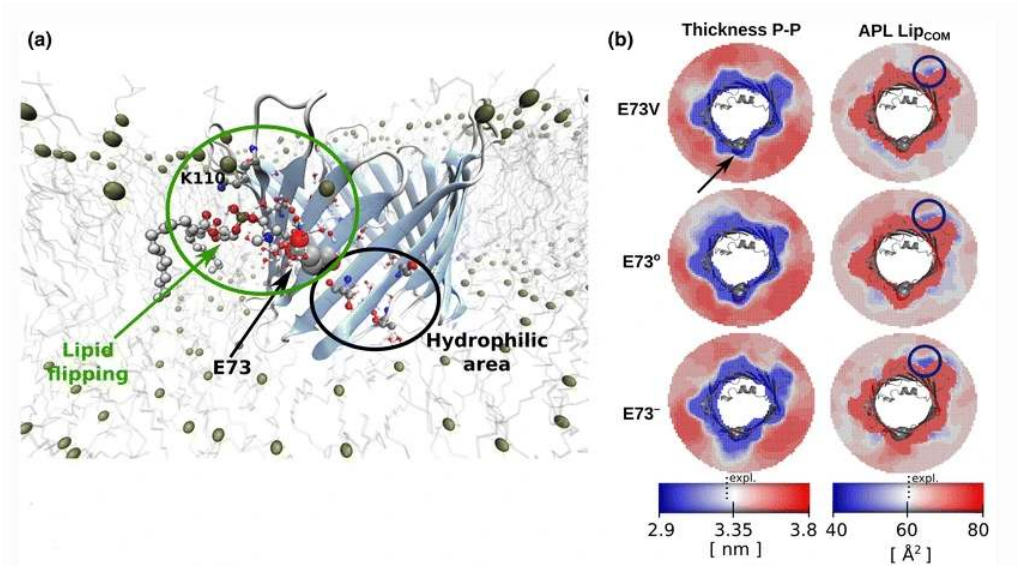


Figure (\PageIndex{29}\): Local membrane properties of a bilayer with an embedded membrane protein VDAC. CVytautas Gapsys et al, https://link.springer.com/article/10...822-013-9684-0. Creative Commons CC BY license.

Panel a shows the voltage-dependent anion channel (VDAC) in a cartoon β-barrel representation. Residue E73 and the water molecules nearby are represented as spheres. The ball and stick representation shows that serine and threonine residues constitute hydrophilic areas close to E73. The snapshot shows a DMPC lipid flipping close to the E73 and K110 residues.

Panel b shows top-view perspectives of the circular representation of the local thickness, calculated considering phosphorus atoms and area per lipid.

12.2.9: Introduction to Lipid Signaling - Chemical Cleavage of Membrane Lipids

Everything in this chapter so far describes the structure and dynamics of the lipid components of a bilayer. The dynamic changes described involve the physical movement of lipid molecules in the membrane. Let's briefly introduce another type of movement involving physical and chemical changes in the cell. What happens when specific lipids in membranes are chemically cleaved by **lipases**, enzymes that are analogous to proteases? It turns out these changes lead to signaling within the cell. We will only briefly introduce lipid signaling in this chapter, but explore it in more detail in Chapter 12.

Lipids are not just used as a passive component of membranes or as a source of stored energy. They are involved in signal transduction at the cell membrane, a process by which the interior components of the cell respond to a signal external to the cell, allowing the cell to respond to its local environment. Usually, a chemical signal on the outside of the cell is the "primary messenger" that causes the cell to respond. Usually, the chemical transmitter of information does not get into the cell. Rather, it

binds to surface receptors on the cell membrane. Somehow, the cells sense that a ligand is bound to the outside. Enzymes, usually in the membrane or at the intracellular surface of the lipid bilayer, are activated. Many of these enzymes cleave lipids in the membrane. The cleaved fragments of the lipid molecules serve as intracellular signals or "secondary messengers," which can bind to intracellular enzymes to activate intracellular processes. Figure (\PageIndex{30}\) below shows some lipid mediators generated by the process and signal the cell to respond.

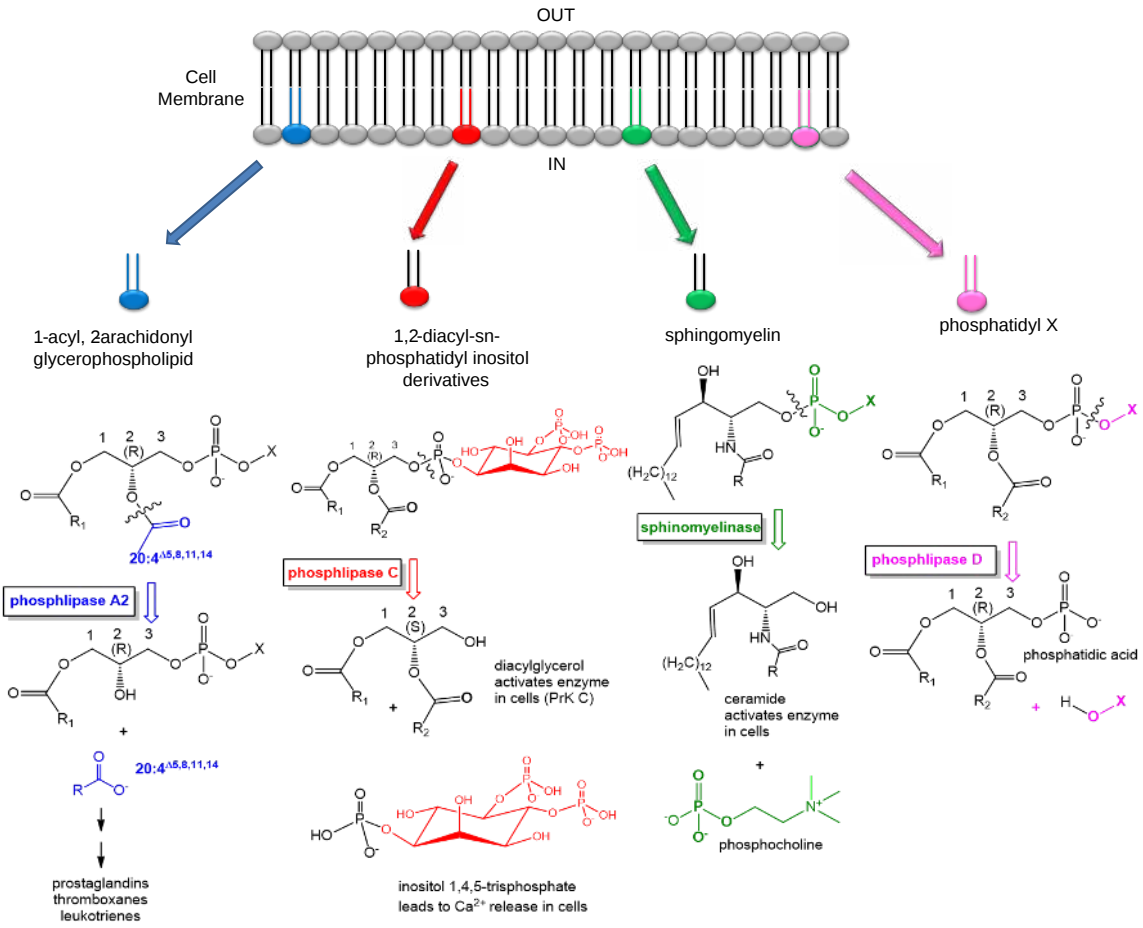
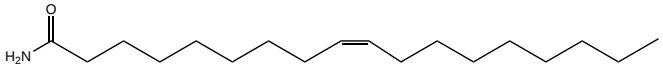


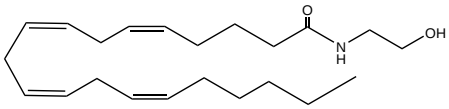
Figure (\PageIndex{30}\): Membrane lipids involved in signaling

Fatty acid amides are potent mediators of neurological processes. In one interesting experiment, sheep were sleep-deprived. Reasoning that the brain might release a biochemical signal into cerebrospinal fluid to induce sleep, scientists at Scripps removed some of this fluid. They isolated a substance that was not found in rested sheep. The structure was shown to be an amide of oleic acid on analysis. Oleylethanolamide has been shown to bind to the Peroxisome-Proliferator-Activated Receptor-a (PPAR-a) in the nucleus. By affecting gene transcription, this ligand appears to regulate body weight and the feeling of fullness after eating (satiety) as it leads to reduced eating.

People have sought the natural neurotransmitter, which binds to the same receptor in the brain as THC, the active ingredient of marijuana. The endogenous cannabinoid is an amide of arachidonic acid, anandamide. Figure (\PageIndex{31}\) below shows the structures of key fatty acid amides and THC.



oleamide -sleep inducing fatty acid amide



arachidonylethanolamide (anandamide)
endogenous THC-like ligand

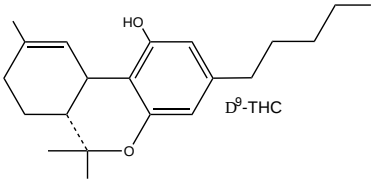


Figure (\PageIndex{31}\): Fatty Acid Amides: Neurochemical Mediators

This fatty acid amide is an example of a class of lipid derivatives called N-acylethanolamines (NAEs). These molecules, with acyl groups that vary in the number of carbons and double bonds, are found widely in nature. Naturally occurring anandamide leads to increased food intake after a short period of reduced food intake. One of the known physiological effects of THC is increased food consumption (the munchies).

Lucanic et al (2011) have shown that decreases in NAEs extend the life span of the small roundworm C. elegans, which has become a model organism to study genes in eukaryotes. Caloric restriction has been shown to increase the life span in various organisms. In invertebrates, anandamide seems to inhibit food intake, even in organisms that lack a receptor similar to the one in which cannabinoids bind. This might seem paradoxical in that anandamide (and THC) in humans seems to induce eating. However, under long periods of caloric restriction (low-level starvation) in rats, anandamide levels are suppressed, leading to a state of low energy consumption.

In general, it appears that reductions in NAEs occur during periods of caloric restriction. Mutant worms that have reduced levels of NAEs through targeted enzyme disruptions that affected either NAE synthesis or degradation have longer life spans. If normal (wild-type) worms were placed under caloric restriction, but given EPA-ethanolamine (the most abundant NAE in these worms), they did not have an extended lifespan.

12.2.10: Membrane Monolayers - Lipid droplets and lipoproteins

We explored membrane bilayers that contain two leaflets above. These bilayers separate the outside and inside of cells and the outside and inside of internal organelles. It turns out that there are two main types of lipid-encapsulated structures in which only one phospholipid leaflet separates the interior contents from the outside. These are lipid droplets and lipoproteins. In both cases, the monolayers encapsulate TAGs and cholesterol esters - hydrophobic molecules, so there is no need for another inner leaflet to stabilize an encapsulated aqueous environment.

12.2.10.1: Lipid Droplet formation

So, how are triacylglycerols stored in cells? In lipid droplets! In contrast to the common single and double-chain amphiphiles, which have charged atoms in their head groups and which form micelles and bilayers, respectively, triacylglycerols (TAGs) and cholesterol esters (CEs), which are almost completely nonpolar, coalesce into lipid droplets in cells. These droplets can range from very big, which are found in adipocytes (fat cells), where they take up almost all of the available space and where they are used for energy storage, to small, which are found in all cells, where they are used mostly for membrane biogenesis and energy mobilization. When esterified into esters, fatty acids and steroids also exhibit reduced toxicity to cells. Lipid droplets are often found in close approximation or attachment to mitochondria, endoplasmic reticulum (ER), and peroxisomes (where plasmalogens with ether-linked fatty acid instead of ester-linked are synthesized), all organelles intimately involved in membrane and energy

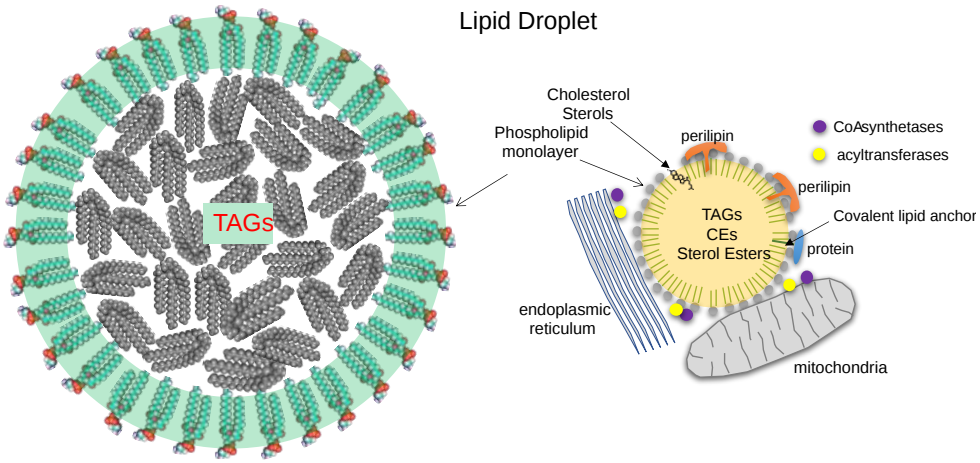
biochemistry. Many enzymes (acyltransferase, for example) required for TAG metabolism are found in the mitochondria and the ER.

Since lipid droplets are specialized for cells and don't form if TAGs are just added to water, we'll discuss their structures and their dynamic assembly in this section.

The droplets are now considered actual cellular organelles. In contrast to other organelles, which a bilayer member surrounds, lipid droplets are surrounded by a monolayer of phospholipids, which prevents exposure of the nonpolar contents to the aqueous cytoplasm. PC and PE appear to be the major phospholipids in the monolayer, and both are synthesized mostly in the ER. Many different proteins are involved in the formation and interaction with lipid drops, including

- perilipins: There are multiple types of perilipins. Perilipin 1 is found in adipocytes and cells synthesizing steroids (adrenals, ovaries, and testes). Perilipin 2 and 3 are found in most cells
- Acyl-CoA synthetases and acyltransferases: These enzymes activate free fatty acids for metabolic processes.
- seipins: These are involved in lipid droplet shape, number, and size. It appears to be involved in forming lipid drops and moving them from the ER to the cytoplasm. It facilitates the initiation of LD formation and ensures that vectorial budding of LDs from the ER is directed toward the cytoplasm.

It appears that lipid droplets arise from the ER, which is involved in membrane biogenesis and "trafficking" of membranes to different locations in the cell. A general structure of a **lipid droplet** is shown in Figure **12.2.32**. The left figure shows just internal triacylglycerols (TAGs). The inside would also contain cholesterol ester (fatty acid esterified to the cholesterol OH), and the monolayer would contain unesterified cholesterol.



(\PageIndex{32}\): Cartoon view of a lipid droplet.

How might these complex lipid droplets form in a cell without phase separation? With the help of proteins, of course! Figure **12.2.31** shows how newly synthesized TAGs (made in the ER membrane) could self-aggregate in the bilayer to form a "lens" which, on further growth and addition of lipid binding proteins, could bud off into the cytoplasm to form the droplets.

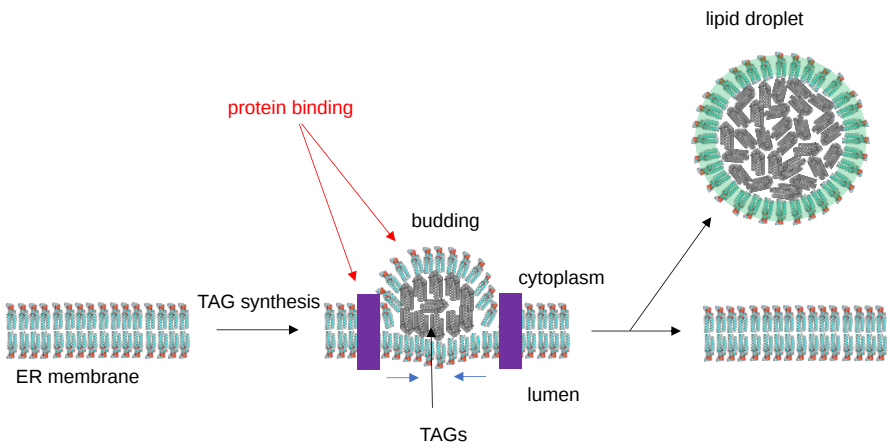


Figure 12.2.33: Formation of a lipid droplet

Figure 12.2.33 shows an incredible image, produced by a molecular dynamics simulation, of diacylglycerols (DAG) accumulating in a bilayer to form a clear lens in the membrane.

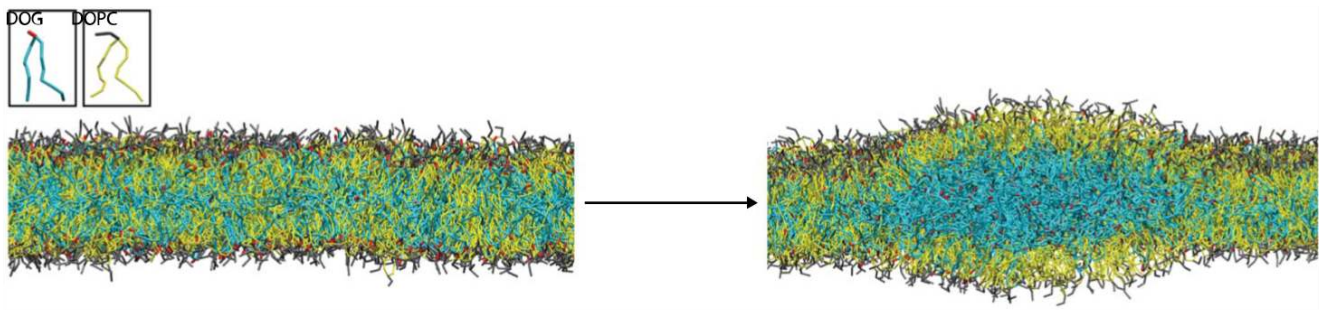


Figure 12.2.33: Molecular dynamic simulation of diacylglycerols (DAG) accumulating in a bilayer to form a clear lens in the membrane. Pablo Campomanes, Valeria Zoni & Stefano Vanni. Nature. https://www.nature.com/articles/s42004-019-0175-7.pdf. Creative Common License: http://creativecommons.org/ licenses/by/4.0/.

Figure (\PageIndex{34}\) below shows a cartoon illustrating the synthesis of a lipid droplet from the ER membrane.

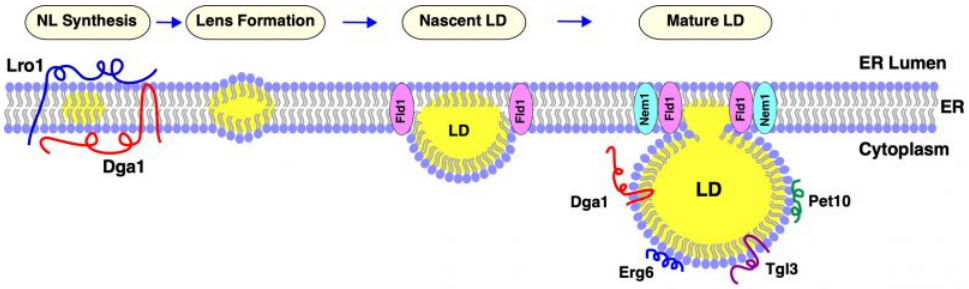


Figure (\PageIndex{34}\): **LD biogenesis in the ER.** ER-localized triacylglycerol-producing enzymes, Lro1, and Dga1 catalyze neutral lipid (NL, indicated in yellow) synthesis from opposite sides of the ER membrane. The NL then accumulates between the two leaflets of the ER membrane leading to the formation of lipid lenses. These NL lenses grow in size to become nascent LDs, which then emerge towards the cytoplasm where they further mature. Fld1 and Nem1 proteins show punctate localization at ER-LD contact sites. The acyltransferase Dga1 and the TAG lipase Tgl3 translocate onto the periphery of mature LDs. Finally, LD-marker proteins such as the perilipin ortholog Pet10 and the sterol biosynthetic enzyme Erg6 decorate the surface of the mature LD. http://microbialcell.com/researchart...icrobial-cell/ . Creative Commons Attribution 4.0 International License.

A membrane protein not displayed in the above figure, seipin, denotes the location for lipid droplet formation in the ER membrane. Mutants of the protein are associated with lipodystrophy. In yeast, seipin and another membrane protein, Ldb16, associate and facilitate lipid droplet formation. Seipin aggregates to form a homo 10-mer in the membrane, but in contrast to human cells, it alone can not concentrate triacylglycerol in the membrane. It requires the binding of Ldb16 for that to happen.

Figure (\PageIndex{35}\) below shows yeast, human, and fly seipin and their properties.

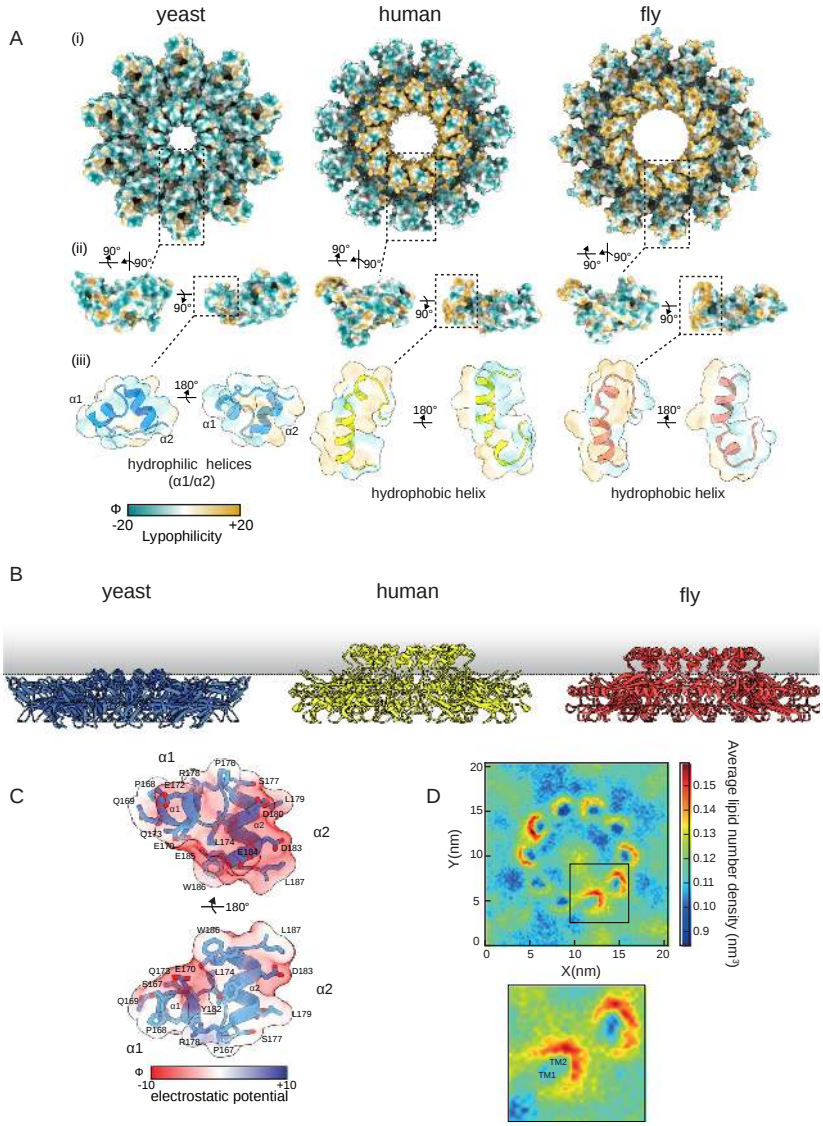


Figure (\PageIndex{35}\): Yeast, human, and fly seipin and their properties. *Nat Commun* **12**, 5892 (2021). https://doi.org/10.1038/s41467-021-26162-6. Klug, Y.A., Deme, J.C., Corey, R.A. *et al*. Mechanism of lipid droplet formation by the yeast Sei1/Ldb16 Seipin complex. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel **A** shows a lipophilicity potential on the surface of yeast (left), human (middle), and fly (right) as viewed as (i) homodecamer assembly from the cytosol, (ii) individual protomers, or (iii) transparent overlay over zoomed-in cartoon representation of the central α1-α2 helices. Surfaces are colored from hydrophilic (dark cyan) to hydrophobic (gold).

Panel B shows a side view of the luminal domains of yeast, human (PDB 6DS5), and fly (PDB 6MLU) Seipin in relation to the plane of the ER membrane (indicated by a dotted line).

Panel C shows the charge distribution of the yeast Sei1 central helices (α1, α2), depicted as a transparent Coulombic electrostatic potential surface representation (Red, negative charge; blue, positive charge; white, no charge) overlayed over a cartoon representation (light blue) to show acidic side chains.

Panel D shows a top view of molecular dynamics simulations of Sei1 in a POPC membrane with 3% trioleylglycerol. Images depict the average lipid number density of trioleylglycerol. Inset – zoom of the corresponding box showing positions of TM1 and TM2.

The lumenal domains form the ring with a floor, as shown in Panel A above. In addition, the transmembrane segments for the cage top and sides. A switch area between the lumenal and transmembrane segments occupying two different conformations appears important for function. The closed cage allows accumulation and hence phase-separation of triacylglycerols, while the open form allows the nascent droplet to grow and then bud.

The Ldb16 has helical regions enriched in serine and threonine, as is required for TAG loading. These -OH-containing amino acids are present in seipins in humans and flies. Site-specific mutations of the serine and threonines in the region of Ldb16 lead to problems with lipid droplet formation.

Figure (\PageIndex{36}\) below shows a cartoon of the serine- and threonine-enriched helix in Ldb16.

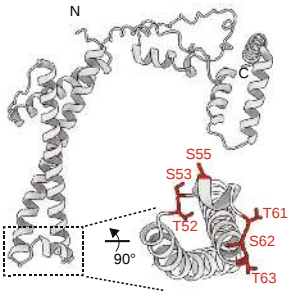


Figure (\PageIndex{36}\): Cartoon of Ldb16 predicted helix enriched in serine and threonine. Klug et al., ibid.

Another model for the assembly of lipid droplets by the Sei1-Ldb16 yeast complex is shown in Figure (\PageIndex{37}\) below.

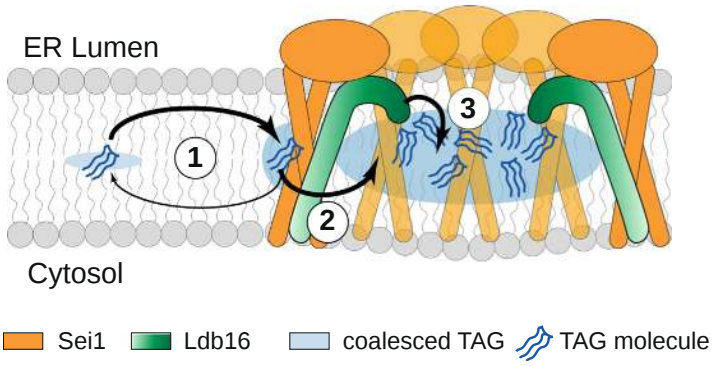
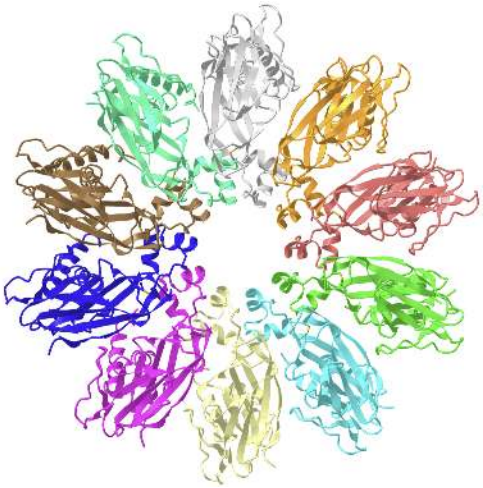


Figure (\PageIndex{37}\): Sequential TAG interactions mediate LD assembly by the Sei1-Ldb16 complex. In the ER bilayer, TAG molecules (blue) concentrate near Seipin oligomers (orange) via weak interaction with Sei1 TMs. TAG molecules within the ring interact strongly with Ldb16 (green) hydroxyl-containing residues, facilitating TAG coalescence and lens formation. Klug et al., ibid.

Figure **12.2.38** shows an interactive iCn3D model of the homo 10-mer yeast seipin membrane complex **(7OXP).**



NCBI iCn3D Figure 12.2.38: Yeast seipinhomo 10-mer yeast seipin membrane complex (7OXP). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wMLHBncjFZMbFA

The two transmembrane segments (helices) from each of the monomers are evident by rotating the structure so the homo 10-mer ring is viewed from the side.

12.2.10.2: Lipoproteins

We eat, digest, and transport dietary fat. We also make fat and transport it through the blood. We saw that free fatty acids are carried in the blood by the most abundant protein in the blood, albumin. What about the very insoluble triacylglycerols and cholesterol esters? They are also transported in the blood by nanoparticles similar to lipid droplets. They are called lipoproteins since, like lipid droplets, proteins are associated with them.

Lipoproteins vary in density and size. The densest is called high-density lipoproteins (HDL). As they get larger and more filled, they form less dense lipoproteins (low-density LDL, intermediate-density IDL, and very-low-density VLDL). These contain nondietary lipids made by organs like the liver. Dietary fats are processed in the intestine into very large particles called chylomicrons. Figure 12.2.39 shows their relative size and density.

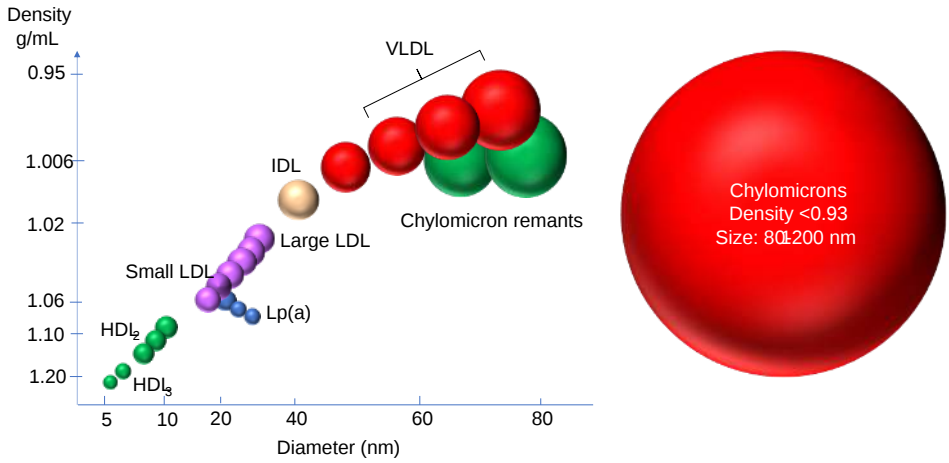


Figure 12.2.39: Size distribution of lipoproteins

Introduction to Lipids and Lipoproteins. Kenneth R. Feingold, MD. Creative Commons (CC-BY-NC-ND) license. A copy of the license can be viewed at http://creativecommons.org/licenses/by-nc-nd/2.0/. With permission.

As with lipid droplets, lipoproteins have a single outer monolayer leaflet containing double-chain membrane lipids like phosphatidylcholine and free cholesterol. Inside are the triacylglycerols and cholesterol esters. Proteins are bound to the outer monolayer. Figure 12.2.40 shows two renderings of discoidal HDL particles containing a single type of protein, Apo-A1. The TAGs are shown in cyan line rendering on the inside, along with cholesterol esters (in spacefill). The bottom part of the figure shows the polar Ns and Os decorating the outer part of the monolayer of phosphatidylcholine surrounding the TAGs and cholesterol esters.

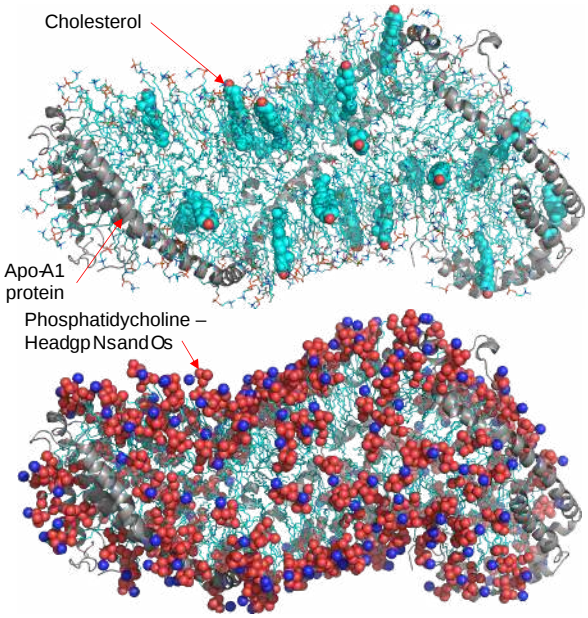
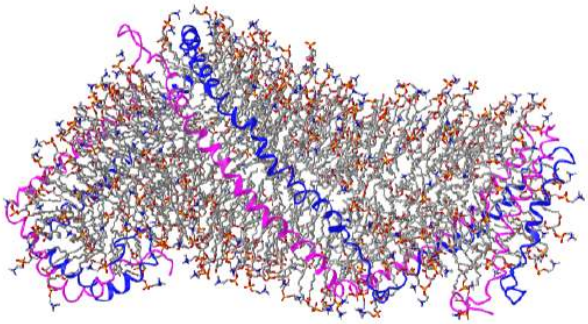


Figure 12.2.40: Two renderings of a discoidal HDL particle containing a single type of protein, Apo-A1

Figure 12.2.41 shows an interactive iCn3D model of discoidal HDL (3k2s)



NCBI iCn3D Figure 12.2.41: discoidal HDL (3k2s) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...BE5XYUopbjGALA

Table 12.2.1 below shows the proteins associated with the different types of lipoproteins.

Lipoprotein	Density (g/ml)	Size (nm)	Major Lipids	Major Apoproteins
Chylomicrons	<0.930	75-1200	Triglycerides	Apo B-48, Apo C, Apo E, Apo A-I, A-II, A-IV
Chylomicron Remnants	0.930- 1.006	30-80	Triglycerides Cholesterol	Apo B-48, Apo E
VLDL	0.930- 1.006	30-80	Triglycerides	Apo B-100, Apo E, Apo C
IDL	1.006- 1.019	25-35	Triglycerides Cholesterol	Apo B-100, Apo E, Apo C
LDL	1.019- 1.063	18- 25	Cholesterol	Apo B-100
HDL	1.063- 1.210	5- 12	Cholesterol Phospholipids	Apo A-I, Apo A-II, Apo C, Apo E
Lp (a)	1.055- 1.085	~30	Cholesterol	Apo B-100, Apo (a)

Table 12.2.1: Proteins associated with the different types of lipoproteins.Introduction to Lipids and Lipoproteins. Kenneth R. Feingold, MD. Creative Commons (CC-BY-NC-ND) license. A copy of the license can be viewed at http://creativecommons.org/licenses/by-nc-nd/2.0/.

12.2.11: Summary

Chapter Summary: Overview of Lipid Bilayers

This chapter presents a comprehensive look at the structure, composition, dynamics, and functional complexity of lipid bilayers—the fundamental barriers that define cellular boundaries and compartmentalize intracellular processes.

It begins by describing the basic architecture of the membrane bilayer as an assembly of amphiphilic lipids arranged into two distinct leaflets. These lipids are not distributed randomly; they exhibit significant asymmetry, with different classes (e.g., phosphatidylcholine, sphingomyelin in the outer leaflet versus phosphatidylethanolamine and phosphatidylserine in the inner leaflet) contributing to unique physical and functional properties in various cellular compartments. The chapter highlights how lipid composition varies among organelles and cell types, emphasizing the central role of the endoplasmic reticulum in lipid synthesis and distribution.

The dynamic nature of membranes is a key theme. Lipids can laterally diffuse within a leaflet at high rates, while transverse (flip-flop) movement between leaflets is more restricted and often catalyzed by specialized enzymes such as flippases, floppases, and scramblases. These dynamics are crucial for maintaining membrane asymmetry and for processes like vesicle formation and membrane fusion.

Thermodynamic principles are used to explain how noncovalent interactions drive the self-assembly of lipids into organized structures. The chapter discusses phase transitions in membranes—such as the shift from a rigid gel phase to a more fluid liquid crystalline phase—as a function of temperature and lipid composition. Cholesterol plays a pivotal role here, modulating membrane fluidity by inhibiting both excessive ordering at low temperatures and over-flexibility at high temperatures.

Lipid rafts, which are dynamic, nanoscale domains enriched in cholesterol, sphingolipids, and specific proteins, are introduced as platforms for signaling and protein sorting. These rafts arise from lateral phase separation within the membrane and are essential for many cell signaling pathways and membrane trafficking events.

The chapter also extends its discussion to more complex lipid assemblies beyond bilayers. Lipid droplets, which store triacylglycerols and cholesterol esters, are described as organelles encapsulated by a phospholipid monolayer. Lipoproteins, the vehicles for lipid transport in blood, are similarly detailed—highlighting their structural variability and biological importance.

Overall, the chapter ties together how the unique chemical properties of lipid molecules—determined by their fatty acid composition, head group characteristics, and interactions with cholesterol—dictate the physical state, dynamics, and biological

functions of membranes. This “structure determines function” paradigm is central to understanding the behavior of membranes in both in vitro experiments and in the complex environment of living cells.

This page titled 12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **10.3: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.3: Working with Lipids

Learning Goals

1. **Describe Liposome Formation and Lipid Self-Assembly:**
   • Explain the principles of liposome formation, including how dried lipids are rehydrated and dispersed into multilamellar vesicles (MLVs) and how these can be processed into unilamellar vesicles (SUVs, IUVs, and LUVs) using techniques such as extrusion, freeze-thaw cycles, and detergent removal via dialysis or gel filtration.
2. **Understand the Role of Lipid Composition in Membrane Properties:**
   • Discuss how the chemical composition (e.g., the balance of phospholipids like PC, PE, and sphingolipids, along with cholesterol and detergents) affects liposome properties such as transition temperature (Tm), membrane fluidity, permeability, and stability.
3. **Relate Liposome Characteristics to Their Applications:**
   • Identify how liposomes are used as models for biological membranes, as vehicles for drug and vaccine delivery (e.g., mRNA vaccines for SARS-CoV-2), and for the solubilization of membrane proteins, emphasizing the importance of controlling vesicle size, lamellarity, and encapsulation efficiency.
4. **Explain the Methods for Liposome Preparation:**
   • Describe the three primary methods for dispersing lipids into aqueous solutions—mechanical dispersion, organic solvent dispersion, and detergent dispersion—and evaluate their advantages and limitations for encapsulating water-soluble compounds.
5. **Understand Analytical Techniques for Lipid Characterization:**
   • Explain the principles behind gas chromatography (GC) and mass spectrometry (MS) for lipid analysis, including the need for derivatization (e.g., transesterification and silylation) to render lipids volatile or amenable to ionization.
6. **Differentiate Advanced Mass Spectrometric Approaches in Lipidomics:**
   • Discuss how methods such as ozone-induced dissociation (OzID), ultraviolet photodissociation (UVPD), and the Paternò–Büchi (PB) photochemical reaction coupled with tandem MS are used to resolve lipid isomers by determining the locations and geometry of carbon–carbon double bonds and sn-positions.
7. **Integrate Thermodynamic Principles with Lipid Aggregation:**
   • Apply concepts of free energy, enthalpy, and entropy to understand the driving forces behind micelle and liposome formation, including the role of the critical micelle concentration (CMC) and how noncovalent interactions govern the self-assembly of amphiphiles.
8. **Assess the Impact of Lipid Structure on Health and Disease:**
   • Evaluate how subtle differences in lipid structure, such as degree of unsaturation and fatty acid chain length, influence membrane function and are linked to health outcomes and diseases, as revealed through detailed lipidomic studies.

These goals will help you synthesize knowledge about the physical and chemical properties of lipids, their assembly into functional structures like liposomes, and the advanced analytical techniques used in lipidomics to study their roles in biological systems.

12.3.1: Introduction

Lipids, although small compared to large biopolymers like proteins, nucleic acids, and large glycans, are very heterogeneous in structure, given the large array of fatty acid and isoprenoid chain lengths, numbers of double bonds, etc, that appear in different lipid classes. In addition to analyzing lipid structure, lipids are used in the laboratory to create liposomes (vesicles), which serve as models for membrane bilayers and encapsulate chemical species (drugs, vaccines) for medical use and solubilization of membrane proteins. This section will concentrate on creating lipid vesicles (critical for encapsulating RNA vaccines against the SARS-CoV-2 spike protein) and the chemical analysis of biological lipids, whose composition affects health and disease states.

12.3.2: Liposomes

Liposomes produced in the lab can be unilamellar, which consists of a single bilayer surrounding the internal aqueous compartment, or multilamellar, which consists of multiple bilayers surrounding the enclosed aqueous solution. You can imagine the multilamellar vesicles resemble an onion with its multiple layers. Cartoons of unilamellar and multilamellar liposomes are shown in Figure 12.3.1, where each concentric circle represents a bilayer.

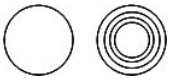


Figure 12.3.1: Cartoons of unilamellar and multilamellar liposomes (each concentric circle represents a bilayer)

Liposomes vary in diameter. They can be generally categorized into small (S, diameter < 25 nm), intermediate (I, diameter around 100 nm), and large (L, diameter from 250-1000 nm). If these vesicles are unilamellar, they are abbreviated as SUV, IUV, and LUV, respectively. Their various sizes are shown in Table 12.3.1 below, compared to other large biological structures.

50 nm	200 nm	2000 nm	others
SUV nuclear pore Cell membrane: 7 Ribosome: 20	LUV lysosome HIV: 100	E. Coli Mitochondria: 3000	Red blood cell: 7,000 animal cell: 10,000 – 30,000 Human egg: 120,000

Table 12.3.1: Sizes of liposomes/vesicles compared to other biological structures.

The chemical composition of liposomes can be widely varied. Most contain neutral phospholipids like phosphatidylcholine (PC), phosphatidylethanolamine (PE), or sphingomyelin (SM), supplemented, if desired, with negatively charged phospholipids, like phosphatidylserine (PS) and phosphatidylglycerol (PG). In addition, single-chain amphiphiles like cholesterol (C) and detergents can be incorporated into the bilayer membrane, which modulates the fluidity and transition temperature (Tm) of the bilayer. If present in too great a concentration, single-chain amphiphiles like detergents, which form micelles, can disrupt the membrane so completely that the double-chain amphiphiles become incorporated into detergent micelles, now called mixed micelles, in a process that effectively destroys the membrane bilayer.

Given the large degree of unsaturation at C2, what do you expect the transition temperature of a liposome composed only of egg yolk PC to be? (Vesicles made using more saturated PC from mammalian sources have a Tm of around 40 ^{o}C.) This high degree of unsaturation makes egg yolk PC very susceptible to oxidation, which could dramatically alter the liposome's properties. Synthetic PC made with saturated fatty acids could alleviate that problem.

The properties of liposomes (charge density, membrane fluidity, and permeability) are determined by the lipid composition and size of the vesicle. The desired properties will be, in turn, determined by the use of the particular liposome. The vesicles offer wonderful, simple models to study the biochemistry and biophysics of natural membranes. Membrane proteins can be incorporated into the liposome bilayer using the exact method you will be using. However, apart from these purposes, liposomes can encapsulate water-soluble molecules such as nucleic acids, proteins, and toxic drugs. These liposomes can be targeted to specific cells if antibodies or other molecules that will bind specifically to the target cell can be incorporated into the bilayer of the vesicle. Intraliposomal material may then be transferred into the cell by fusion of the vesicle with the cell or by endocytosis of the vesicle.

Since phospholipids will spontaneously form a bilayer structure when placed in water, most efforts in liposome production involve producing vesicles with the desired size, lamellar structure, and physical characteristics, which, as previously stated, are controlled by liposome size and chemical composition. Also, ways must be developed to entrap the desired molecule inside the vesicle cost-effectively and with minimal contents leaking. All production methods involve drying organic solvent-solubilized lipids, dispersion of the lipids in the appropriate aqueous solution, and formation of monolamellar (one bilayer) liposomes or vesicles. Finally, the vesicles are characterized (chemical composition, T_m, permeability, size, etc.)

Drying of lipids

Purified lipids of the desired composition (often egg PC:cholesterol:PS in molar ratios of 0.9:1.0:0.1) are dissolved in a purified, water-free organic solvent mixture (often chloroform/methanol, 2:1 v/v) and dried down in a round bottom flask on a rotary evaporator under reduced pressure (using a water aspirator) and slightly elevated temperature (20-40 ^{o}C). The rapid rotation of the flask will ensure that the lipid is dispersed over a large surface area and will increase the evaporation rate. The dried flask is usually placed under a high vacuum overnight to remove the last traces of solvent. If a small volume (< 1 ml) of lipid solution is used, the solvent can be evaporated under a stream of nitrogen. To avoid entrapment of residual chloroform in the lipid film, the film is dissolved in t-butyl-methyl ether or diethyl ether, and dried several times. Alternatively, the residual solvent can be removed under a high vacuum.

Dispersion of the lipids

There are three main methods of dispersing lipids into an aqueous solution to form liposomes.

a. mechanical dispersion - in this method, lipid dried onto the inside glass surface of a container is hydrated with an aqueous solution, which peels off the lipid to form multilamellar - MLV - (multiple bilayers separated by water) vesicles. Only a small part of the aqueous solution is encapsulated inside the liposome, so this is not the method of choice for the encapsulation of expensive or rather insoluble solutes. Different-sized liposomes can be prepared depending on the degree of agitation and the nature of the lipid used.

b. organic solvent dispersion - In these methods, the lipids, which are dissolved in organic solvents, are injected through a fine needle, at a slow rate, into an aqueous solution in which the organic solvent may be miscible (such as ethanol) or immiscible (such as ether). In each case, the lipids are oriented at the interface between the organic solvent and aqueous solution to form bilayer structures. Injection of ethanol-dissolved lipids provides a simple way to produce SUVs, but because liposome formation can not occur at an ethanol concentration greater than 7.5%, only a fraction of the total aqueous phase can be entrapped in the vesicle; hence, this technique is not cost-effective for the entrapment of an expensive solute. Alternatively, the lipid can be dissolved in ether and slowly injected into an aqueous solution, which is warmed so that the ether evaporates at the rate at which it is injected. Since the ether is volatilized, large amounts of lipid can be introduced, and the encapsulation efficiency of the aqueous solution is high.

c. detergent dispersion and solubilization - In this method, lipids are solubilized in an aqueous solution by adding detergents. The detergents are removed slowly from the solution, resulting in the spontaneous formation of liposomes. Detergents are single-chain amphiphiles that spontaneously form micelles in an aqueous solution when the concentration of free lipid rises to a minimum critical value, the critical micelle concentration (CMC); at this concentration, self-association of the detergent results in the formation of a stable aggregate, the micelle. This is illustrated in Figure 12.3.2.

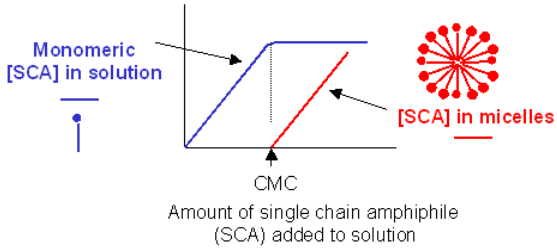


Figure 12.3.2: [Single chain amphiphile] in solution vs [Single chain amphiphile] added - the CMC

Table 12.3.2 below shows the properties and CMC of various detergents.

Table 12.3.2: Properties and CMC of various detergents (data from Avanti Polar Lipids)

Name	mM	mg/ml	MW
n-hepty glucopyranoside	70	19.5	278
n-octyl glucopyranoside	23.2	6.8	292
n-nonyl glucopyranoside	6.5	2.0	306
n-decyl maltoside	2.19	1.1	499
n-dodecyl maltotrioside	0.2	0.16	825
Triton X-100 (a)	0.24	0.15	625
Nonidet P-40 (b)	0.29	0.02	603
Tween 20 (c)	0.033	0.04	1364
Brij 98 (d)	0.025	0.04	1527
sodium deoxycholate	2-6	1.7	415
sodium taurocholate	10-15	6.7	538
sodium cholate	14	6.0	431
sodium dodecyl sulfate	8.3	2.4	289

12.3.3: Making liposomes by dialysis

Lipids are deposited in a small container. An aqueous solution is then added, containing water-soluble molecules for encapsulation. Detergent is added at a concentration over the lipid concentration and greater than its CMC. The lipid molecules are then "emulsified" in the detergent micelle. The solubilized mixture is then placed in a semipermeable dialysis bag, which is placed in a large volume of an aqueous solution. The free detergent in solution is in equilibrium with the detergent in the micelle. The bag contains microscopic holes large enough for the monomeric detergent molecule to pass through, but small enough so that the large micelle can not. The lipid is embedded in the micelle during this process, forming a detergent-lipid mixed micelle. As dialysis continues, the monomeric detergent partitions throughout the volume in the bag and surrounds the bag while the mixed micelle remains in the bag. If the aqueous solution surrounding the bag is changed several times with fresh solution, the equilibrium in the bag is shifted to the monomeric form. Alternatively, detergent-adsorbing beads (such as Bio Bead SM-2 by Bio-Rad) can be placed in the aqueous solution surrounding the bag to speed up the process of detergent re-equilibration. Eventually, all the detergent is in this form, and during the process, which occurs slowly, the lipid in the mixed micelle self-associates to form a liposome. A detergent of low monomer molecular weight and a high CMC is most desirable for this method of liposome production. Another method of removing the free detergent is through gel filtration chromatography. In this technique, molecules of disparate molecular weights can be separated. An explanation follows this discussion. This method of formation of unilamellar liposomes is the method of choice if membrane proteins are to be inserted into the liposome bilayer to target the liposome. However, it is not the best method for quantitative encapsulation of expensive soluble molecules.

12.3.4: Making Liposomes by Extrusion

Several techniques can further process these multilamellar liposomes to form unilamellar liposomes. These include probe or bath sonication of the MLVs, extrusion at high pressure of the MLVs through membrane filters of defined pore size, or pH-induced vesiculation in which a transient change in pH destabilizes the MLVs in favor of unilamellar liposomes. Another technique involves the fusion of SUVs by repeated freezing and thawing or by fusion of SUVs containing acidic phospholipids (such as PS) through Ca^{2+}-mediated aggregation.

Figure **12.3.3**shows the structure of vesicles as they undergo multiple freeze/thaw cycles.

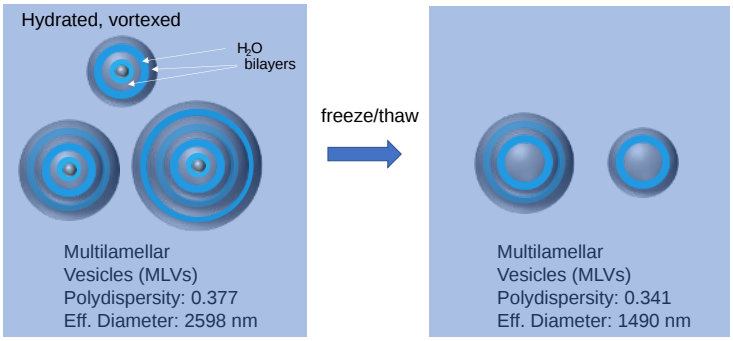


Figure 12.3.3: Structure of multilamellar vesicles as they undergo multiple freeze/thaw cycles

Figure 12.3.4 shows the final step in making large unilamellar vesicles by extrusion of freeze/thaw intermediates.

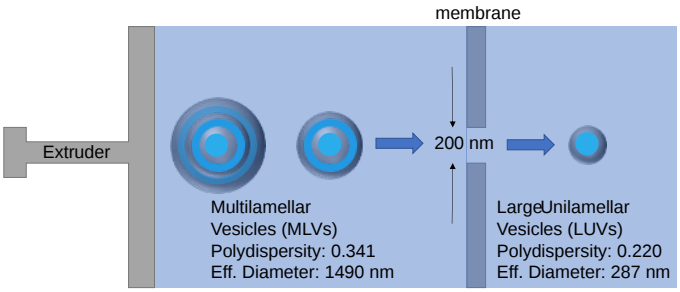


Figure 12.3.4: Conversion of multilamellar vesicles to unilamellar vesicles by extrusion

Once the liposomes are formed, they must be separated from free monomeric lipids, detergents, and unencapsulated solutes. This can be done by dialysis or, more easily, by size exclusion chromatography. Macromolecules of different sizes can be separated on a column in which the stationary phase is a polymerized agarose or acrylamide bead containing pores of various sizes. A small molecule (such as a monomeric detergent, free lipid, or small aqueous solute) in the mobile phase (aqueous buffered solution) may enter the pores in the bead, while a larger macromolecule or aggregate (such as a large protein, a micelle, or a liposome) may not, due to size restriction. The result is that a larger fraction of the overall volume of the column is available to the smaller molecules, which thus spend a longer time on the column and are eluted by the mobile solvent after the larger species. Liposomes can be characterized chemically to determine the average lipid and protein (if they were incorporated) composition of the bilayer and physically to determine the size, permeability, lamellarity, and amount of encapsulated material. Size is usually determined by electron microscopy or indirectly by light scattering from these large species.

12.3.5: Chemical Analysis of Lipids

There are so many lipids and their derivatives that the most sensitive techniques are needed for separation and analysis. The field of lipidomics focuses on analyzing the structure and function of all lipids in the cell. The most valuable techniques are analysis (and separation) by gas chromatography (GC) followed by mass spectrometry (MS). NMR spectroscopy is also essential. For GC analysis, the lipid must be made volatile, which limits its use in some circumstances. MS analysis is the most sensitive. MS requires ion formation, and techniques like electrospray ionization and matrix-assisted laser desorption/ionization (MALDI) are used.

12.3.5.1: Gas chromatography

(This section is adapted from https://www.intechopen.com/chapters/64008. Creative Commons Attribution 3.0 License). In GC, gas is the mobile phase that carries lipids through the column's stationary phase. Gas chromatography (GC) separates organic compounds from a mixture in the gas form. For this purpose, the GC uses interactions among the sample components and the stationary/mobile phases. After lipid extraction with chloroform and/or methanol, the samples (lipid mixture) are usually liquids and must be exposed to a high temperature at the gas chromatograph entrance (injector). Vaporized, the samples are carried by a

gas, which is usually a light and inert gas (i.e., hydrogen, helium), through a long capillary column containing a high or low polarity material (stationary phase)

The gaseous compounds generated from the vaporized sample interact with the stationary phase, allowing each compound to elute/separate at a different time (retention time). Because GC considers both the chemical and physical properties of the vaporized compounds, those with more chemical affinity to the stationary phase will take longer to be removed from the column, and the temperature will influence the overall process. This explains why the column stays in an oven, which is programmed to work at different temperature ranges (i.e., temperature programming). The mobile phase gas carries the compounds, and their boiling point determines their elution time.

At the end of GC analysis, the electronic detector generates a chromatogram based on retention time by intensity. This allows a qualitative identification of the lipid compounds by comparing their retention times with certified standards using the flame ionization detector (FID) or by deduction of spectral information using a mass spectrometer as a detector. Lipid quantification can also be performed using analytical procedures of external or internal certified standards in GC analysis.

The main points to be considered when assessing FAs by GC analysis are the carrier gas flow rate, column length, and temperature, because these can influence the order or retention time of the lipid compounds and must be precisely standardized. The column length of the stationary phase influences the resolution of the analytes once the number of theoretical plates (hypothetical zone in which two phases establish an equilibrium with each other) is higher in the longer column. As fats and oils have high boiling points not supported by the stationary phase, a previous derivatization reaction step is required after lipid extraction from the biological sample, in which triacylglycerol and free fatty acids are transformed into their respective free fatty esters with lower boiling points (transesterification/esterification reaction). Several methods are available for FA derivatization.

Particularly for cholesterol analysis, the sample preparation involves a derivatization reaction. This increases the volatility of the compounds. It also generates products with reduced polarity. Cholesterol derivatization is usually achieved by using trimethylsilyl compounds through silylation reactions. A common method for this purpose uses *N*,*O*-bis(trimethylsilyl-trifluoroacetamide/trimethylchlorosilane).

12.3.5.2: Mass spectrometry

(This section is derived from https://doi.org/10.1038/s41467-019-14180-4. Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0/).

Mass spectrometry (MS) has become the method of choice for lipid analysis, offering label-free detection at high sensitivity and structural characterization capability. However, large-scale lipid analysis with a comprehensive capability of revealing all levels of structure information still represents a significant analytical challenge for lipidomics. General protocols, for instance, have five levels of structure information, including lipid class, fatty acyl identities, fatty acyl *sn*-positions, and C=C location/geometry (viz cis/trans) in the fatty acyl. Successful attempts to determine C=C locations in fatty acyls or their *sn*-positions have already been reported for MS analysis, enabling the characterization of detailed structure moieties and the identification of lipid structure isomers. An extremely useful feature offered by lipid isomer analysis is the relative quantitation achieved at high precision without using lipid standards, which are not readily available. Remarkably, our recent study demonstrated a close correlation between the lipid C=C location isomer compositions and Type II diabetes, which owes to tighter regulation on lipid desaturation, allowing efficient elimination of interferences due to variations among samples."

Various methods have been explored for differentiating the lipid C=C location and *sn*-position isomers. Ozone-induced dissociation (OzID) and ultraviolet photodissociation (UVPD) have been used to determine both *sn*-positions and C=C locations in GPs. By coupling the Paternò-Bǔchi (PB) photochemical reaction with tandem MS (MS/MS), qualitative and quantitative analysis of lipids with C=C specificity from complex biological samples can be accomplished. PB reaction converts the C=C to an oxetane, which can be preferentially fragmented by low-energy collision-induced dissociation (CID).

An ideal analytical tool for lipidomics to survey a wide range of lipids in discovery work should provide detailed information at multiple lipid structure isomer levels (e.g., C=C location/geometry and *sn*-position) and be feasible for large-scale quantitative analysis. UVPD can assign C=C locations and *sn*-positions of fatty acyls, while OzID may be the only one that has been demonstrated to assign C=C locations in *sn*-specific fatty acyls well. One problem of OzID is the long reaction time required for the ion trap implementation; however, recent work has demonstrated the ability to perform OzID on LC-compatible time scales

in the high-pressure regions of the MS system. For the PB reaction method, both shotgun analysis and HPLC-PB-MS/MS workflow have been developed to identify many C=C location isomers. Figure **12.3.5** describes methods for determining C=C location and sn-isomers in lipids.

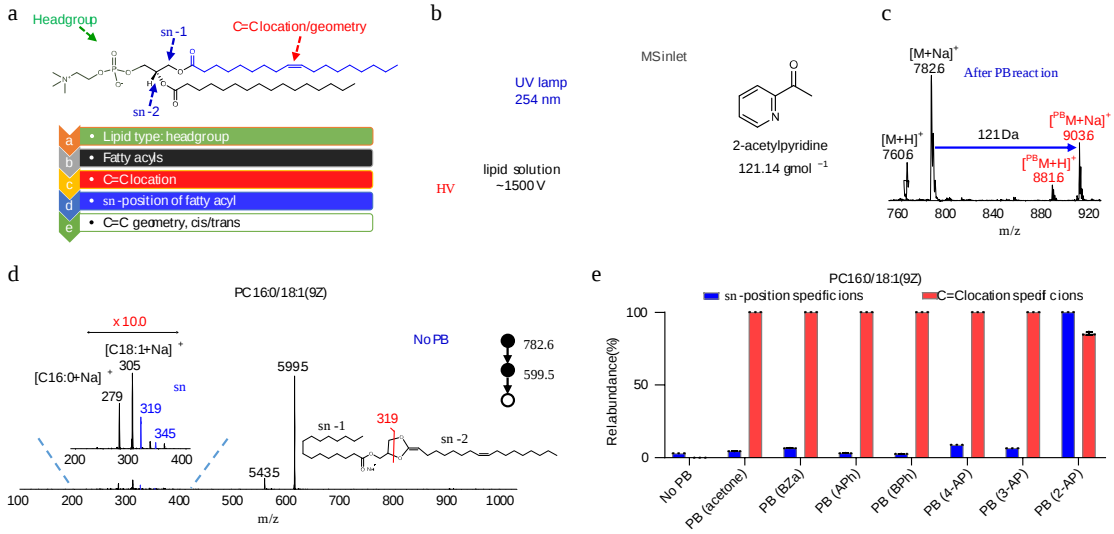


Figure **12.3.5**: Methods for the determination of C=C location and sn-isomers in lipids. **a** The chemical structure of PC 18:1(9Z)/16:0. **b** Schematic of the experimental setup for online derivatization of unsaturated GPs by coupling 254 nm irradiation with nanoESI-MS. **c** MS spectrum of PC 16:0/18:1(9Z) after 30 s reaction. **d** MS^3 spectrum of PC 16:0/18:1(9Z) without PB derivatization. Sodiated lipid precursors were first fragmented to generate product ions after headgroup loss (−183 Da). Product ion at *m/z* 599.5 was further fragmented to release *sn*-1-specific diagnostic ions at *m/z* 319 (C16:0) and 345 (C18:1). **e** Comparison the relative abundance of *sn*-position and C=C location specific ions of PC 16:0/18:1(9Z) without and with PB derivatization using different PB reagents (Bza: benzaldehyde, APh: acetophenone, BPh: benzophenone, AP: acetylpyridine). The error bar represents the standard deviation, $n = 3$. Cao, W., Cheng, S., Yang, J. *et al.* Large-scale lipid analysis with C=C location and *sn*-position isomer resolving power. *Nat Commun* **11,** 375 (2020). https://doi.org/10.1038/s41467-019-14180-4. DOI: https://doi.org/10.1038/s41467-019-14180-4. Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0/.

12.3.6: Summary

Chapter Summary: Liposome Formation and Lipid Analysis

This chapter provides an in-depth exploration of the formation, manipulation, and chemical analysis of lipid assemblies, with a focus on liposomes as model membrane systems and vehicles for drug and vaccine delivery. It begins by highlighting the inherent heterogeneity of lipids, driven by variations in fatty acid and isoprenoid chain lengths and degrees of unsaturation, which underlie the diverse functions of lipids in biological systems.

Liposome Formation and Characterization:

The chapter explains that when lipids are rehydrated, they spontaneously self-assemble into vesicular structures. Depending on the method of preparation and the lipid composition, liposomes can be multilamellar (multiple concentric bilayers, like an onion) or unilamellar (a single bilayer encircling an aqueous core). Techniques such as mechanical dispersion, organic solvent injection, and detergent-mediated dispersion are described in detail. Each method influences the size (ranging from small to large vesicles), lamellarity, and encapsulation efficiency of the resulting liposomes, which are crucial for their applications in encapsulating hydrophilic drugs, nucleic acids, or membrane proteins.

Chemical Composition and Membrane Properties:

The chemical composition of liposomes—typically containing neutral phospholipids (PC, PE, or sphingomyelin), possibly supplemented with charged lipids (PS, PG), cholesterol, or detergents—plays a critical role in determining their transition temperature (Tm), fluidity, permeability, and overall stability. The chapter emphasizes that the degree of unsaturation in the fatty acids affects not only the physical state of the liposome but also its susceptibility to oxidation, which is important in both natural membranes and synthetic vesicle applications (e.g., in RNA vaccines against SARS-CoV-2).

Analytical Techniques in Lipidomics:
A significant portion of the chapter is devoted to the chemical analysis of lipids, highlighting the importance of lipidomics—the study of the structure and function of cellular lipids. Key analytical techniques include gas chromatography (GC) and mass spectrometry (MS), which often require prior derivatization (transesterification for fatty acid methyl esters, silylation for cholesterol derivatives) to enhance volatility and detectability. Advanced MS methods, such as ozone-induced dissociation (OzID), ultraviolet photodissociation (UVPD), and the Paternò–Büchi (PB) photochemical reaction, are introduced. These approaches allow researchers to resolve lipid isomers by pinpointing double-bond positions, geometries (cis/trans), and the specific sn-positions of fatty acyl chains, which are critical for understanding lipid function in health and disease.

Thermodynamics and Self-Assembly:
The chapter also discusses the thermodynamic principles underlying lipid self-assembly. It explains how noncovalent interactions, including hydrophobic forces, van der Waals contacts, and hydrogen bonding, drive the formation of micelles and bilayers, despite the entropic cost of organizing monomers into ordered structures. The concept of the critical micelle concentration (CMC) is introduced as a key parameter in determining when lipids transition from dispersed monomers to aggregated states.

Applications and Biological Relevance:
Finally, the chapter ties these fundamental concepts to practical applications. Liposomes serve as valuable experimental models for studying membrane dynamics and are extensively used in the encapsulation and targeted delivery of therapeutic agents. The detailed analysis of lipid composition and structure provides insights into membrane function, lipid-mediated signaling, and the molecular basis of various diseases, reinforcing the central theme that structure governs function.

Overall, this chapter integrates principles of lipid chemistry, thermodynamics, and advanced analytical methodologies, offering a comprehensive framework for understanding both the physical behavior of lipid assemblies and their critical roles in biological systems.

This page titled 12.3: Working with Lipids is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **10.4: Working with Lipids** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.4: Transport in Membranes

Source: BiochemFFA_3_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Movement of materials across membranes

As noted earlier, it is essential for cells to be able to uptake nutrients. This function along with movement of ions and other substances is provided by proteins/protein complexes that are highly specific for the compounds they move.

Selective movement of ions by membrane proteins and the ions' extremely low permeability across the lipid bilayer are important for helping to maintain the osmotic balance of the cell and also for providing for the most important mechanism for it to make ATP - the process of oxidative phosphorylation.

Terminology

A protein involved in moving only one molecule across a membrane is called a uniport (Figure 3.25). Proteins that move two molecules in the same direction across the membrane are called symports (also called synporters, synports, or symporters). If two molecules are moved in opposite directions across the bilayer, the protein is called an antiport. Proteins involved in moving ions are called ionophores.

If the action of a protein in moving ions across a membrane results in a net change in charge, the protein is described as electrogenic and if there is no change in charge the protein is described as electroneutral (Figure 3.26). When the driving force for movement through the membrane protein is simply diffusion, the process is called facilitated diffusion or passive transport and when the process requires other energy input, the process is called active transport.

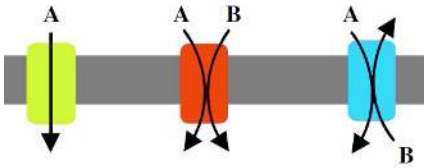


Figure 3.25 - A uniport, a symport, and an antiport

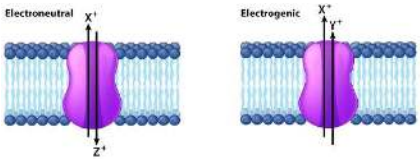


Figure 3.26 - Electroneutral and electrogenic transporters - Image by Aleia Kim

Channels and transporters

With respect to movement of materials through membrane proteins, there is a difference between channels (sometimes called pores) and transporters. Channels largely provide openings with some specificity and molecules pass through them at close to the rate of diffusion. They usually involve movement of water or ions. Examples would be the sodium or potassium channels of nerve cells. Transporters have high specificity and transfer rates that are orders of magnitude slower. Transport proteins include the sodium-potassium pump, the sodium-calcium exchanger, and lactose permease, amongst many others).

Facilitated diffusion

As noted, the driving force for facilitated diffusion is concentration, meaning that in facilitated diffusion, materials will only move from a higher concentration to a lower concentration and that at the end of the process, the concentration of materials on each side of a bilayer will be equal (Figure 3.28). This may work well in many cases.

For example, the blood concentration of glucose is sufficiently high that red blood cells can use facilitated diffusion as a means of acquiring glucose. Other cells, further removed from the blood supply where the glucose concentration is lower, must use active transport mechanisms because there is not a sufficient concentration of glucose to provide cells with the glucose they need.

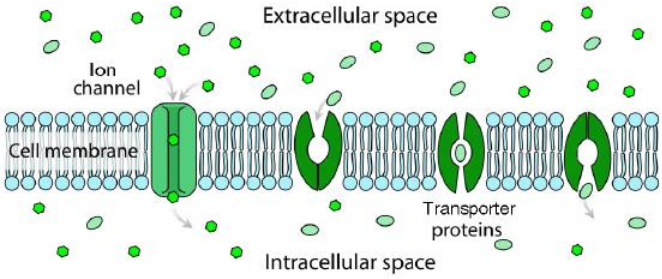


Figure 3.27 - Ion channel and transporter proteins

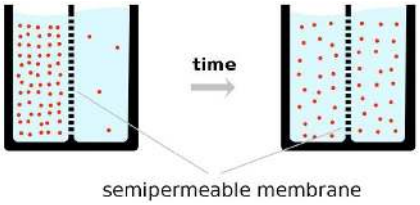


Figure 3.28 - In diffusion, solutes move from high concentration to lower concentration and eventually equalize

Ion channels

Ion channels are pore-forming membrane proteins in the membranes of all cells that regulate movement of selected ions across a membrane (Figures 3.29 & 3.30). They help to establish the resting membrane potential and to affect action potentials and other electrical signals. They are very important in the process of nerve transmission. Ion channels control the flow of ions across secretory and epithelial cells, and consequently help to regulate cell volume by affecting osmotic pressure.

Ion channels are essential features of almost all cells, functioning as selective “tunnels” that restrict movement through them to ions with specific characteristics (typically size). The size of the opening is very narrow (usually one or two atoms wide) and is able to select even against ions that are too small.

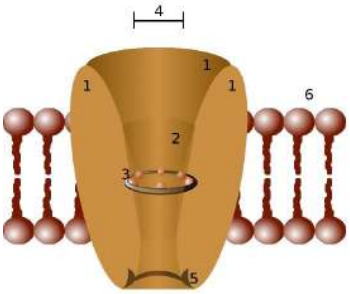


Figure 3.29 - Schematic structure of an ion channel protein - 1 = General structure / 2 = Entrance / 3 = Selectivity filter / 4 = Effective diameter / 5 = Controllable opening

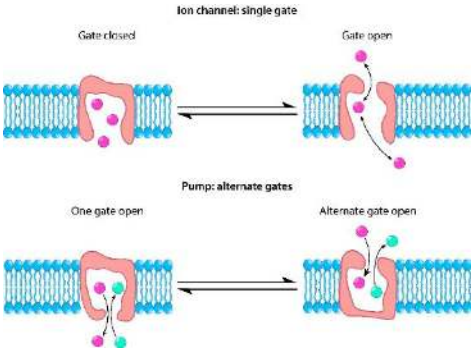


Figure 3.30 - Two types of ion channel - single and double gates - Image by Pehr Jacobson

Control mechanisms

Ion channels are controlled by mechanisms that include voltage, ligands, light, temperature, and mechanical deformation (stretch activated). Ligand-gated ion channels (LGICs) are transmembrane proteins which open to selectively allow ions such as Na^+, K^+, Ca^{++}, or Cl^- to pass through the membrane in response to the binding of a ligand messenger.

Sound waves cause mechanical deformation of hair cells in the ear. This results in the opening of ion channels and initiation of a nerve signal to the brain.

Sodium ion channels in the tongue for sugar receptors open in response to binding of sucrose, allowing sodium concentration in the nerve cell to increase and initiate a nerve signal to the brain. In this case, the default for the gate is to be closed and it opens in response to binding of a ligand (sucrose).

In light sensing cells of the eye, calcium gates are open by default, but stimulation by light causes them to close, triggering a series of events that result in a signal being sent the brain about the perception of light. Thus, in this case, the stimulus (light) causes an open channel to close.

Moving the other direction, nerve signals originating in the brain travel to muscle tissue and through a complicated set of exchanges, result in the opening of calcium gates of muscle cells, increasing the concentration of calcium and stimulating muscular contraction (see HERE).

Voltage gated channels are essential for transmission of nerve signals, a process discussed in more depth HERE.

Ion movement through channels

The ability of ion channels to select against ions too large is intuitive - the size of the opening in the ion channel simply isn't big enough for a larger ion to fit through the opening. Potassium, for example, passes through sodium channels rarely because the opening is too small.

Potassium channels that are selective for potassium ions must be big enough to allow potassium to enter, but if size were the only selection means, then sodium ions would also readily pass through potassium channels, since sodium ions (0.95 Å) are smaller than potassium ions (1.33 Å). In order for potassium channels to select against sodium ions and favor potassium ions, other considerations come into play.

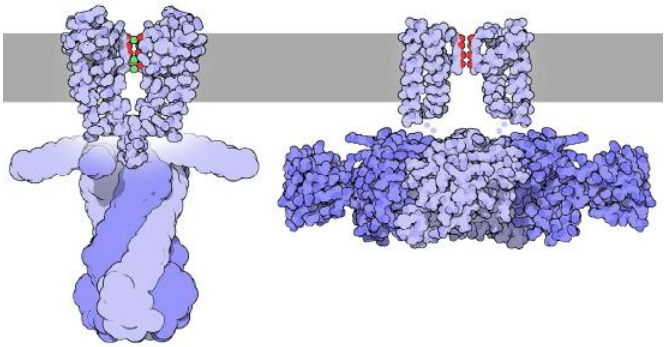


Figure 3.31 - A potassium channel closed (left) and open (right) - David Goodsell - "Molecule of the Month: Potassium Channels". (from Wikipedia)

Hydration shell

To understand this unique selectivity, it is important to understand how ions move through channels. Before an ion can pass through a channel, it must first be dissociated from (stripped of) the water molecules in its hydration shell - water molecules surrounding ions in aqueous solutions (Figure 3.32). This process requires an input of energy. The initial energy required to strip the water molecules from the hydration shell has been compared to the activation energy of an enzymatic reaction.

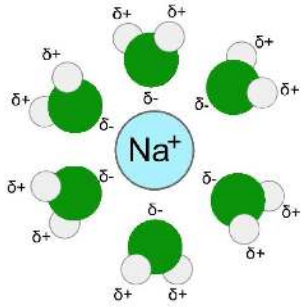


Figure 3.32 - Sodium surrounded by water molecules in a hydration shell

Comparable to enzymes

Just as enzymes lower the activation energy of enzymatic reactions and thus allow them to more readily occur, so too do channel proteins lower the energy requirements for a molecule to traverse a lipid bilayer. In the absence of the channel protein, the dehydration energy is mostly prohibitive for most polar molecules to occur, so very few make it across the lipid bilayer without the channel protein. This is why ion channel/transport proteins are so important to the cell.

After the water has been stripped, the ion can pass through the channel and when it arrives at the other side of the channel, the diffusing ion becomes rehydrated, thus regaining the energy that was required initially to strip away the water molecules from the ion.

Selectivity of the potassium channel

The potassium channel (Figure 3.33) uses the dimensions of the potassium ion precisely to shepherd it through the channel. The sodium ion, which has different dimensions has a more difficult time making it through the channel despite its smaller size. The reason this is rooted in the energy required for dehydration.

For potassium ions, after the water has been stripped off, precisely positioned carbonyl groups along the channel help to stabilize the ion as it moves. The sodium ion, on the other hand is too small and does not make efficient connections with carbonyl groups and thus has a more difficult path. Because of this, the energy difference between dehydration and rehydration of a sodium ion in a potassium channel is energetically unfavorable (requires net input of energy) but the same process for a potassium ion is energetically favorable (results in a net gain of energy).

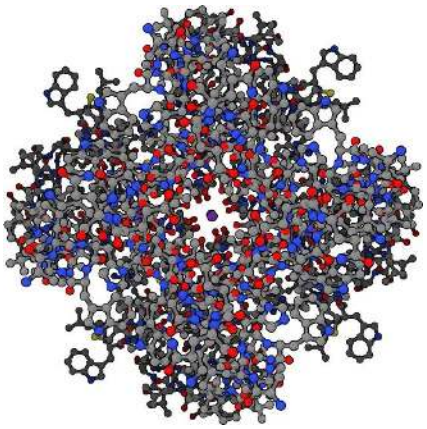
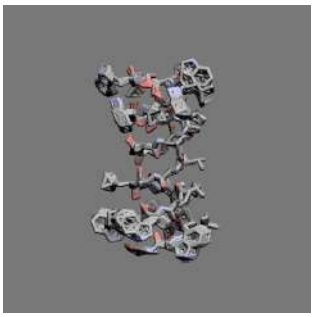


Figure 3.33 - A potassium ion (center) transiting the potassium channel - David Goodsell



Movie 3.1 - Gramicidin A Wikipedia (animated gif, download to view)

Energy factor

Thus the selection in favor of potassium and against sodium ions in a potassium channel is based on energy, not physical size, whereas in the selection of sodium ions over potassium ions in a sodium channel, size is the primary consideration.

Ion balance

The movement of ions across a lipid bilayer is tightly regulated, and with good reason. Maintaining a proper balance of ions inside and outside of cells is important for maintaining osmotic balance. It is also important inside and outside of organelles like the mitochondria and chloroplasts for energy generation. If the ionic balance of a cell is sufficiently disturbed by an uncontrolled ionophore, a cell may die.

Gramicidin

Gramicidins (Movie 3.1) are antibiotic polypeptides synthesized by the soil bacterium known as Bacillus brevis. These small pentadecapeptides (15 amino acids) are synthesized by the bacterium to kill other bacteria.

When released by the Bacillus brevis, the gramicidins insert themselves in the membranes of Gram positive bacteria and allow the movement of sodium ions into the target cells, ultimately killing them. Gramicidins can also cause hemolysis in humans so they cannot be used internally, but instead are used topically.

Aquaporins

Aquaporins are pore-containing integral membrane proteins that selectively permit passage of water molecules in and out of the cell, while preventing ions and other solutes from moving (Figures 3.34 & 3.35). Some aquaporins called aquaglyceroporins, also transport other small uncharged entities, such as glycerol, ammonia, urea, and CO2, across the membrane,. The water pores are completely impermeable to charged molecules, such as protons, which is important for the preserving the membrane's electrochemical potential difference.

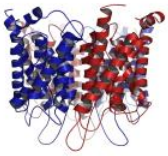


Figure 3.34 - View of aquaporin from the side - Wikipedia

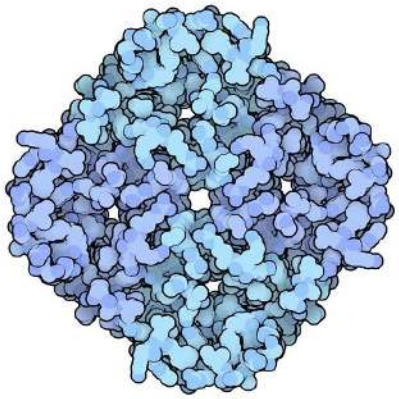


Figure 3.35 - View of aquaporin from the top - David Goodsell

Porins

Porins are proteins containing a β-barrel structure that crosses the cell membrane/wall and acts as a pore/channel through which specific molecules diffuse. Porins are found in the outer membrane of Gram-negative bacteria and some Gram-positive bacteria, mitochondria, and chloroplasts.

Porins typically transport only one group of molecules or, in some cases, one specific molecule. Antibiotics, such as β-lactam and fluoroquinolone pass through porins to reach the cytosol of Gram negative bacteria. Bacteria may develop resistance to these antibiotics when a mutation occurs to the porin involved that results in exclusion of the antibiotics that would otherwise pass through.

Transporter proteins

Not all facilitated transport occurs through ion channel proteins. Transporter proteins, as noted earlier (HERE and Figure 3.27) facilitate movement of materials across a lipid bilayer, but are slower than ion channels. Figure 3.36 illustrates a transporter protein in action. As can be seen, transporter proteins rely on a specific receptor site for proper recognition of the molecule to be moved.

Binding of the proper molecule causes a conformational change in the shape of the protein (an eversion) which results in a flipping of the open side of the protein to the other side of the lipid bilayer. In this way, the molecule is moved. Like ion channels, transporter proteins facilitate movement of materials in either direction, driven only by the concentration difference between one side and the other.

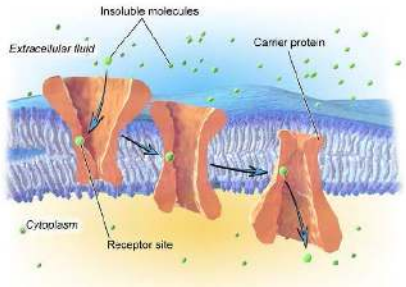


Figure 3.36 - Facilitated diffusion with a transporter protein - Wikipedia

Active transport

All of the transport mechanisms described so far are driven solely by a concentration gradient - moving from higher concentrations in the direction of lower concentrations. These movements can occur in either direction and, as noted, result in equal concentrations on either side of the bilayer, if allowed to go to completion. Many times, however, cells must move materials against a concentration gradient and when this occurs, another source of energy is required. This process is known as active transport.

A good definition of active transport is that in active transport, at least one molecule is being moved against a concentration gradient. A common, but not exclusive, energy source is ATP (see Na+/K+ ATPase), but other energy sources are also employed.

For example, the sodium-glucose transporter uses a sodium gradient as a force for actively transporting glucose into a cell. Thus, it is important to know that not all active transport uses ATP energy.

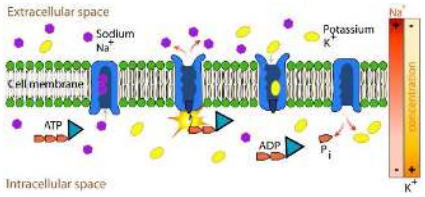


Figure 3.37 - An overview of active transport by the Na+K+ ATPase

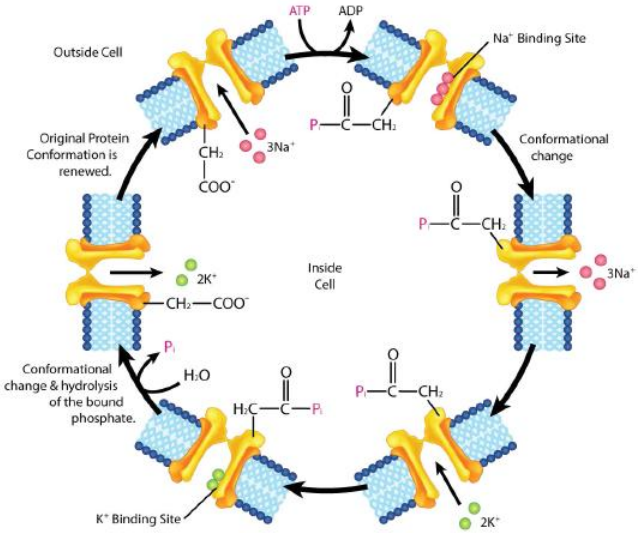


Figure 3.38 - Sequential steps in the active transport of ions by theNa+K+ ATPase - Wikipedia

Na+/K+ ATPase

An important integral membrane transport protein is the Na+/K+ ATPase antiport (Figures 3.37 and 3.38), which moves three sodium ions out of the cell and two potassium ions into the cell with each cycle of action. In each case, the movement of ions is against the concentration gradient. Since three positive charges are moved out for each two positive charges moved in, the system is electrogenic.

The protein uses the energy of ATP to create ion gradients that are important both in maintaining cellular osmotic pressure and (in nerve cells) for creating the sodium and potassium gradients necessary for signal transmission. Failure of the system to function results in swelling of the cell due to movement of water into the cell through osmotic pressure. The transporter expends about one fifth of the ATP energy of animal cells. The cycle of action occurs as follows:

1. Pump binds ATP followed by binding of 3 Na+ ions from cytoplasm of cell
2. ATP hydrolysis results in phosphorylation of aspartate residue of pump. ADP is released
3. Phosphorylated pump undergoes conformational change to expose Na+ ions to exterior of cell. Na+ ions are released.
4. Pump binds 2 extracellular K+ ions.
5. Pump dephosphorylates causing it to expose K+ ions to cytoplasm as pump returns to original shape.
6. Pump binds 3 Na+ ions, binds ATP and releases 2 K+ ions to restart process

The Na+/K+ ATPase is classified as a P-type ATPase. This category of pump is notable for having a phosphorylated aspartate intermediate and is present across the biological kingdoms - bacteria, archaeans, and eukaryotes.

ATPase types

ATPases have roles in either the synthesis or hydrolysis of ATP and come in several different forms.

- F-ATPases (F1FO-ATPases) are present in mitochondria, chloroplasts and bacterial plasma membranes and are the prime ATP synthesizers for these systems. Each uses a proton gradient as its energy source for ATP production. Complex V of the

mitochondrion is an F-type ATPase.

- V-ATPases (V1VO-ATPases) are mostly found in vacuoles of eukaryotes . They utilize energy from ATP hydrolysis to transport solutes and protons into vacuoles and lysosomes, thus lowering their pH values.

The V-type and F-type ATPases are very similar in structure. The V-type (Figure 3.39) uses ATP to pump protons into vacuoles and lysosomes, whereas F-types use proton gradients of the mitochondria and chloroplasts to make ATP.

- A-ATPases (A1AO-ATPases) are found in archaeans and are similar to F-ATPases in function.
- P-ATPases (E1E2-ATPases) are in bacteria, fungi and in eukaryotic plasma membranes and organelles. They transport a diversity of ions across membranes. Each has a common mechanism of action which include autophosphorylation of a conserved aspartic acid side chain within it. Examples of P-type ATPases include the Na+/K+ ATPase and the calcium pump.
- E-ATPases are enzymes found on the cell surface. They hydrolyze a range of extracellular nucleoside triphosphates, including ATP.

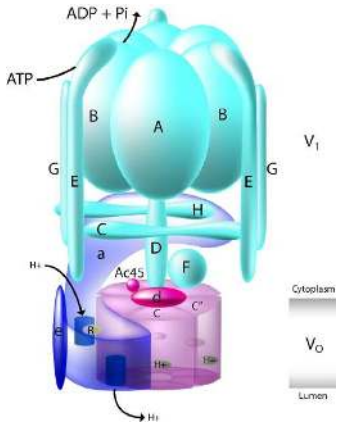


Figure 3.39 - Schematic structure of V-ATPases- Wikipedia

Nerve transmission

Now that you have seen how the Na+/K+ ATPase functions, it is appropriate to discuss how nerve cells use ion gradients created with it to generate and transmit nerve signals. Neurons are cells of the nervous system that use chemical and electrical signals to rapidly transmit information across the body (Figure 3.40). The sensory nerve system links receptors for vision, hearing, touch, taste, and smell to the brain for perception. Motor neurons run from the spinal cord to muscle cells. These neurons have a cell body and a very long, thin extension called an axon, that stretches from the cell body in the spinal cord all the way to the muscles they control. Nerve impulses travel down the axon to stimulate muscle contraction.

Signals travel through neurons, ultimately arriving at junctions with other nerve cells or target cells such as muscle cells. Note that neurons do not make physical contact with each other or with muscle cells. The tiny space between two neurons or between a neuron and a muscle cell is called the synaptic cleft. At the synaptic cleft, the neuron releases neurotransmitters that exit the nerve cell and travel across the junction to a recipient cell where a response is generated. That response may be creating another nerve signal, if the adjacent cell is a nerve cell or it may be a muscular contraction if the recipient is a muscle cell (Figure 3.41).

In considering information movement via nerve cells, then, we will discuss two steps - 1) creation and propagation of a signal in a nerve cell and 2) action of neurotransmitters exiting a nerve cell and transiting a synaptic junction.

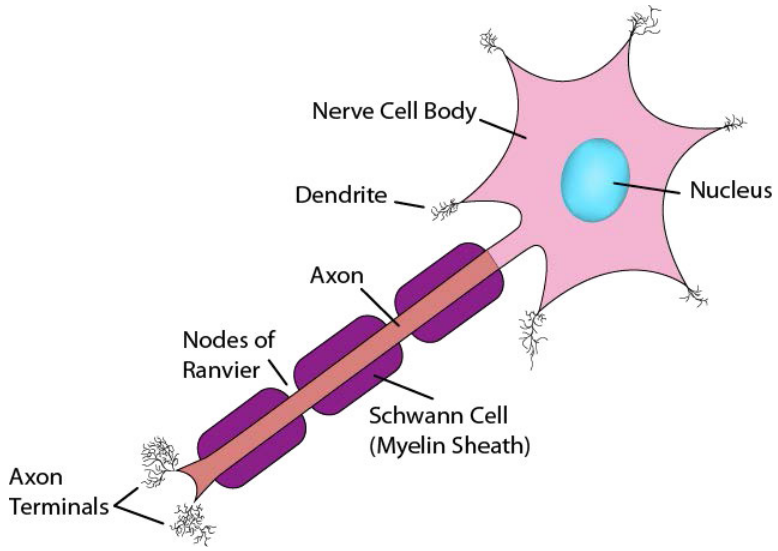


Figure 3.40 - Anatomy of a nerve cell - Image by Pehr Jacobson

Signal source

Creation of a nerve signal begins with a stimulus to the nerve cell. In the case of muscle contraction, the motor cortex of the brain sends signals to the appropriate motor neurons, stimulating them to generate a nerve impulse. How is such an impulse generated?

Resting potential

In the unstimulated state, all cells, including nerve cells, have a small voltage difference (called the resting potential) across the plasma membrane, arising from unequal pumping of ions across the membrane. The Na+/K+ ATPase, for example, pumps sodium ions out of the cell and potassium ions into cells. Since three sodium ions get pumped out for every two potassium ions pumped in, a charge and chemical gradient is created. It is the charge gradient that gives rise to the resting potential.

Altering the gradients of ions across membranes provide the driving force for nerve signals. This happens as a result of opening and closing of gated ion channels. Opening of gates to allow ions to pass through the membrane swiftly changes the ionic balance across the membrane resulting in a new voltage difference called the action potential. It is the action potential that is the impetus of nerve transmission.

Initiation of signal

The signal generated by a motor neuron begins with opening of sodium channels in the membrane of the nerve cell body causing a rapid influx of sodium ions into the nerve cell. This step, called depolarization (Figure 3.42), triggers an electrochemical signal - the action potential. Remember that the Na+/K+ ATPase has created a large sodium gradient, so sodium ions rush into the cell when sodium channels open. After the initial depolarization, potassium channel gates, responding to the depolarization, open, allowing potassium ions to rapidly diffuse out of the cell (remember K+ ions are more abundant inside of the cell). This phase is called the repolarization phase and during it, the sodium gates close.

The rapid exit of potassium ions causes the voltage difference to “overshoot” the resting potential and potassium gates close. This followed by the so-called refractory period, when the Na+/K+ ATPase begins its work to re-establish the original conditions by pumping sodium ions out and potassium ions into the nerve cell. Eventually, the system recovers and the resting potential is re-established. The initiating end of the nerve cell is then ready for another signal.

Propagation of action potential

What we have described here is only the initiation of the nerve signal in one part of the nerve cell. For the signal to be received, the action potential must travel the entirety of the length of the nerve cell (the axon) and cause a chemical signal to be released into the synaptic cleft to get to its target. Propagating the nerve signal (action potential) in the original nerve cell is the function of all of the rest of the gated ion channels (Figure 3.43) positioned on the sides of the nerve cell. The sodium and potassium gates involved in propagation of the signal all act in response to voltage changes created by the electrochemical gradient moving down the nerve cell

(Figure 3.44). Remember that opening of the initial gates at initiation of the signal created an influx of sodium ions and an efflux of potassium ions.

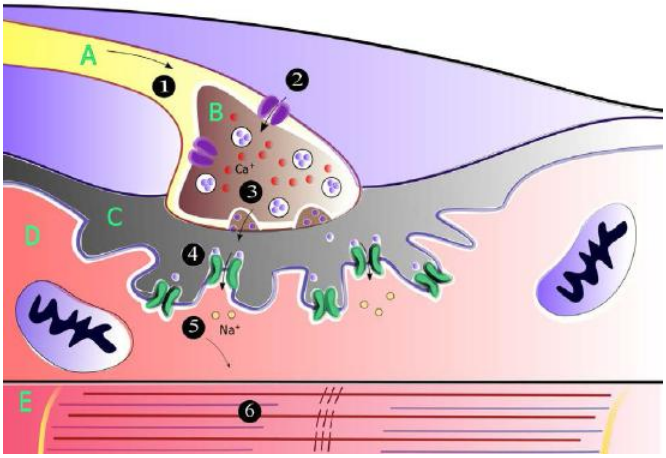


Figure 3.41 - (1) The action potential reaches the axon terminal; (2) voltage-dependent calcium gates open - calcium enters the axon terminal; (3) neurotransmitter vesicles fuse with the presynaptic membrane and release acetylcholine (ACh); (4) ACh binds to postsynaptic receptors on the sarcolemma; (5) ion channels open in response and allow sodium ions to flow into the muscle cell; (6) The flow of sodium ions into the muscle cell generates an action potential which travels to the myofibril and results in muscle contraction. - Wikipedia

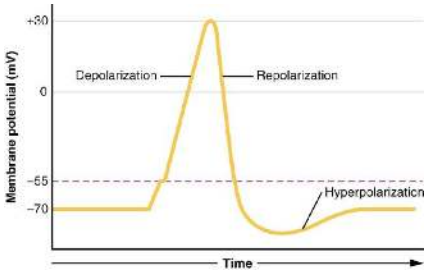


Figure 3.42 - Depolarization and repolarization of a nerve cell - Wikipedia

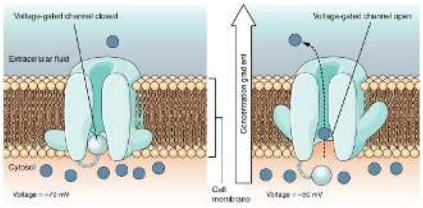


Figure 3.43 - Voltage gated ion channels - Wikipedia

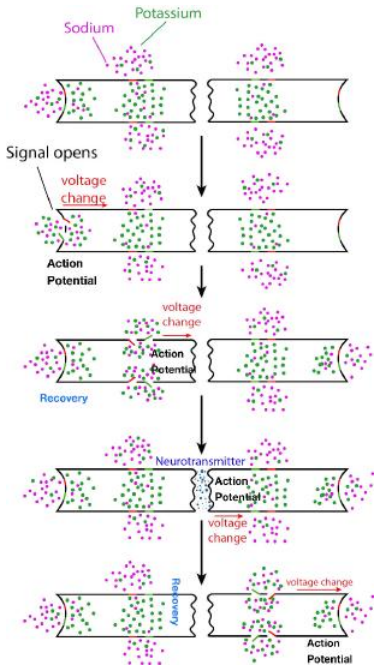


Figure 3.44 - Action potential moving across a synaptic junction of two nerve cells - Image by Pehr Jacobson

Moving signal

This chemical and electrical change that creates the action potential leaves the end of the nerve cell where it started and travels down the axon towards the other end of the nerve cell. Along the way, it encounters more sodium and potassium gated channels. In each case, these respond simply to the voltage change of the action potential and open and close, exactly in the same way the gates opened to start the signal. Thus, a rapid wave of increasing sodium ions and decreasing potassium ions moves along the nerve cell, propagated (and amplified) by gates opening and closing as the ions and charges move down the nerve cell. Eventually, the ionic tidal wave reaches the end of the nerve cell (axon terminal) facing the synaptic cleft.

Crossing the synaptic cleft

For the signal to be received by the intended target (postsynaptic cell) from the originating neuron (presynaptic neuron), it must cross the synaptic cleft and stimulate the neighboring cell (Figure 3.45). Communicating information across a synaptic cleft is the job of neurotransmitters. These are small molecules synthesized in nerve cells that are packaged in membrane vesicles called synaptic vesicles in the nerve cell. Neurotransmitters come in all shapes and chemical forms, from small chemicals like acetylcholine to peptides like neuropeptide Y. The most abundant neurotransmitter is glutamate, which acts at over 90% of the synapses in the human brain.

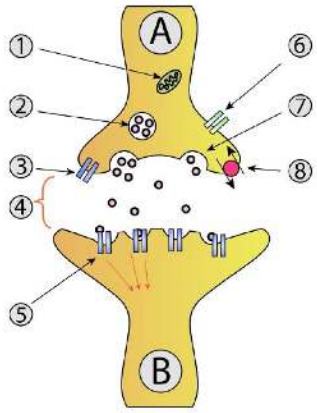


Figure 3.45 - A) Pre-synaptic neuron; B) post-synaptic neuron; 1) Mitochondria; 2) synaptic vesicle with neurotransmitters; 3) autoreceptor; 4) synapse with neurotransmitter released (serotonin); 5) postsynaptic receptors activated by neurotransmitter; 6) calcium channel; 7) exocytosis of a vesicle; 8) recaptured neurotransmitter. Wikipedia

Movie 3.2 - Movement of an action potential down a nerve cell - Wikipedia

Into the cleft

As the action potential in the presynaptic neuron approaches the axon terminus, synaptic vesicles begin to fuse with the membrane and their neurotransmitter contents spill into the synaptic cleft. Once in the cleft, the neurotransmitters diffuse, some of them reaching receptors on the postsynaptic cell. Binding of the neurotransmitter to the receptors on the membrane of the postsynaptic cell stimulates a response.

For motor neurons, the postsynaptic cell will be a muscle cell, and the response will be muscle contraction/relaxation. At this point, the originating nerve cell has done its job and communicated its information to its immediate target. If the postsynaptic cell is a nerve cell, the process repeats in that cell until it gets to its destination.

Na+/glucose transporter

Absorbing nutrients from the digestive system is necessary for animal life. The sodium/glucose transport protein is an electrogenic symporter that moves glucose into intestinal cells. It is found in the intestinal mucosa and the proximal tubule of the nephron of the kidney. The sodium/glucose transport system functions in the latter to promote reabsorption of glucose.

The pump works in conjunction with the Na+/K+ transport system. The gradient of sodium ions built up by the Na+/K+ pump is used as an energy source to drive movement of glucose into cells (see Figure 3.38). Use of an ion gradient established by a separate pump is known as secondary active transport. For intestinal mucosa, the pump transports glucose out of the gut and into gut cells. Later, the glucose is exported out the other side of the gut cells to the interstitial space for use in the body.

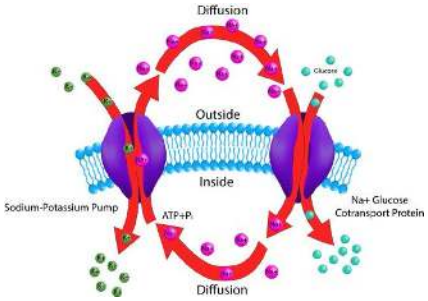


Figure 3.46 - The sodium/glucose pump (right) and the Na+/K+ ATPase (left). The sodium gradient generated by the Na+/K+ ATPase is used by the sodium/glucose pump to import glucose into the cell. Wikipedia

Calcium pumps

Calcium ions are necessary for muscular contraction and play important roles as signaling molecules within cells. In addition, when calcium concentrations rise too high, DNA in chromosomes can precipitate. Calcium concentration in cells is therefore managed carefully. It is kept very low in the cytoplasm as a result of action of pumps, both in the plasma membrane, which pump calcium outwards from the cytoplasm and in organelles, such as the endoplasmic reticulum (sarcoplasmic reticulum of muscle cells), which pump calcium out of the cytoplasm and into these organelles.

Opening of calcium channels, then, increases calcium concentration quickly in the cytoplasm resulting in a quick response, whether the intention is signaling or contraction of a muscle. After the response is generated, the calcium is pumped back out of the cytoplasm by the respective calcium pumps.

Some calcium pumps use ATP as an energy source to move calcium and others use ion gradients, such as sodium for the same purpose.

Na+/Ca++ transporter

One calcium pump of interest uses the sodium gradient as an energy source. It is the sodium/calcium pump. This electrogenic antiport system uses sodium's movement into the cell as a driving force to move calcium out of the cell, although its direction can reverse in some circumstances. The pump is a high capacity system to move a lot of calcium quickly, moving up to 5000 calcium ions per second and is found in many tissues with many functions.

Digitalis

One important function of the Na+/Ca++ pump occurs in heart cells. Ca++ is important for contraction of heart muscle. Calcium efflux from the cells is the normal operation of the pump, however, during the upstroke of the cycle, there is a large movement of sodium ions into the heart cell. When this occurs, the pump reverses and pumps Na+ out and Ca++ in briefly. Since calcium helps stimulate contraction of cardiac muscle, this can help make the heart beat stronger and is the focus of the use of digitalis to treat congestive heart failure.

Digitalis blocks the sodium-potassium ATPase and interferes with the sodium ion gradient. As noted above, when the Na+ gradient is oriented in the wrong direction, calcium is pumped in. Digitalis is therefore used to treat congestive heart failure because it increases the concentration of calcium in the heart cells, favoring more forceful beats.

ABC transporters

ABC transporters are another class of transmembrane proteins that use ATP energy to transport things against concentration gradients (Figures 3.47 & 3.48). This protein superfamily includes hundreds of proteins (48 in humans alone) and spans all extant phyla from prokaryotes to humans. These proteins function not only in membrane transport, but also in processes that include DNA repair and the process of translation.

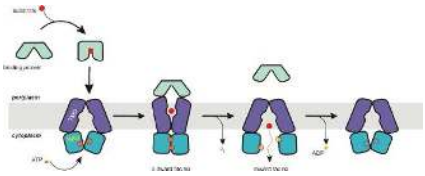


Figure 3.47 - Mechanism of an ABC importer

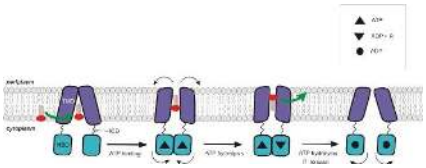


Figure 3.48 - Another view of an ABC exporter. NBD = Nucleotide Binding Domain

Transport

Substances that ABC transporters move across membranes include metabolic products, lipids, sterols, and drugs. ABC transporters function in multi-drug resistance of many cells, and provide resistance to antibiotics in bacteria as well as resistance to chemotherapy in higher cells by exporting drugs used to treat both of these types of cells.

ABC transporters are divided into three main groups - 1) importers (prokaryotes only); 2) exporters (prokaryotes and eukaryotes), and 3) non-transporters with roles in DNA repair and translation. All ABC transport proteins have four protein domains - two that are cytoplasmic and two that are membrane bound. They are alternately open to the cytoplasmic or extracellular (or periplasmic) regions and this is controlled by hydrolysis of ATP.

Disease

ABC transporters have roles in cystic fibrosis and other inherited human diseases. They are very involved in development of resistance to multiple drugs by a diverse group of cells. ABC transporters provide multi-drug resistance by expelling drug(s) from cells. ABCB1 protein, for example, pumps tumor suppression drugs out of the cell. Another ABC transporter known as Pgp transports organic cationic or neutral compounds.

Cystic fibrosis

Cystic fibrosis (CF) is a an autosomal recessive genetic disorder arising from mutations in both copies of the gene for the cystic fibrosis transmembrane conductance regulator (CFTR) protein. This ABC transporter system, which moves chloride and thiocyanate ions across epithelial tissue membranes exerts its effect mostly in the lungs but the pancreas, liver, kidneys, and intestine are all also affected by it.

Function

CFTR has roles in the production of sweat, mucus, and digestive fluids. Manifestations of the disease include breathing difficulty and overproduction of mucus in the lungs. When CFTR is functional, these fluids are normally thin, but when the gene is non-functional, they become much thicker and are points of infection.

CFTR contains two ATP-hydrolyzing domains and two cell membrane-crossing domains with 6 α-helices each. It can be activated by phosphorylation by a cAMP-dependent protein kinase. The carboxyl end of CFTR is linked to the cytoskeleton by a PDZ domain.

Lactose permease

Another integral membrane protein performing active transport is lactose permease. It facilitates the movement of the sugar lactose across the lipid bilayer of the cell membrane (Figures 3.49- 3.51). The transport mechanism is classified as a secondary active transport since it exploits the inwardly directed H+ electrochemical gradient as an energy source. When lactose is transported into cells, it is broken down into its substituent monosaccharide sugars - glucose and galactose - for energy creation.

The enzyme catalyzing this reaction is known as lactase and deficiency of it in humans leads to lactose intolerance (see HERE).

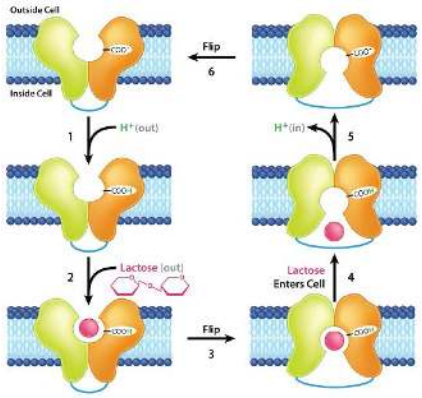


Figure 3.49 - Mechanism of action of lactose permease - Note that the proton acts by altering the charge of a carboxyl group - Image by Aleia Kim

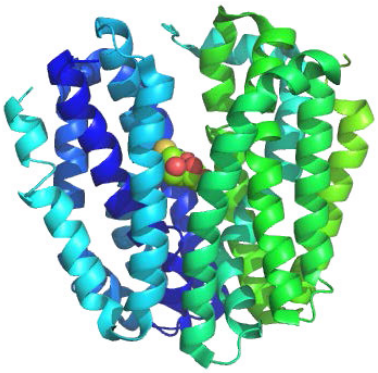


Figure 3.50 - Structure of lactose permease - Wikipedia

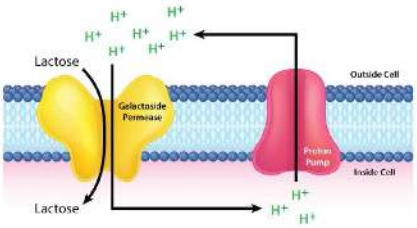


Figure 3.51 - Lactose (galactoside) permease is a secondary transporter, using a proton gradient as an energy source to pump lactose into the cell - Image by Aleia Kim

GLUTs

GLUTs (GLUcose Transport proteins) are uniport, type III integral membrane proteins that participate in the transport of glucose across membranes into cells. GLUTs are found in all phyla and are abundant in humans, with 12 GLUT genes. GLUT1, in erythrocytes is well-studied. Through GLUT 1, glucose enters and passes through it via facilitated diffusion at a rate that is 50,000 higher than in its absence. GLUTs of various types are found in different cells of the body. The one in red blood cells is known as GLUT 1 and has 12 membrane-spanning hydrophobic helices.

Though the structure of GLUT 1 is not known, it is speculated that the 12 helices form a chamber able to form hydrophilic bonds with glucose to facilitate its passage.

GLUT 1 levels in erythrocytes go up as glucose levels decrease and decrease when glucose levels go down. GLUT 1 can also transport ascorbate (vitamin C) in addition to glucose in mammals (such as humans) that do not produce their own vitamin C.

Glut 4

GLUT 4 is regulated by insulin and is found primarily in adipose and striated muscle tissue. Insulin alters intracellular trafficking pathways in response to increases in blood sugar to favor movement of various GLUT proteins (including GLUT 4) from intracellular vesicles to the cell membrane, thus stimulating uptake of the glucose. GLUT 4 is also found in the hippocampus where, if trafficking is disrupted, the result can be depressive behavior and cognitive dysfunction.

For all of the GLUT proteins, a key to keeping the glucose in the cell is phosphorylation of it by the glycolysis enzyme, hexokinase, in the cytoplasm. Phosphorylated molecules cannot enter GLUTs and don't have an easy means of exiting the cell.

Glucose Transporters

Name	Tissue	K_M (μM)	Comments
GLUT1	All mammalian tissues	1	Basal glucose uptake
GLUT2	Liver and pancreatic β cells	15-20	Participates in insulin regulation in the pancreas; removes excess glucose from the blood in the liver
GLUT3	All mammalian tissues	1	Basal glucose uptake
GLUT4	Muscle and fat cells	5	Concentration in plasma membrane of muscles rises with endurance training
GLUT5	Small intestine	---	Fructose transport

Figure 3.52 - GLUTs found in cells - Image by Aleia Kim

This page titled 12.4: Transport in Membranes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Transport in Membranes** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.
- **9.3: Transport in Membranes** by Kevin Ahern, Indira Rajagopal, & Taralyn Tan is licensed CC BY-NC-SA 4.0. Original source: http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy.

12.5: Problems

10.1 Questions

This page titled 12.5: Problems is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **10.5: Problems** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.6: Membranes and Membrane Proteins

Learning Goals (ChatGPT o3-mini, 2/2/25)

1. **Explain the Functional Role of Membrane Bilayers:**
   • Describe how membrane bilayers not only serve as physical barriers separating cellular compartments but also provide platforms for signal transduction and molecular trafficking.
2. **Differentiate Types of Membrane Proteins:**
   • Compare and contrast peripheral (extrinsic) and integral (intrinsic) membrane proteins, including the concepts of monotopic, bitopic (single-pass), and polytopic (multi-pass) proteins, and explain their respective modes of association with the lipid bilayer.
3. **Describe Mechanisms of Membrane Protein Localization:**
   • Explain how various post-translational modifications (e.g., N-myristoylation, S-palmitoylation, prenylation, and glycosylphosphatidylinositol (GPI) anchoring) target proteins to membranes and influence their function and interactions.
4. **Analyze Protein-Lipid Interactions:**
   • Understand how peripheral proteins such as MMP-12 and prothrombin interact with lipid bilayers via electrostatic interactions (e.g., Ca^{2+}-mediated binding of GLA domains) and hydrophobic patches, and how these interactions affect their catalytic activity.
5. **Explore Structural Classes of Integral Membrane Proteins:**
   • Describe the structural features of transmembrane proteins, including alpha-helical polytopic proteins (such as GPCRs and kinases) and beta-barrel proteins (such as OMF and VDAC), and discuss how their structure supports their function in transport and signal transduction.
6. **Examine Membrane Protein Models and Mimetics:**
   • Evaluate the different experimental systems used to study membrane proteins, including detergent micelles, bicelles, nanodiscs, and SMALPs, and understand how these systems mimic the native lipid bilayer environment.
7. **Discuss Protein Complexes Involved in Cell Adhesion and Signaling:**
   • Analyze the roles of adhesive proteins such as cadherins and dystroglycan in maintaining cell–cell interactions and connecting the extracellular matrix to the cytoskeleton, as well as their implications in disease when mutated.
8. **Understand the Role of Membrane Proteins in Signal Transduction:**
   • Explain how transmembrane receptors, including GPCRs and receptor tyrosine kinases, transduce extracellular signals into intracellular responses through conformational changes and downstream signaling cascades.
9. **Interpret Structural Data from Molecular Models:**
   • Utilize interactive 3D molecular visualization tools (e.g., iCn3D) to analyze the architecture of membrane proteins and their interactions with lipids, highlighting how structural insights can inform understanding of protein function and pharmacological targeting.
10. **Relate Membrane Protein Function to Therapeutic Applications:**
    • Discuss how targeting membrane-associated enzymes and transporters (e.g., P-glycoprotein, Ras proteins) can influence disease outcomes, particularly in cancer and other signal transduction disorders, and explore strategies for drug design based on membrane protein structures.

These learning goals will help you integrate theoretical knowledge with practical examples, providing a comprehensive understanding of how membrane proteins contribute to cellular function and how their structural features are harnessed for biological regulation and therapeutic intervention.

12.6.1: Introduction

One easily understandable function of membrane bilayers is to separate the inside and outside of the cell or intracellular organelles. Yet, as we mentioned before, such barriers can not be so rigid and impenetrable that they prevent the movement of materials across the membrane. Also, all cells must sense and respond to their environment through signal transduction. We have already discussed lipid molecules involved in signaling. Now, let's turn our attention to proteins associated with the membrane and confer added functionalities. Figure **12.6.1** reviews some of the features of membranes we've discussed before and shows a simple bilayer (top) to the complicated membrane/cell wall of bacteria.

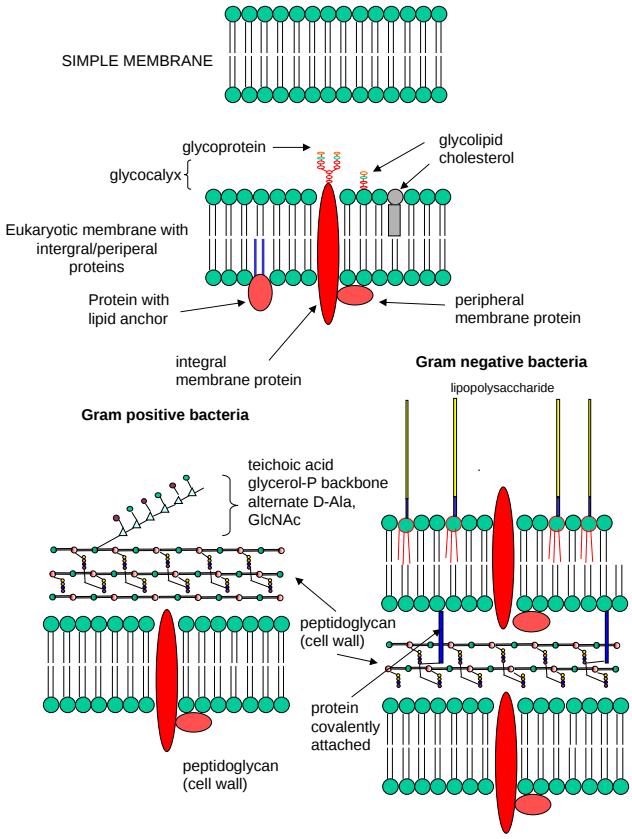


Figure **12.6.1**: Membrane from simple to complex

12.6.2: Types of Membrane Proteins

Although we presented this image earlier, Figure **12.6.2** reviews the details that should now be clearer. In this section, we will explore membrane proteins in more detail.

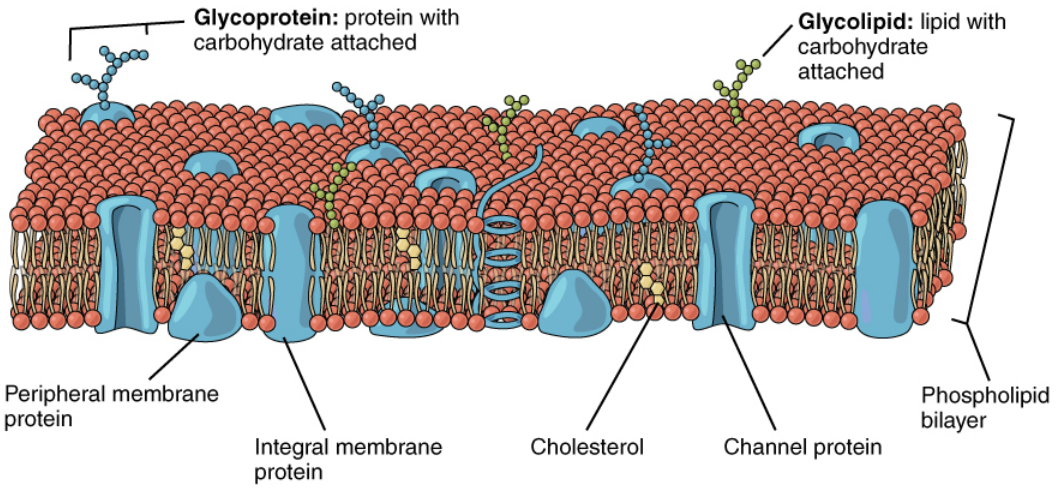


Figure 12.6.2: A cell membrane with peripheral and integral membrane proteins. https://commons.wikimedia.org/wiki/F...Components.jpg. This file is licensed under the Creative Commons Attribution 4.0 International license.

Proteins can be loosely associated with the membrane (**peripheral or extrinsic**), embedded deeply, or pass through the membrane and become a transmembrane (also called i**ntegral or intrinsic)** protein. Sometimes, they pass through using a single alpha helix, while others pass through multiple times (for example, seven times in G-protein coupled receptors). They can also be classified based on the number of membrane leaflets they cross, as shown in Figure 12.6.3.

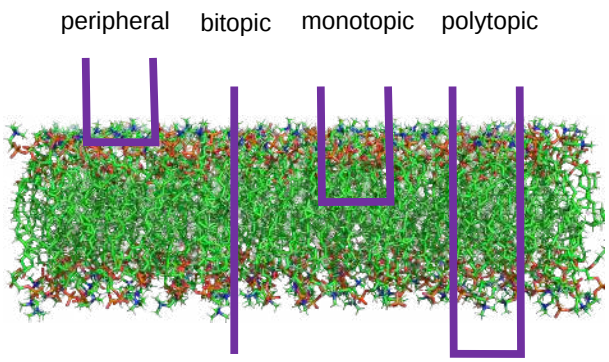


Figure 12.6.3: Peripheral and biotopic, monotopic, and polytopic integral membrane proteins

12.6.3: Peripheral Proteins

The proteins interact with a membrane through protein-lipid head group interactions, but might slightly penetrate the membrane. Those that do would be classified as monotopic peripheral. Peripheral proteins are generally easy to remove from a membrane in vitro by changing the solution ion concentration, as the interactions are often of an ion-ion nature. The first model below shows a matrix Metalloproteinase (MPP) 12 binding to a lipid bilayer. This protein is involved in inflammation, wound healing, arthritis, cardiovascular disease, and remodeling of neural synapses, suggesting a broad role in recovery from cell and tissue aberrations. Macrophages secrete MMP-12, so it is considered a water-soluble (aqueous) protein. It travels to viral cells and appears to display activity not in aqueous solution but near membranes, implying activation of the enzyme through binding to the bilayer.

Studies show the catalytic domain of MMP 12 can bind bilayers through both α- and β-secondary structure regions of the protein. Figure 12.6.4 shows an interactive iCn3D model of the protein and its interaction with the membrane through the alpha-helical region. Once bound to the membranes, catalytic activity increases.

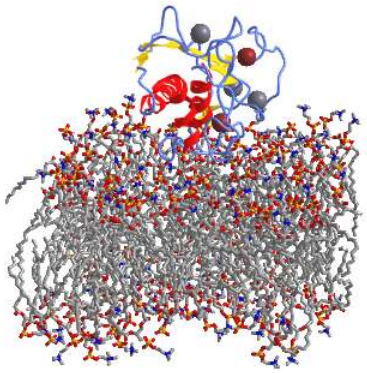


Figure 12.6.4: Catalytic domain of peripheral membrane Protein MMP 12 interacting with a bilayer. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wyF6WvRMx9JP6A

Other MMPs localize to membranes in other ways. MMP-7 interacts with heparan sulfate proteoglycans (CD44) and lipid rafts. Others bind transmembrane proteins like integrins.

Other examples of peripheral proteins include many precursor forms of protein clotting factors. Clotting is initiated when the serine protease thrombin cleaves fibrinogen to form fibrin, which self-associates to form a fibrin clot, or when thrombin activates platelet receptors. The soluble precursor of thrombin, prothrombin, a zymogen, is activated on membrane binding through interactions with several proteins assembled on a negatively charged phospholipid (like phosphatidylserine) bilayer in the prothrombinase complex. How does the precursor zymogen interact with the membrane? It requires calcium ions, which bind to a series of **gamma-carboxylated glutamic acid (GLA)** residues on the zymogen. The enzyme that carboxylates the zymogen depends on Vitamin K.

Figure 12.6.5 shows an interactive iCn3D model of bovine prothrombin Fragment 1 (N terminal) bound to a bilayer through its GLA domain (1NL2).

Figure 12.6.5: Bovine prothrombin Fragment 1 (N terminal) bound to a bilayer through its GLA domain (1NL2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...SG6qByvpAsvwe6

The Gla sidechains are shown in CPK-colored sticks and interact with Ca^{2+} ions (gray spheres). Click on this link to see a zoomed view of just the calcium ions and Gla sidechains: https://structure.ncbi.nlm.nih.gov/i...xydtj8a4WHgdb7

The Gla domain, in the absence of calcium ions, is disordered. On binding, an ordered linear alignment of bound calcium ions is formed, stabilizing the ordered structure of the Gla domain and allowing interaction with the membrane. Three nonpolar amino acid side chains, Phe 5, Leu 6, and Val 9, are now clustered and exposed, allowing penetration of this hydrophobic patch partway into the membrane. They are represented in cyan spacefill just underneath the surface of the red dots in the model above (the red dots are dummy atoms that represent the outer bilayer leaflet). Given this penetration, this protein domain would then be considered monotopic.

What is not shown in the model is the role of negatively charged phosphatidyl serine. Studies have shown that the head group of serine in lysophosphatidylserine (which has only one acyl group) provides additional ion-ion interactions with the Ca^{2+} ions that bind Gla residues 17 and 21. Arg 10 and Arg 16 also interact with the phosphatidyl serine head group. Phosphatidylcholine could also spatially fit into the active site, but electrostatic interactions would prevent it. Why?

12.6.4: Lipid-Anchored Proteins

We have studied lipids, proteins, and carbohydrates. Although phospholipids can spontaneously form bilayers, the actual structure of biological membranes is made much more complicated by adding protein and carbohydrate substituents to the membrane. Soluble proteins can be made to insert into bilayers by the addition of nonpolar attachments. Localization to a membrane changes the functional expression of the protein. Several examples of such attachments are described below.

12.6.4.1: Fatty acid linkers

Two common covalent modifications of proteins are N-myristoylation (attached myristic acid - 14:0 - through an amide link) and S-palmitoylation (attached palmitic acid - 16:0 - through a thioester link with a Cys).

Myristoylation is usually a cotranslational modification in eukaryotic and viral proteins after cleavage of the N-terminal methionine. Figure **12.6.6** shows an image of the serine/threonine phosphatase 2C (1A6Q) with its N-terminal glycine myristoylated. It should be obvious how this post-translationally modified protein interacts with a membrane.

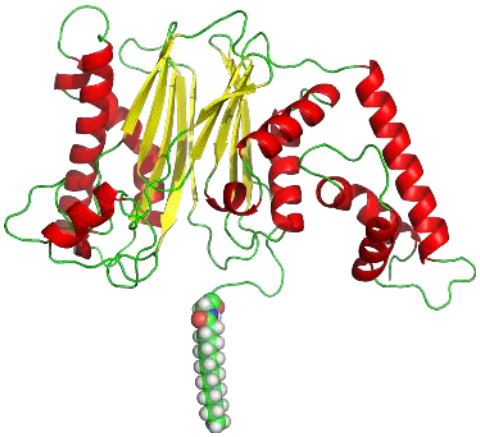


Figure **12.6.6**: The serine/threonine phosphatase 2C (1A6Q) with its N-terminal glycine myrisoylated

This modification is a key part of initiating immune system signal transduction pathways. The modification is catalyzed by N-myristoyltransferase (NMT) using myristoyl-coenzyme as the fatty acid acyl donor. This activates the protein's function partly by reducing the dimensionality of substrate diffusion to the protein to the 2D surface of the membrane instead of a 3D search in the cytoplasm. NMT acylates protein at this consensus sequence: $G^1X^2X^3X^4S/T^5X^6R^7R^8$.

Likewise, many signaling proteins are palmitoylated, leading to protein recruitment to membranes. Small G proteins like Ras, Rho, and the alpha subunit of heterotrimeric G proteins are often palmitoylated. This modification can also be found in transmembrane proteins where localization is not an issue (see example xx below). In such circumstances, the modification might help target the proteins to rafts within the membrane. Palmitic acid is saturated, and attaching it to a protein might target the modified protein to more ordered membrane regions with cholesterol and sphingolipids within rafts.

12.6.4.2: Isoprenoids linkers:

The isoprenoids farnesyl (15C) or geranylgeranyl (20C)are added to a CAAX carboxy-terminal sequence in a target protein like RAS, where C is Cys, A is aliphatic, and X is any amino acid, which helps target proteins to the membrane. The enzymes used for these modifications are farnesyltransferase (FTase) and protein geranylgeranyltransferase I (GGTase I), respectively. This and the other modifications have the potential to do more than target proteins to the membrane. The modification can also modulate protein-ligand interactions and protein stability. Ras, a key signaling protein, is a target of prenylation.

Ras and other small G proteins are involved in many human cancers. As the G protein Ras has a somewhat billiard ball-like surface with obvious sites to target drugs that would affect its aberrant function in cancers, efforts have been made to target the prenyltransferases necessary to target it to the membrane.

In humans, there are three genes in the Ras family, H-Ras and N-Ras, whose gene products localize to both plasma and Golgi membranes, and K-Ras, which localizes predominantly to the plasma membrane. These and other G proteins bind GTP and possess

GTPase activity. The GTP-bound form is active, while the GDP form is inactive. Point mutations that attenuate or prevent GTP cleavage leave the protein continually activated, contributing to oncogenesis.

KRas has two predominant isoforms, 4A, the canonical form (also called 2A), and 4B (also called 2B), that arise from alternative splicing of the primary RNA transcript. The C-terminal protein sequences of isoforms 4A and 4B differ significantly.

Isoform 4A: QYRLKKISKEEKTPGCVKIKKCIIM

Isoform 4B: KHKEKMSKDGKKKKKKSKTKCVIM

The farnesylation motif site containing the modified Cys are highlighted in yellow above. The same cysteine is also often carboxymethylated. The Cys six residues from the farnesylated Cys in isoform 4A are also often palmitoylated.

Figure 12.6.7 shows isoform KRas 4B bound to a membrane bilayer through its farnesylated tail (PDB file provided by Gorfe, Nair, and McCammon). The tail is essential for its function at the plasma membrane, where KRAS-mediated signaling events occur. Phosphodiesterase-δ (PDEδ) binds to KRAS4b and is important in targeting it to cellular membranes. Note that the farnesyl attachments only penetrate part of the upper leaflet.

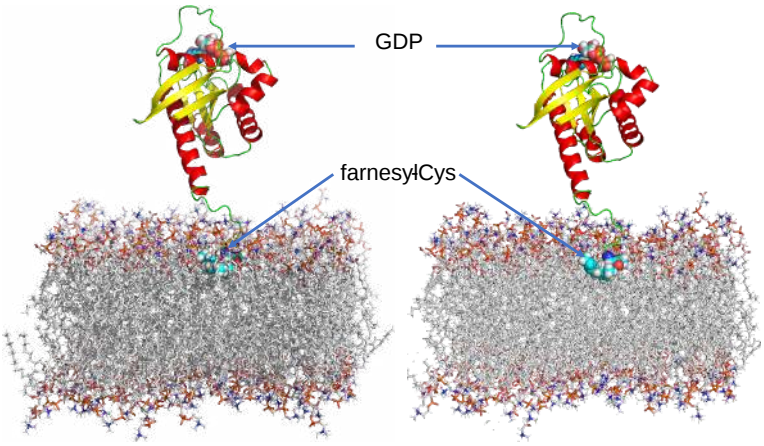


Figure 12.6.7: Isoform KRas 4B bound to a membrane bilayer through its farnesylated tail

12.6.4.3: Glycosyl-phosphatidylinositol linkers

Soluble cytosolic proteins can attach to membranes normally by adding glycosylphosphatidylinositol (GPI). The attachment usually contains a conserved tetrasaccharide core of three mannoses (Man) and one unacetylated glucosamine (GlcN) linked to the carboxy terminus of the protein. The GPI can be further modified with extra galactoses and mannoses and additions to the PI group, which secures the protein in the membrane. Figure 12.6.8 shows the common backbone for GPI anchors. Note the additions of the phosphoethanolamines to the core polysaccharide.

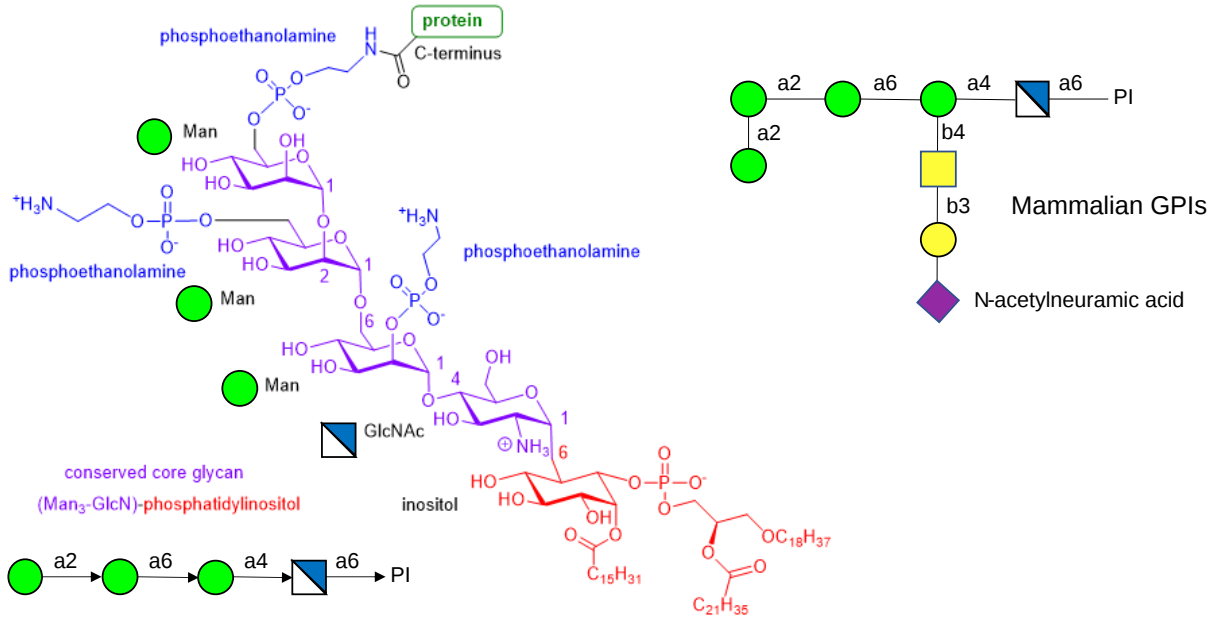


Figure 12.6.8: Common backbone of GPI anchors

GPIs in eukaryotic cells link many surface antigens, adhesion molecules, and hydrolases to the membrane. GPIs from Plasmodium falciparum, the malarial parasite that kills about two million people yearly, appear to act as a toxin and are the most common CHO modification of the parasite protein. Mice immunized against the GPI sequence, NH_2-CH_2-CH_2-PO_4-Man (α1-2) 6Man (α1-2) Man (α1-6) Man (α1-4) GlcNH2 (α1-6) myo-inositol-1,2-cyclic-phosphate, were substantially protected from malarial symptoms and death after they were exposed to the actual parasite.

Figure 12.6.9 shows a cross-section of a membrane (with cholesterol, PE, SM) containing the glycosylated form of the human complement regulatory protein CD59 protein (1cdr) with a GPI anchor attached at its C-terminus. Note that the middle part of the anchor (glycan) holds the protein well above the top of the lipid bilayer. The soluble protein is also glycosylated. The protein binds to complement proteins C8 and/or C9, which are effector immune proteins that assemble on the surface of a cell undergoing lysis.

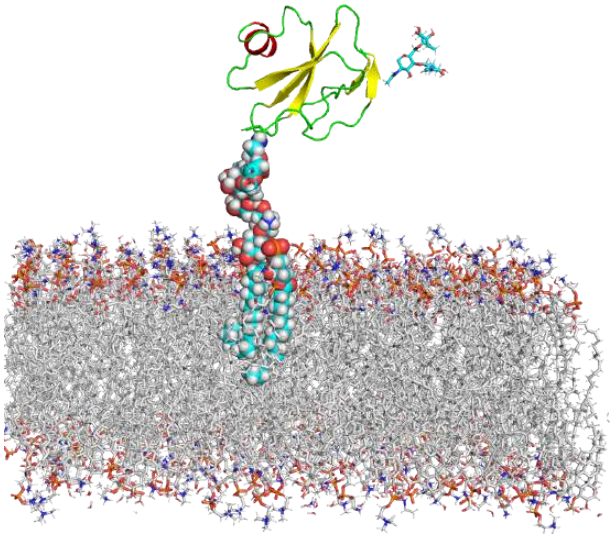


Figure 12.6.9: Cross section of a membrane containing the glycosylated form of the human complement regulatory protein CD59 protein (1cdr) with a GPI anchor attached at its C-terminus

The GPI anchor is shown in spacefill. Note that it only extends halfway into the bilayer, as expected from the size of the fatty acids attached to the phosphatidyl inositol. The glycan part of the GPI is shown in spacefill between the lipid and its protein attachment site. The protein is also glycosylated in the extracellular domain.

Something new!

An unexpected type of glycosylated molecule has been found in the outer leaflet of mammalian cells - a glycosylated RNA, as shown in Figure 12.6.10. This adds RNA to the lipids and proteins as a target for glycosylation. These surface glycoRNAs interact with antibodies against dsRNA and the Siglec lectin family. They are found in cells in vivo and cultured cells in vitro.

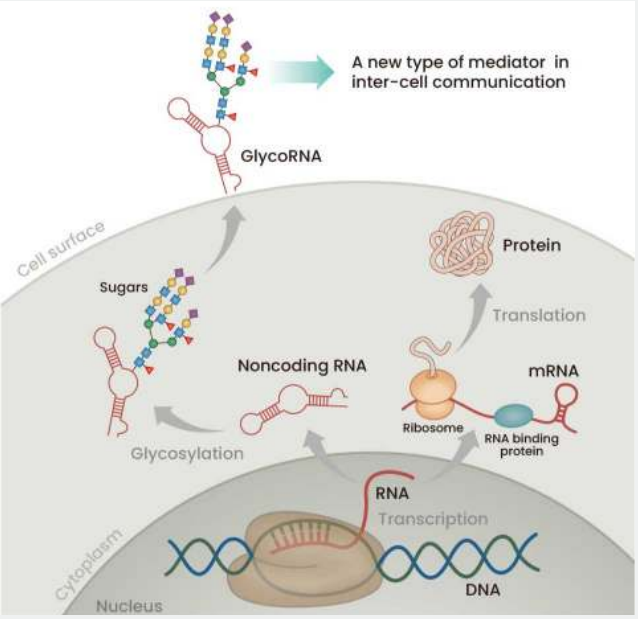


Figure 12.6.10: A glycoRNA - a small noncoding RNAs with sialylated glycans. Park. https://doi.org/10.14348/molcells.2021.0178 www.molcells.org. Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/.

Flynn et al., Small RNAs are modified with N-glycans and displayed on the surface of living cells, Cell (2021), https://doi.org/10.1016/j.cell.2021.04.023

12.6.5: Transmembrane (Integral) Proteins

These proteins pass through the membrane either in a single pass, usually with a single alpha-helix or many membrane-spanning helices. For example, G protein-coupled receptors, often called serpentine receptors, cross the membrane seven times. There are three types based on the number of times the protein crosses the membrane and the type of secondary structure used in crossing: biotopic (single pass), alpha-helical polytopic, and beta-barrel. These proteins are found in all types of membranes and have many types of functions, from receptors, receptor ligands, structural, adhesion, transport, gene regulation, and transport.

12.6.5.1: Transmembrane Biotopic - Single Pass Proteins

There are four types of single-pass transmembrane proteins:

- Type I: N-terminal outside of the cell (extracellular), and the precursor signal sequence on the N-terminus, which is a localization sequence, is removed
- Type II: N-terminal intracellular, with the transmembrane domain close to the N-terminus
- Type III: N-terminus extracellular and no signal sequence in precursor protein
- Type IV: N-terminus intracellular and the transmembrane domain close to the C-terminus

The transmembrane domain of single-pass integral membrane proteins consists of a single alpha-helix with nonpolar side chains extending outward from the helical axis, interacting with the nonpolar lipid parts of the membrane. These nonpolar sides are more stable in nonpolar environments.

To study such proteins in a less complex environment, membranes are often "dissolved" in nonpolar, single-chain amphiphilic detergents. These single-chain amphiphiles form micelles without membrane proteins but form **mixed micelles** in which the nonpolar part of the protein is surrounded in the detergent micelle by the nonpolar acyl chains of the detergent.

Figure 12.6.11 (top) below shows just the transmembrane and juxtamembrane (next to the membrane) domains of the single-pass Notch protein, which is critical in many signal transduction pathways.

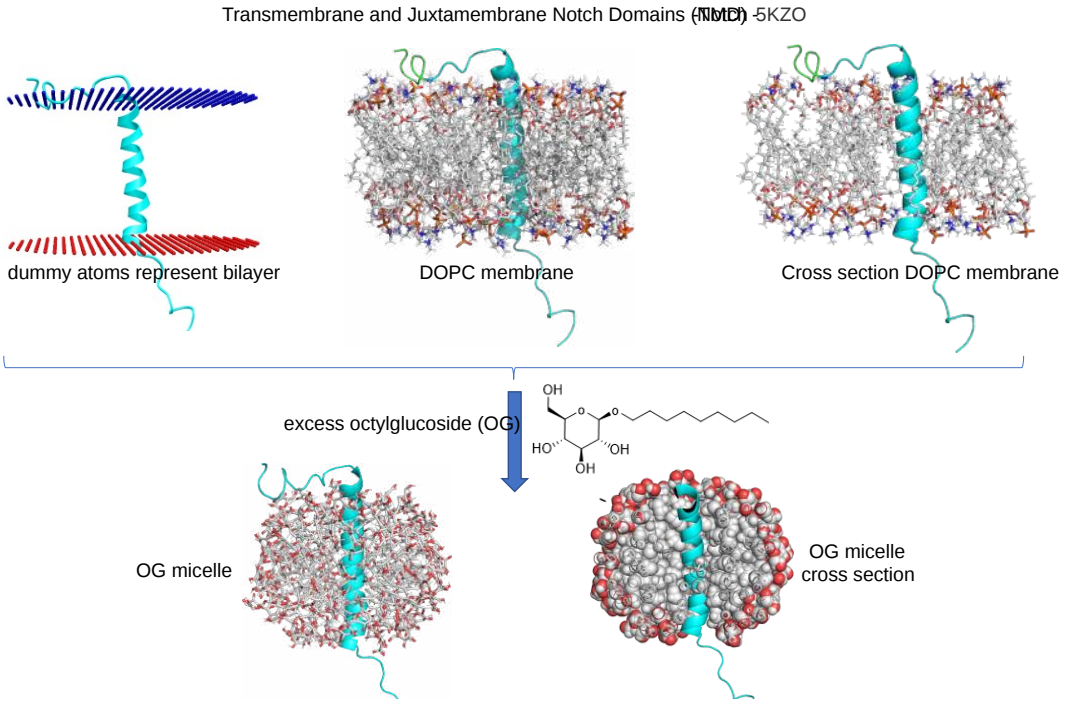


Figure 12.6.11: Transmembrane and juxtamembrane domains of the single pass Notch protein in a bilayer and an octylglucoside micelle

The top images in the figure above show different ways to represent the protein in the bilayer, with the right-hand image showing a cross-section through the membrane to better show how the protein passes through the bilayer.

The bottom images in the figure above show the protein after excess detergent, in this case, octylglucoside, is added to the protein-containing bilayer.

Studying membrane proteins in lipid aggregates - NMR applications

Recent Updates: 9/26/24

Various discrete lipid aggregates are used to conduct NMR studies of membrane proteins in a more realistic biological environment. Examples are shown in the Figure 12.6.12*A* below.

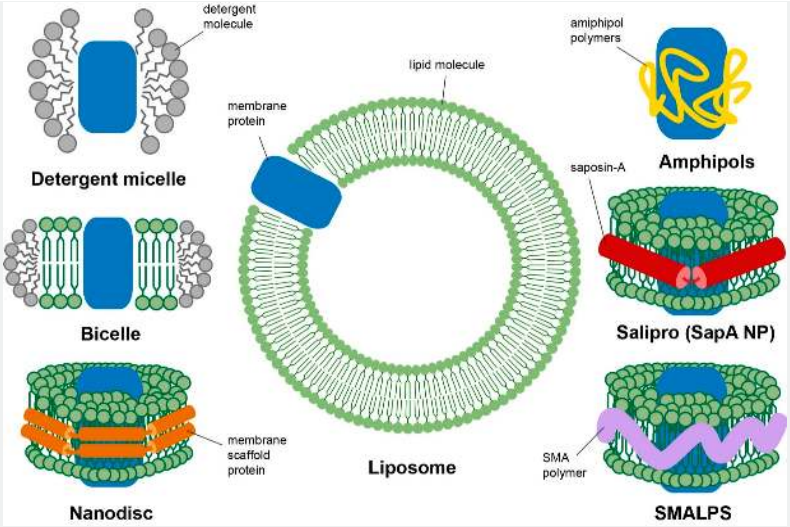


Figure 12.6.12*A*: Schematic diagram of different membrane mimetic systems used in NMR studies: detergent micelle, bicelle, nanodisc, liposome, amphipols, salipro, and styrene–malic acid co-polymer lipid particles (SMALPS). The membrane protein is represented as a blue cartoon block. Yeh, V.; Goode, A.; Bonev, B.B. Membrane Protein Structure Determination and Characterization by Solution and Solid-State NMR. *Biology* **2020**, *9*, 396. https://doi.org/10.3390/biology9110396. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**). For an alternative image see https://www.cell.com/chem/fulltext/S...2822%2900092-4.

An alternative description and use for general membrane protein purification and studies include a general polymer form and virus-like particles, as shown in Figure 12.6.12*B* below.

Detergent form	Liposome form	Nanodiscs form	Polymers form	Virus-like Particles form
• X-ray Protein crystallization • Solution-state NMR • Enzyme kinetics • Protein engineering and mutagenesis studies • Drug screening • Drug target discovery • Diagnostic application	• Electron microscopy • Solid-state NMR • Transport assays • Therapeutic application • Drug screening • Protein-protein interaction experiment • Small molecule or drug target discovery	• Immunization & antibody selection • Studies of protein-protein/protein-lipid/protein-ligand/protein-additive interactions • Membrane protein delivery • Studies of membrane protein conformation dynamics • Discovery of small-molecule drugs • Very suitable formats for various techniques including but not limited to: Binding assays (SPR, BLI, *etc.*) • Solution-state NMR • Cryo-EM • Single-molecule characterizations (SMFS, TIRFM *etc.*)	• Protein structure with Cryo-EM • Protein-lipid interaction	• Immunization for antibody production • Phage and yeast display for antibody discovery • Hybridoma screening for antibody discovery • Antibody characterization • Membrane protein binding assays • Membrane protein functional assays

Figure 12.6.12*B*: Alternative descriptions and uses in membrane protein solubilization, purification, and characterization. Creative BioLabs. MAGICTM MEMBRANE PROTEIN SERVICES & PRODUCTS. With permission (10/20/24)

Micelles: We've studied these in Chapter 10.2: Lipid Aggregates and the above descriptions. They form from the self-assembly of single-chain amphiphiles (detergents, such as octylglucoside and **[mono]**dodecylphosphocholine) as well as double-chain amphiphiles with short alkyl chains, such as hexanol-phosphatidylcholine (DHPC, also abbreviated C6DH).

Bicelle: Bicelles are made of phospholipids (such as dimyristoyl phosphatidylcholine - DMPC) that can form a bilayer and a detergent (a single-chain amphiphile that can form a micelle. The detergent fills in the hydrophobic "edges" of the bilayer part

to form the final structure. Bicelle size varies with a Q value, the [phospholipid]/[detergent] ratio. When Q=0 (high detergent), a micelle forms.

A phospholipid with short acyl chains such as dihexyl phosphatidylcholine (DHPC or C6DH) can replace the detergent. The structures of these two phospholipids (DMPC and DHPC) are shown in Figure 12.6.13 below.

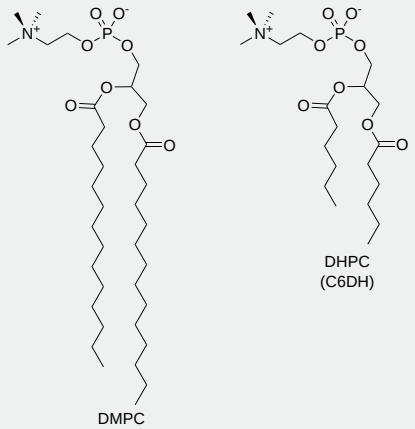


Figure 12.6.13: Structures of dimyristoyl phosphatidylcholine (DMPC) and dihexyl phosphatidylcholine (DHPC or C6DH)

Figure 12.6.14 below represents a cross-sectional view of a bicelle made from the above lipids. The central white/gray chains represent the alkyl chains of the DMPC phospholipid bilayer. The cyan spheres represent the carbon atoms of the smaller DHPC "detergent" that fill the edges of the central bilayer.

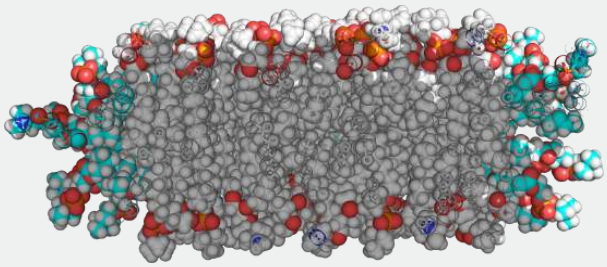


Figure 12.6.14: Cross-sectional view of a bicelle made with DMPC (white carbons) and DHPC (cyan carbons) which are seen around the periphery.

Figure 12.6.15 below shows an iCn3D model of the same bicelle. The black lines represent the alkyl chains of the central DMPC bilayers. The cyan sticks represent the alkyl chains of the DHPCs that fill the edges. Other atoms are shown in their own CPK colors.

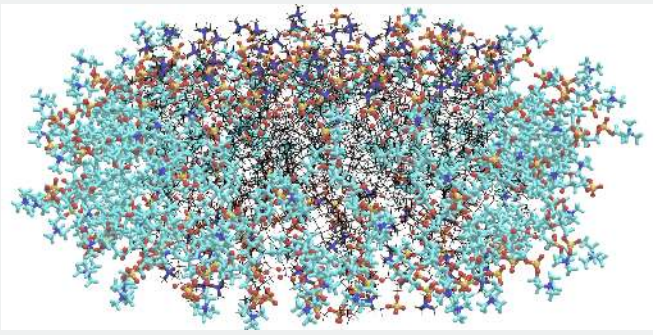


Figure 12.6.15: Bicelle

Open iCn3D and follow these prompts to get an interactive model of this bicelle:

- Download this file to your computer

- in iCn3D, **File, Open File**, and then open the iCn3D PNG Image

Nanodiscs: Nanodiscs are like bicelles that use two copies of an amphiphilic protein (membrane scaffold protein (MSP) from high-density lipoprotein (HDL) instead of a detergent or small double-chain amphiphile to protect the exposed hydrophobic edges of a small section of a bilayer. Figure 12.6.16 shows a cross-sectional view of a nanodisc composed of DMPC and two copies of the protein MSP.

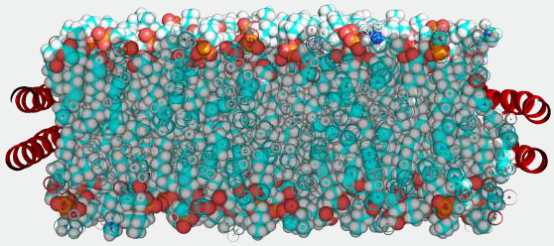


Figure 12.6.16: Cross-sectional view of a nanodisc made with DMPC (white carbons) and the membrane scaffold protein (MSP, red helices). Carbon atoms are shown in cyan. Image made with Pymol.

Figure 12.6.17 below shows an iCn3D model of the same nanodisc. The cyan atoms represent the carbons in the alkyl chains of the DMPCs. Other atoms are shown in their own CPK colors. The two helical amphiphilic proteins are shown in cartoon form in red.

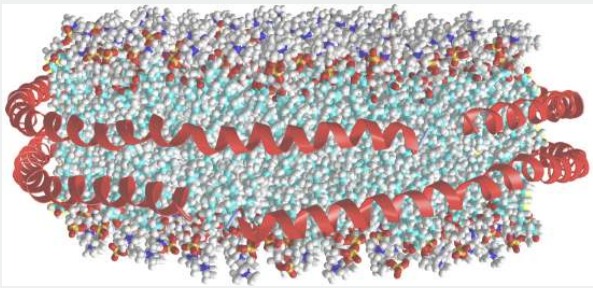


Figure 12.6.17: A nanodisc with DMPC and the protein MSP.

Open iCn3D and follow these prompts to get an interactive model of this nanodisc:

- Download this file to your computer
- in iCn3D, **File, Open File**, and then open the iCn3D PNG Image

Follow these instructions to view the hydrophobicity of the two alpha helices.

With the full model in the iCn3D window:

- **Analysis, Defined Sets**, choose Proteins
- **View, View Selection**,
- **Style, SideChains, Sticks**
- **Color, Hydrophobicity, Normalized**

Rotate the structure to see a top-down view, and you will observe that the hydrophobic side chains are all pointed inward in a fashion that would stabilize the exposed nonpolar alkyl chains of the phospholipids.

SMALPS (styrene–malic acid co-polymer lipid particles): These are nanodiscs wrapped with this amphiphilic polymer instead of an amphiphilic protein. The structure of SMALPS is shown in Figure 12.6.18 below.

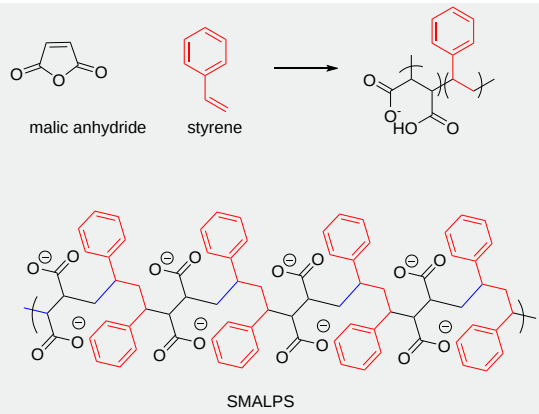


Figure 12.6.18

Figure 12.6.19 below shows an iCn3D model of SMALPS. Rotate the image to confirm its that it is amphiphilic.

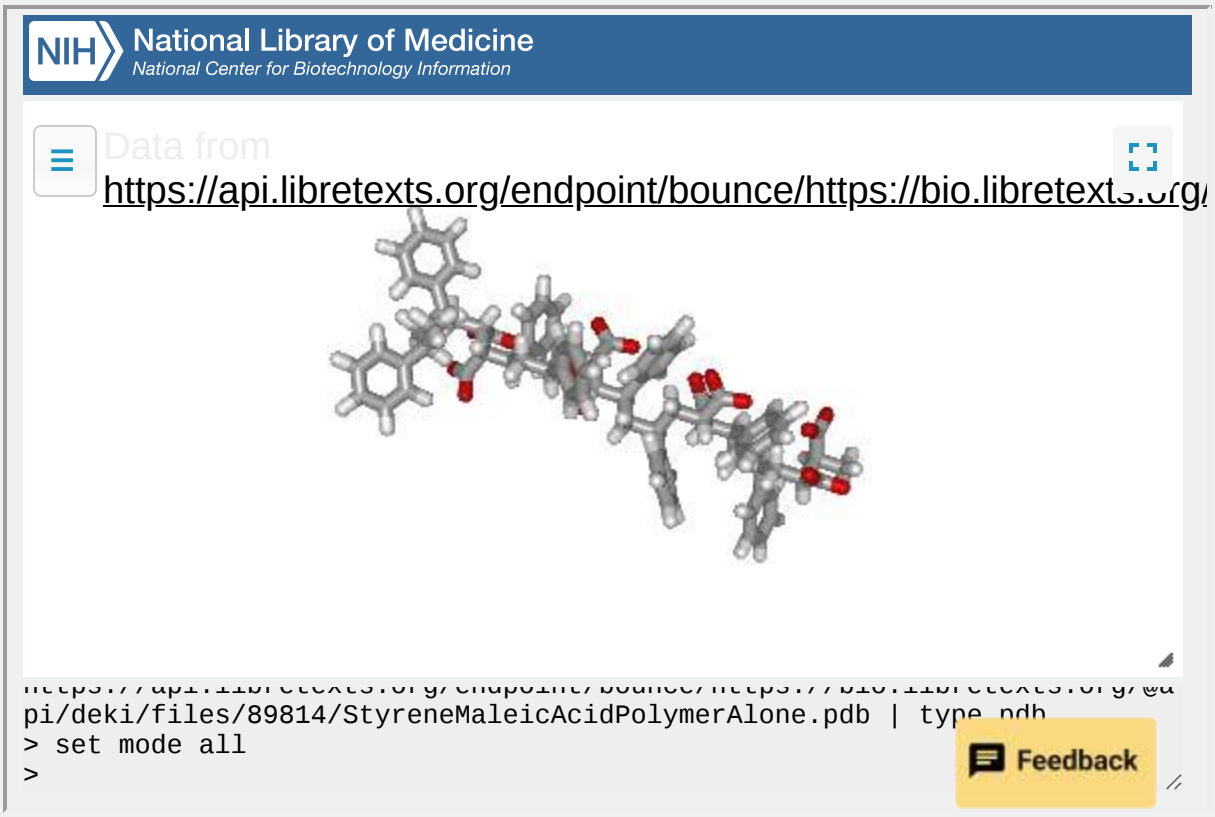


Figure 12.6.19: SMALPS (styrene–malic acid co-polymer lipid particles)

Here are some examples of bitopic single-pass transmembrane proteins

Cadherins

All structures need support and connections. At the macro level, the skeleton supports the mass and organization of organs and tissues in whole organisms. Within an organ, how can cells hold together? How do they adhere to each other? Certainly not through outer leaflet lipid contacts, as the outer surface of the leaflet is typically charged. The extracellular matrix does provide some of the glue that holds cells together. At a more detailed level, transmembrane proteins are involved. One class of adhesive proteins is cadherins, a calcium-dependent cell adhesion molecule. There are over 100 human cadherins. They are mostly ditopic, single transmembrane-pass proteins. Their cytoplasmic domains interact with proteins like catenin, which then bind to the interior

cytoskeletal network composed of actin and other proteins. This provides a way for the intracellular region to regulate the extracellular interactions of the cell.

The extracellular domain is composed of five repeating "cadherin" domains, each around 110 amino acids, that can fold independently. Calcium ions bind at the domain interfaces. A cadherin can interact with other cadherin domains on other cadherins on other cells, leading to cell adhesion. Essentially, the receptor cadherin on one cell binds to the ligand cadherin on the other. As metastatic tumor cells lose their adhering feature and leave the primary tumor site, you would expect mutations in cadherins to be often involved. They may also be involved in cell sorting during morphogenesis, regulation of tight and gap junctions, and in the control of intercellular spacing.

Figure 12.6.20 shows a "constructed" image of cadherin-1 (1L3W) interacting with cytoplasmic β-catenin (1I7X) through a modeled transmembrane helix (amino acids QIPAILGILGGILALLILILLLLLFLRR, amino acids 706-731). No full-length structure of cadherin in a membrane is available.

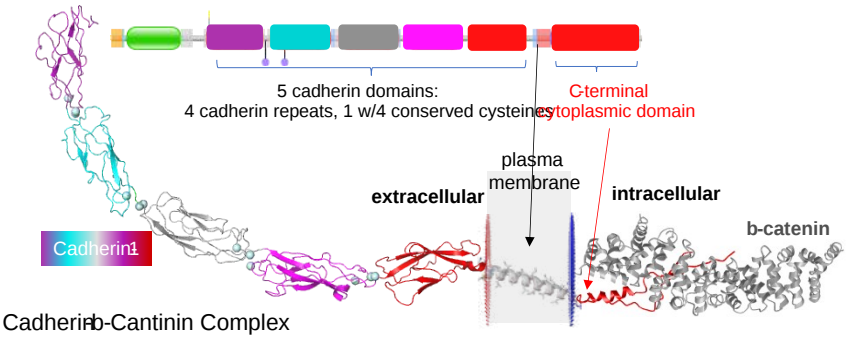


Figure 12.6.20: A constructed" image of cadherin-1 (1L3W) interacting with cytoplasmic β-catenin (1I7X) through a modeled transmembrane helix

Membrane Protein Kinases

Kinases are enzymes that phosphorylate substrates. Hexokinase is a protein enzyme that catalyzes the phosphorylation of a hexose substrate, such as glucose. A **protein kinase** is a protein enzyme that phosphorylates a protein substrate. That protein could be another copy of itself or another protein. We will see in Chapter 12 that many protein kinases are involved in cell signaling. Many tyrosine protein kinases are bitopic single-pass integral membrane proteins that become active when binding a ligand. Typically, on binding an extracellular ligand, two monomeric copies of the kinase form a dimer in the membrane, activating a tyrosine kinase cytoplasmic domain, which typically phosphorylates (using ATP as a substrate), the other member of the dimer in an "autophosphorylation" reaction. Sometimes, the dimers are held together by disulfide bonds.

Figure 12.6.21 shows a "constructed" image of the human dimeric insulin receptor.

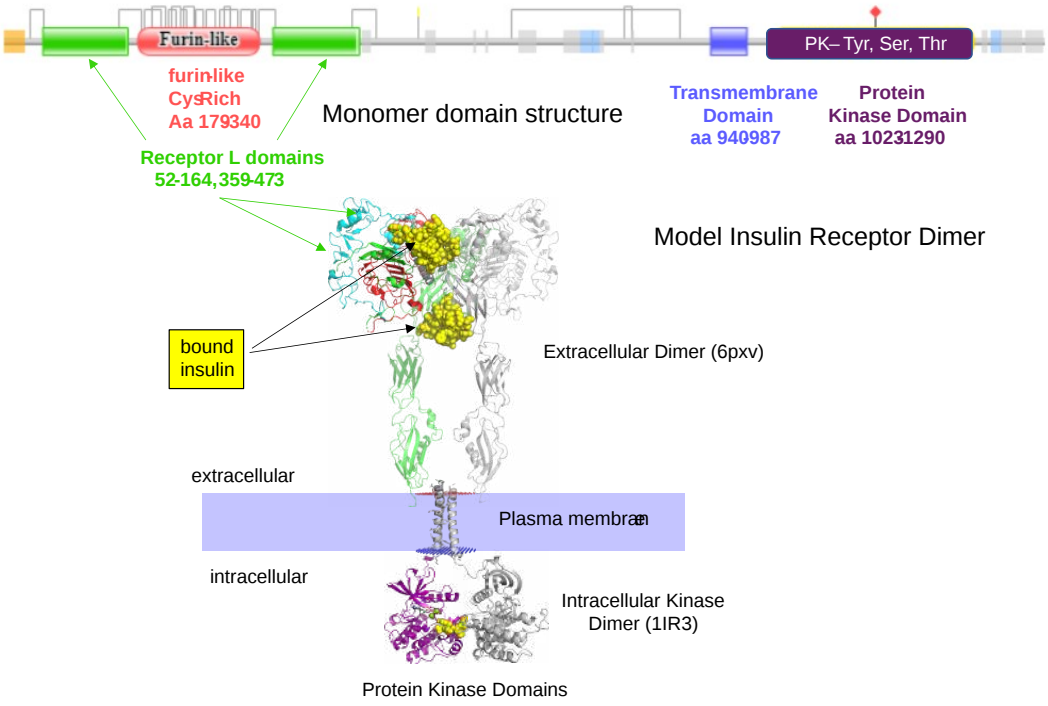


Figure 12.6.21: aA "constructed" image of the human dimeric insulin receptor

One of the monomers is shown in gray. The other monomer is shown in colors corresponding to the domain organization of the protein. Each extracellular dimer (6PXV) has two insulins bound (yellow spacefill). The intracellular domains (1IR3) are activated on insulin binding. No full-length structure of the full insulin receptor in a membrane is available.

Almost half of all helical membrane proteins in humans are bitopic, compared to 20-25% in prokaryotes. Humans have 10-20 fold more bitopic proteins than E.Coli. There appear to be about 196 bitopic proteins in E. coli (located in the inner membrane) and 70 in M. jannaschii (Archaea in the plasma membrane). In humans, 57% are in the plasma membrane, with the rest distributed between the Golgi, ER, nuclear, mitochondrial, and chloroplast membranes. In single-celled yeast, only 8% are in the plasma membrane.

Beta-Dystroglycan

This protein is another bitopic protein with a single alpha-helix membrane domain. Dystroglycan is a dimer of alpha and beta subunits. Alpha-dystroglycan is a peripheral protein that binds beta-dystroglycan, a transmembrane protein. Alpha dystroglycan also binds Lassa virus and lymphocytic choriomeningitis virus glycoprotein, serving as viral receptors. It also binds the protein dystrophin, which is missing in Duchenne muscular dystrophy and affects 1 out of 5000 live male births. As an integral transmembrane protein, beta-dystroglycan connects the extracellular matrix to the cytoskeleton through dystrophin. Alpha- and beta-dystroglycan share the same gene, which codes one long protein proteolyzed post-translationally to form the alpha (N-terminal end) and beta subunits (C-terminal end).

Figure 12.6.22 shows a schematic outline of dystrophin and the dystrophin-associated glycoprotein complex (DAGC).

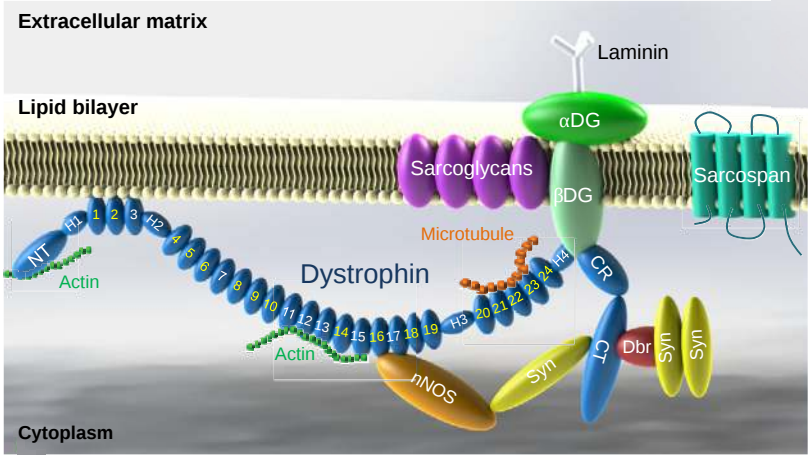


Figure 12.6.22: Schematic outline of dystrophin and the dystrophin-associated glycoprotein complex (DAGC). Dystrophin contains N-terminal (NT), middle rod, cysteine-rich (CR), and C-terminal (CT) domains. The middle rod domain is composed of 24 spectrin-like repeats (numbers in the cartoon, positively charged repeats are marked in white color) and four hinges (H1, H2, H3, and H4). Dystrophin has two actin-binding domains at NT and repeats 11-15, respectively. Repeats 1-3 interact with the negatively charged lipid bilayer. Repeats 16 and 17 form the neuronal nitric oxide synthase (nNOS)-binding domain. Dystrophin interacts with microtubules through repeats 20-23. Part of H4 and the CR domain binds to the β-subunit of dystroglycan (βDG). The CT domain of dystrophin interacts with syntrophin (Syn) and dystrobrevin (Dbr). Dystrophin links cytoskeleton components (actin and microtubules) to laminin in the extracellular matrix. Sarcoglycans and sarcospan do not interact with dystrophin directly, but they strengthen the entire DAGC, which consists of dystrophin, DG, sarcoglycans, sarcospan, Syn, Dbr, and nNOS.Disease Models & Mechanisms (2015) doi:10.1242/dmm.018424vailable via license: CC BY 3.0

12.6.6: Transmembrane - Alpha-helical polytopic

There are so many intriguing examples of these proteins. We'll illustrate just two.

Rhodopsin-like receptors and pumps

These proteins are involved in cell signaling and are the target of most pharmaceutical drugs. **G protein-coupled receptors (GPCRs)** are incredibly important, and we will discuss them extensively in the capstone chapter (28.2: At the cell membrane-receptors and receptor enzymes) for Vol 1.

GPCRs are cell receptors that span the membrane seven times in a serpentine fashion. They bind ligands (neurotransmitters, hormones, etc.) in the extracellular or internal membrane domains (the latter for hydrophobic ligands) and, through propagated conformational changes, alter the cytoplasmic domain where they functionally interact with a heterotrimeric G protein.

Figure 12.6.23 shows an interactive iCn3D model of the human cannabinoid receptor with bound cholesterol and Δ^9-tetrahydrocannabinol (Δ^9-THC) in spacefill (5xra). The red dummy atoms represent the outer leaflet, and the blue the inner.

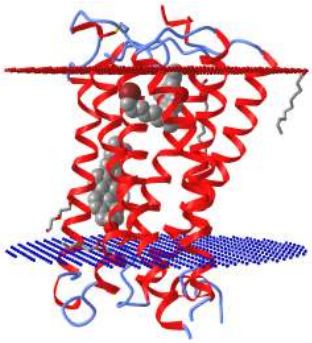


Figure 12.6.23: Human cannabinoid receptor with bound cholesterol and Δ^9-tetrahydrocannabinol (Δ^9-THC). (5xra). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PaTFrkGFgKeP2A

Δ^9-THC is a partial agonist and tunes the response of the receptor. The active site is conformationally somewhat flexible or plastic. Other ligands bound to it act as antagonists instead of agonists and must do so by eliciting inactive conformations.

ABC Transporter

Figure 12.6.24 shows an interactive iCn3D model of the P-glycoprotein multidrug resistance transporter protein (6nf1).

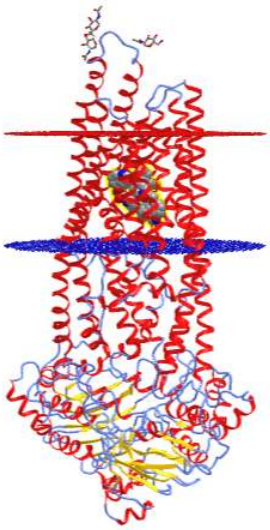


Figure 12.6.24: P-glycoprotein multidrug resistance transporter protein (6nf1). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...zzXduCZjhCMaw9

The spacefill ligands represent Zosuquidar, which binds with high affinity to P-glycoprotein and inhibits its activity, making it a cancer agent as it prevents chemotherapeutic drugs that have entered the cell from being pumped out. The protein chain interacting with it on the cytoplasmic face is an antibody fragment that stabilizes the P-glycoprotein so crystals can form.

12.6.7: Transmembrane Beta-barrel Fold

We will focus on two of these proteins.

Outer Membrane Factor (OMF) - Gram-negative bacteria

Figure 12.6.25 shows an interactive iCn3D model of a beta-barrel transmembrane protein OPRM - Outer Membrane Factor (4y1k) from Pseudomonas aeruginosa that acts as a pore. The transmembrane part of the protein consists of a beta-barrel fold (BBF). The protein also has a palmitoyl fatty acid in thioester linkage to Cys 1 of the protein for extra but unneeded anchorage.

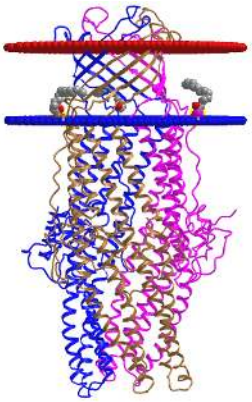


Figure 12.6.25: Beta-barrel transmembrane protein OPRM - Outer Membrane Factor (4y1k) from Pseudomonas aeruginosa (4y1k). (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...xrAuW1oFtY5FL7

Use this link for another view: https://structure.ncbi.nlm.nih.gov/i...FtY5FL7&t=4Y1K (OPM) in iCn3D

This protein is part of a large complex of proteins that spans both the inner and outer membranes of Gram-negative (examples, E. Coli and Pseudomonas aeruginosa) bacterial cell walls. Unfortunately for humans, this protein complex pumps out toxins (to the bacteria) like antibiotics, which makes bacteria resistant to these drugs. The OPRM acts as the outer passageway or duct for the pumped molecules. The Bacterial Outer Membrane Factor (OMF) protein differs in sequence, but they all form the beta-barrel duct. The E. Coli version of OMF has a triacylated lipid modification of the N-terminus. The N-terminal lipid modification might be necessary for the initial attachment of the protein to a membrane before the insertion of the beta-barrel. As such, the enzymes involved in tail attachment could be targets for new antibiotics.

Voltage-dependent anion channel (VDAC) - mouse

This protein, which also contains a beta-barrel fold, regulates the movement of molecules between the cytoplasm and the interior of the mitochondria across the outer mitochondrial membrane. VDAC also serves as a docking site or scaffold for assembling molecules into a complex that regulates mitochondrial function. The protein's conformation and function are regulated by changes in the transmembrane potential, which we will explore in the following sections. Hence, the protein and its function are voltage-dependent. Figure 12.6.26 shows an interactive iCn3D model of mouse VDAC with a beta-barrel fold formed by 19 beta-strands (3emn).

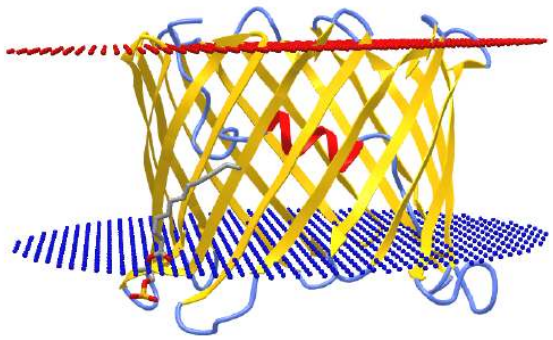


Figure 12.6.26: Mouse voltage-dependent anion channel (VDAC) (3emn). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...WxUqWaGPxgHH79

Note the N-terminal alpha-helix resident inside the channel opening. This helical section moves on changes in membrane potential, gating open and regulating the flow of metabolites and ions across the membrane through the pore. The conductance is high at a low transmembrane potential (10 mV) as the channel is in the open state. When the potential increases to 30 mV (either + or -), conductance drops as the protein forms the closed state.

An alternative beta-barrel fold is found in the protein enzyme **T**riose phosphate **I**so**M**erase (**TIM**), which we will see in Chapter x. That fold has eight α-helices connecting 8 parallel β-strands that form the barrel. The **TIM beta-barrel fold (TBF)** is found in up to 10% of protein enzymes. The transmembrane pore proteins above superficially resemble the TIM barrels, but they have loops, not helices, connecting the β-strands, which are antiparallel, not parallel as in TIM. They represent a different fold called the **beta-barrel fold (BBF)**.

Membrane proteins containing the **beta-barrel fold (BBF)** can now be designed de novo (see Chapter 4.14:Predicting Structure from Sequence and Sequence from Structure/Function) to create nanopores of predefined sizes. These could serve as designed sensors. The links below will take you to iCn3D models containing designed proteins with TIM barrel folds (TBF).

- 8UZL: Designed Transmembrane beta-barrel- TMB10_163 (10-stranded)
- 9FDG: Solution structure of a de novo designed 12-stranded transmembrane beta-barrel in LDAO micelles

Now that you understand the structure of membrane proteins, let's explore a key function of a subset of integral membrane proteins in the next chapter section: the movement of molecules/ions across the membrane.

12.6.8: Summary

This chapter provides a comprehensive overview of membrane proteins and their vital roles in cellular membranes. While the primary function of a lipid bilayer is to act as a semi-permeable barrier that distinguishes the inside from the outside of a cell or

organelle, the incorporation of proteins adds layers of functionality, including signal transduction, adhesion, transport, and enzymatic activity.

Key topics include:

- **Types of Membrane Proteins:**
  The chapter categorizes proteins based on their association with the membrane. Peripheral proteins, which interact with the membrane through transient, non-covalent bonds (e.g., via lipid head groups or through binding to other membrane proteins), are contrasted with integral (transmembrane) proteins that span the bilayer. These integral proteins are further classified as:
  - **Bitopic (Single-Pass) Proteins:**
    These include various types (I–IV) depending on the orientation of their domains relative to the membrane. Examples, such as prothrombin's GLA domain and the Notch protein, illustrate how even single-pass proteins can play crucial roles in signaling and cellular regulation.
  - **Polytopic (Multi-Pass) Proteins:**
    These proteins, including G protein-coupled receptors (GPCRs) and ion channels, typically traverse the membrane multiple times (e.g., seven transmembrane domains in GPCRs), mediating a wide array of cellular functions.
  - **Beta-Barrel Proteins:**
    Found primarily in the outer membranes of Gram-negative bacteria and in mitochondrial membranes (e.g., VDAC), these proteins form pore-like structures that facilitate the transport of molecules.
- **Protein Modifications and Membrane Targeting:**
  The chapter describes how covalent modifications such as fatty acid attachment (myristoylation, palmitoylation), isoprenoid addition (farnesylation, geranylgeranylation), and glycosylphosphatidylinositol (GPI) anchoring serve to target proteins to the membrane. These modifications not only anchor proteins to specific membrane regions but also influence their interactions and functional states.
- **Membrane Dynamics and Protein Interactions:**
  The dynamic nature of membranes is underscored by the lateral and transverse mobility of lipids and proteins, essential for processes such as vesicle trafficking, membrane fusion, and signal transduction. Detailed examples, including the interaction of matrix metalloproteinases (MMPs) and clotting factors with lipid bilayers, illustrate how protein binding can enhance catalytic activity or trigger signaling cascades.
- **Experimental Approaches for Studying Membrane Proteins:**
  Various membrane mimetics—such as detergent micelles, bicelles, nanodiscs, and SMALPs—are introduced as tools to study membrane proteins in environments that mimic native membranes. These systems facilitate structural and functional analyses using techniques like NMR, electron microscopy, and molecular dynamics simulations.
- **Functional Relevance in Health and Disease:**
  Finally, the chapter connects the structural and functional aspects of membrane proteins to their roles in disease. Examples include the involvement of cadherins in cell adhesion, the critical function of GPCRs in signal transduction (with implications in pharmacology), and the role of transporters like P-glycoprotein in drug resistance.

In summary, this chapter weaves together the principles of membrane biochemistry, highlighting how the unique architecture of lipid bilayers, combined with diverse protein modifications and dynamic interactions, underlies the complex functionalities required for cellular communication, adhesion, transport, and signal transduction. Understanding these interrelationships is crucial for unraveling cellular processes and developing targeted therapeutic interventions.

This page titled 12.6: Membranes and Membrane Proteins is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.1: Membranes and Membrane Proteins** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion

Learning Goals (ChatGPT o3-mini)

1. **Differentiate Diffusion Mechanisms:**
   • Distinguish between passive diffusion, facilitated diffusion, and active transport, and explain the conditions under which each process occurs.
2. **Apply Fick's First Law:**
   • Use Fick's First Law of Diffusion to derive and interpret the relationship between flux, diffusion coefficient, concentration gradient, and membrane thickness.
3. **Understand Permeability Coefficients:**
   • Define the permeability coefficient and explain how it quantitatively relates to the ease with which different molecules (based on size and charge) traverse a membrane.
4. **Analyze Electrochemical Driving Forces:**
   • Explain the concept of electrochemical potential and how the combination of concentration gradients and transmembrane potentials influences the movement of charged species (e.g., ions) across membranes.
5. **Derive and Interpret the Goldman Equation:**
   • Compare and contrast the Goldman Equation with the Nernst Equation, and understand how electric potential and concentration gradients together determine ion flux across membranes.
6. **Explore Facilitated Diffusion Kinetics:**
   • Derive the relationship between receptor-bound substrate concentration and net flux in facilitated diffusion, and recognize how this leads to a hyperbolic (Michaelis–Menten-like) dependence of flux on substrate concentration.
7. **Examine Membrane Transport Proteins:**
   • Identify and describe the roles of different types of membrane proteins (carrier proteins/permeases, channels, ionophores) in facilitating the movement of molecules and ions, with specific examples such as GLUT1 and the ADP-ATP carrier.
8. **Evaluate Fatty Acid Transport Models:**
   • Compare the protein-mediated model (involving FATPs, CD36, and fatty acid binding proteins) versus passive diffusion models for long-chain fatty acid transport, considering the factors that limit free diffusion of hydrophobic molecules.
9. **Interpret Computational and Experimental Data:**
   • Analyze simulation outputs and experimental animations (e.g., PHET and Vcell simulations) to understand the dynamics of solute diffusion, including how equilibrium is reached and the role of kinetic barriers.
10. **Integrate Thermodynamics with Kinetics in Diffusion:**
    • Discuss how changes in free energy (ΔG), enthalpy (ΔH), and entropy (ΔS) influence diffusion rates and how these thermodynamic factors interplay with kinetic parameters in membrane transport.

These learning goals are designed to help you bridge the theoretical foundations of diffusion and membrane transport with practical examples and applications in biological systems.

12.7.1: Diffusion Across a Membrane

We have studied molecular aggregates (micelles and bilayers) and macromolecular structures (mostly proteins). We also studied binding interactions, which are the first steps in the expression of the biological activity of a macromolecule. For some proteins, reversible binding is their sole function (consider the binding of dioxygen to myoglobin and hemoglobin). For many others, it is not. For those, what can happen next?

You already have one possible answer. A bound reactant, which we will call a substrate, can be converted to a product in a chemical step involving the breaking and making of covalent bonds catalyzed by a protein enzyme. However, there is an even simpler

process that does not involve covalent bond changes. If a small molecule is bound to a membrane protein, it could be transported in a purely physical step across a membrane. Just as reactions can proceed with and without an enzyme, a solute can move down a concentration gradient across a semipermeable membrane, driven by diffusion alone in a thermodynamically favorable process, either by itself, in a process called **passive diffusion,** or with the assistance of a membrane protein, in a process called **facilitated diffusion**. Large **pores** made of assemblies of proteins can also be formed that allow the passage of many solutes across the membrane.

There are many occasions when moving a molecule across a membrane from a region of low to high concentration would be optimal. This process is called **active transport**. It is not thermodynamically favored, so it requires an external energy source. This is often the thermodynamically favored cleavage of ATP to ADP and Pi. The uphill transport can also be powered by the downhill diffusion of a "co-transport" molecule from high to low concentration. We will explore these processes in this and the remaining chapter sections. First, we should understand the simplest process, "passive diffusion", that requires no protein "help".

Let's step back and consider how difficult it is for a chemical species to cross a lipid bilayer. Chemical intuition would tell us that both size and polarity are important. The bigger the size and the greater the charge, the more difficult it would be to cross the membrane. The permeability coefficient is related to the ease with which solutes traverse the membrane. Figure 12.7.1 shows the permeability coefficients for relevant biological molecules.

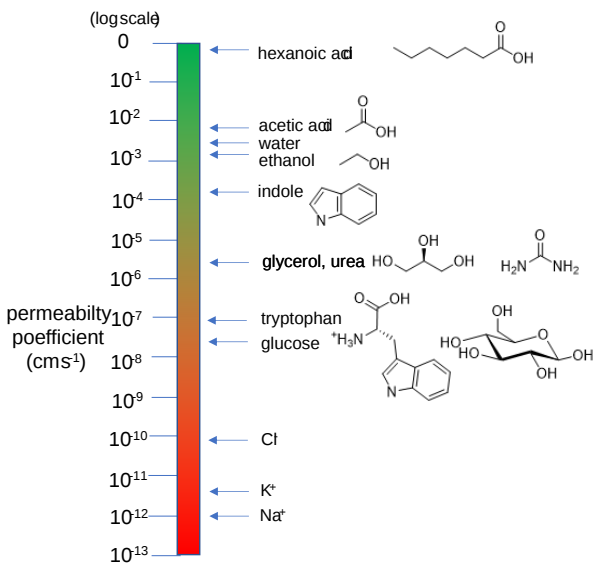


Figure 12.7.1: Permeability coefficients for relevant biological molecules

Smaller, higher charge density ions (like Na^+) have a lower permeability coefficient than do larger, lower charge density ions (like K^+) as seen in (Table 12.7.1). What about natural membranes?

Table 12.7.1: Permeability coefficient (cm/s) of natural and synthetic membranes to D-Glucose and D-mannitol at 25 °C

Membrane Preparation	D-Glucose	D-Mannitol
Synthetic Lipid Bilayer	2.4×10^{-10}	4.4×10^{-11}
Calculated Passive Diffusion	4×10^{-9}	3×10^{-9}
Intact Human Erythrocyte (red blood cell)	2.0×10^{-4}	5×10^{-9}

It looks like D-glucose gets a little help getting across. We'll see the mechanism below.

12.7.2: Passive Diffusion

Let's start with the passive diffusion of uncharged solute A across a membrane, represented by the chemical equation $A_{out} \leftrightarrow A_{in}$. Intuitively, you probably believe that the rate of **net** diffusion or the **flux** of A across the membrane is directly proportional to the

concentration gradient. If concentrations of A are identical across the membrane, the net flux J should be 0. If you double the concentration gradient, the net rate should double. We will see that the net rate is a linear function of [ΔA] across the membrane. Figure 12.7.2 shows the flux of A_{out} across a semipermeable membrane of thickness Δx (we will use dx instead when Δx is very small).

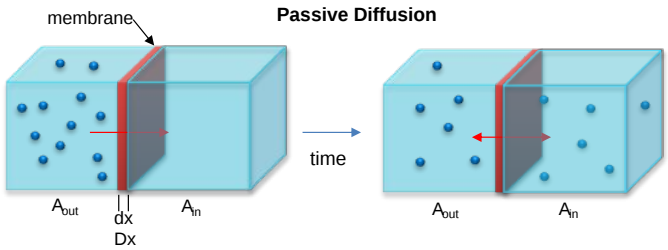


Figure 12.7.2: Flux of A_{out} across a semipermeable membrane of thickness Δx

Let's animate diffusion using a PHET simulation, as shown in Figure 12.7.3.

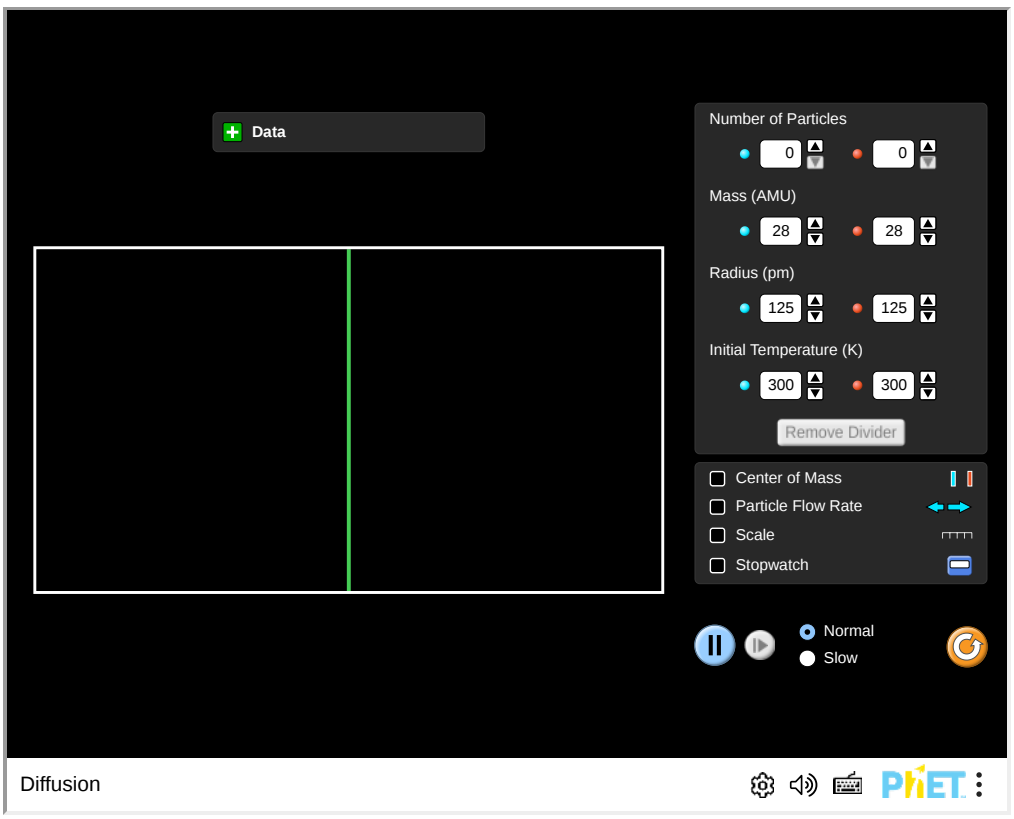


Figure 12.7.3: PHET animation of passive diffusion. PHET: https://phet.colorado.edu/

The flux of molecules A (\(J__A\)) is proportional to the concentration gradient across the membrane, $\Delta A/\Delta x$ (which we will refer to as dA/dx, which is the derivative of A with respect to x). The equation below is Fick's First Law of Diffusion:

$$\mathrm{J_A} \propto \frac{d\,\mathrm{A}}{dx} = -\mathrm{D}\frac{d\,\mathrm{A}}{dx} \tag{12.7.1}$$

where D is the diffusion coefficient. The negative sign is necessary since concentration increases to the left in the figure above in the opposite direction of net flux, which is to the right. For these derivations, we will assume the J_A is the initial flux. That is, the flux is measured for a short enough time that the relative concentrations of A on both sides of the membrane do not change significantly. It should be clear that eventually, the net flux levels off to zero when the concentrations of A on both sides of the

membrane are equal. Under these conditions, the free energy $G_{A\,out}$ = $G_{A\,in}$, so $\Delta G = 0$. This thermodynamic relation can also be expressed as

$$J_A = -L\frac{dG_A}{dx} \tag{12.7.2}$$

This equation bridges the kinetic and thermodynamic aspects of diffusion.

Dimensional analysis of Fick's 1st Law (Equation 12.7.1) shows that the units of D are cm^2/s, which gives the number (dimensionless) of molecules crossing a 1 cm^2 surface area of membrane each second.

J = moles/area/sec = $mol/cm^2 \cdot s$ = - (cm^2/s) $mol/cm^3/cm$. Hence, the units of D are cm^2/s.

Let's rearrange Fick's 1st Law and use a bit of calculus to get Equation 4.

$$\begin{aligned} \mathrm{J_A}\int_0^{\mathrm{x}} \mathrm{dx} &= -D\int_{\mathrm{A_{out}}}^{A_{\mathrm{in}}} \mathrm{dA} \\ \mathrm{J_A x} &= -\mathrm{D}\left(\mathrm{A_{in}} - \mathrm{A_{out}}\right) \\ \mathrm{J_A} &= \frac{\mathrm{D}}{\mathrm{x}}\left(\mathrm{A_{out}} - \mathrm{A_{in}}\right) \end{aligned} \tag{12.7.3}$$

or

$$J_A = P\left(A_{\mathrm{out}} - A_{\mathrm{in}}\right) = P\Delta A \tag{12.7.4}$$

where P is the permeability coefficient, which has units of $cm^2/s/cm$ or cm/s. (We discussed permeability coefficients for different solutes traversing model bilayers when we discussed lipids.) That unit is less intuitive to understand, but the final unit is very intuitive.

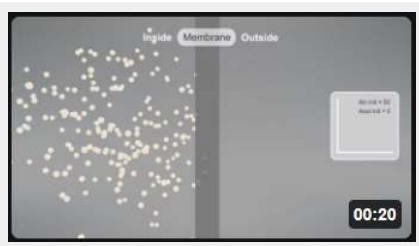
A plot of J_A vs (A_{out} - A_{in}) is linear, with a slope of P = D/x

It's important to remember that there is still diffusion of A across the membrane at equilibrium since the equilibrium is dynamic. There is no net diffusion, however. This is where animations come in handy.

The table below shows the reaction diagram (left), the progress curve (middle), and animations for the reversible diffusion $A_{out} \leftrightarrow A_{in}$ across a membrane with the conditions below. The reactant A and product are called species and are shown as green spheres. The yellow square is a reaction node indicating a reaction connects A to P. The lines connect the species that participate in the reaction. The velocities (slope of the concentration vs time curve at any given time) are called **fluxes, J**, in Vcell and many other similar programs. When we get to metabolism, we will talk about fluxes of metabolites through pathways. Also, fluxes are used to describe the solute's movement rate through membranes.

Animations

Now let's look at animations of the same passive diffusion reaction of a neutral species across a semipermeable membrane. Animations are by Shraddha Nakak and Hui Liu.

$A_{out} \leftrightarrow A_{in}$ $[A_{in}]_{t=0}$ = 100; $[A_{out}]_{t=0}$ = 0; P = 2	$A_{out} \leftrightarrow A_{in}$ $[A_{in}]_{t=0}$ = 50; $[A_{out}]_{t=0}$ = 0; P = 2
	

Note the dynamic nature of the diffusion. An equilibrium is reached when the number of particles inside equals those outside. Diffusion occurs in both directions from compartments of equal volume so that the particles moving to the outside don't escape into a comparatively huge volume. (Animations by Shraddha Nayak and Hui Lui)

12.7.3: Passive Diffusion of ions across a membrane - Transmembrane Potentials

Before moving on to facilitated diffusion, let's alter the scenario and use a charged solute. In the example of passive diffusion above, the only thermodynamic driving force for the movement of A across the membrane was the ΔG_A, the change in free energy/mol of A across the membrane (or more strictly $\Delta\mu$ = change in chemical potential). Solute A moves spontaneously across the membrane from high to low concentration. But what if A was charged?

We could add a bunch of positively or negatively charged species to Figure 12.7.2 and ask what would happen. You can't go to a chemical stockroom and find a bottle of K^+ ions, but you could find a bottle of neutral KCl. Let's make our experiment system a vesicle with 0.1 M KCl in the aqueous inner compartment and 0.1 M NaCl on the outside. We could easily prepare such vesicles by making large unilamellar vesicles (LUVs) with entrapped 0.1M KCl in a solution of 0.1 M KCl and then separating the vesicles from the 0.1 M KCl not encapsulated on liposome formation using a size exclusion gel chromatography column equilibrated and eluted in 0.1 M NaCl. These vesicles are illustrated on the left of Figure 12.7.4.

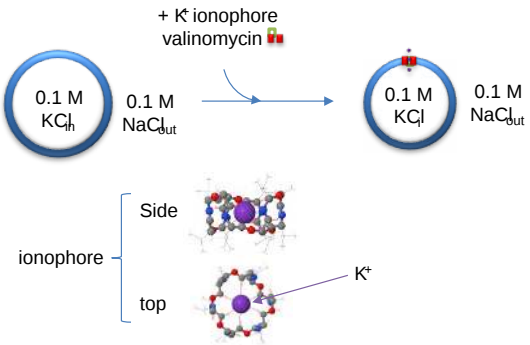


Figure 12.7.4: Making vesicles with a transmembrane potential

In these prepared vesicles, is there a net thermodynamic driving force for K^+ and Cl^- to move from inside to outside the vesicle? Not for Cl^- since its concentration is the same on both sides of the membrane (see Fick's 1st Law above). However, a clear thermodynamic driving force is needed to take $K^+_{in} \rightarrow K^+_{out}$. If the membrane were impermeable, net outward flux would not occur even though it is favored thermodynamically. Think of this as an example of a reaction under complete kinetic control! Note that this example also has a net driving force to move Na^+ ions from outside to inside.

In our next step, let's make the membrane permeable to only K^+ ions. We can do this by adding a small antibiotic, valinomycin, which binds in the membrane and, once there, binds and carries K+ ions across the membrane. It is called an **ionophore.** Figure 12.7.5 shows an interactive iCn3D model of K^+ bound to Valinomycin.

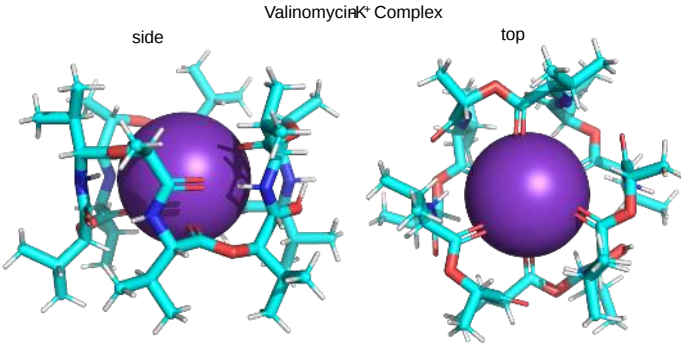


Figure 12.7.5: K^+ bound to Valinomycin (xxxx). (Copyright; author via source). Click the image for a popup or use this external link: Not available.

Valinomycin, from *Streptomyces fulvissimus*, is a cyclic peptide consisting of L and D-Val, L-lactate, and D-hydroxy isovalerate, connected through both ester and amide bonds. The K^+ ion is in the center. The six valine carbonyl oxygens bind the K^+ ion. The hydrophilic groups are pointed toward the center. In contrast, the hydrophobic groups point to the outside of the structure, allowing the K^+ ion to be sequestered in a polar environment as the nonpolar exterior of the complex passes through the membrane. This ionophore is specific for K^+ and binds the smaller Na+ ion weakly. Two factors can account for this. The smaller sodium ion doesn't bind as tightly to the chelating carbonyl oxygens. Also, the sodium ion has a higher charge density, so the Na^+/water

interactions must be more stable and difficult to break than those of K^+. The ion must be desolvated before it binds to the complex. Other ionophores are specific for other ions.

Once the ionophore is bound, the kinetic barriers to K^+ efflux are removed, and it starts moving K^+ from inside to outside. However, as soon as it does, the charge balance across the membrane is lost, with the outside becoming net positive and the inside becoming net negative. A transmembrane electric potential develops across the membrane. This disfavors stops K^+ efflux to the outside and eventually stops it, even as the concentration difference of K^+ across the membrane still favors efflux.

If you were a positive ion stuck in the middle of a membrane, as illustrated in Figure 12.7.6, which way would you move?

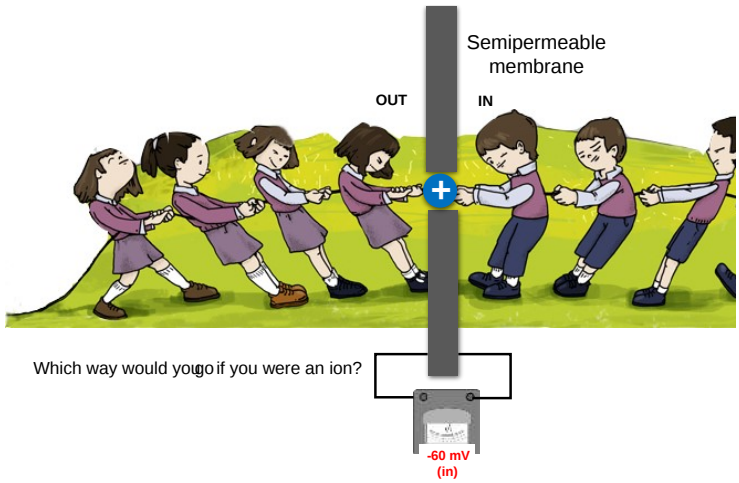


Figure 12.7.6: Which way does a positive ion move across a membrane with a transmembrane potential? (adapted from https://commons.m.wikimedia.org/wiki/File:Dib_1.jpg

There are now two thermodynamic driving forces for K^+ movement from inside to outside:

- a $\Delta G_{concentration}$, which favors K^+ efflux. At time t=0, $\Delta G_{concentration} << 0$, and it becomes a bit less negative (less favored) with efflux.
- a $\Delta G_{membrane\ pot}$, which is zero to start and slowly becomes positive, increasingly disfavoring K^+ efflux.

When these driving forces are equal and opposite, **net** K^+ movement across the membrane stops, and the system is in dynamic equilibrium.

Since we use electric potential to describe electrical phenomena (electron, ion movement), we often use the word **chemical potential,** in this case, to describe the movement of ions across a concentration gradient. Add them together, and we call it the **electrochemical potential**.

$$\begin{aligned} \Delta G_{electrochemical} &= \Delta G_{chemical} + \Delta G_{electrical} \text{ or} \\ \Delta \mu_{electrochemical} &= \Delta \mu_{chemical} + \Delta \mu_{electrical} \end{aligned} \tag{12.7.5}$$

We can use this understanding to derive an equation for flux J of a charged solute across a membrane of a given potential. The equation is called the Goldman Equation and is shown below.

$$J = \frac{P\frac{ZF}{RT}(E_1 - E_2)\,C_1\left(1 - \frac{C_2}{C_1}e^{\frac{ZF}{RT}(E_2-E_1)}\right)}{1 - e^{\frac{ZF}{RT}(E_2-E_1)}} \tag{12.7.6}$$

where

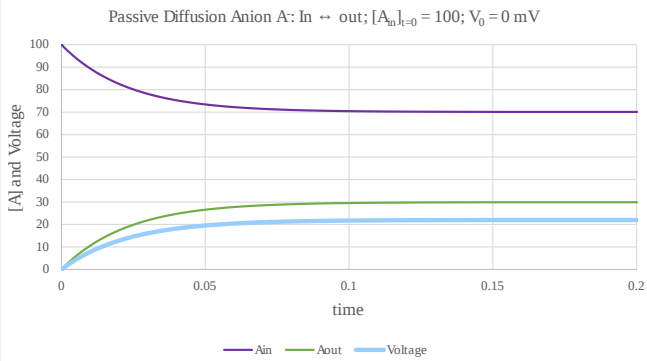
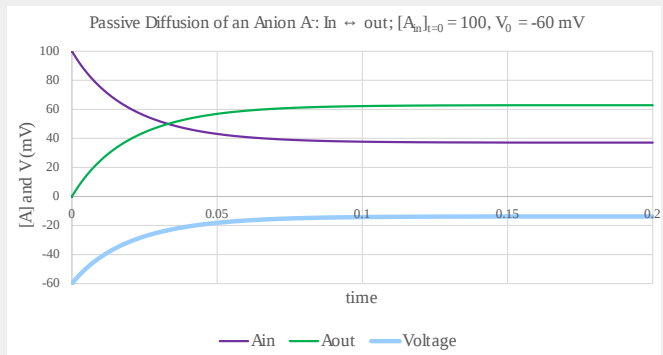
- P is the permeability coefficient
- Z is the charge or valence on the ion
- F is the Faraday constant
- R is the ideal gas constant
- T is temperature
- E2-E1 and the reverse is the transmembrane potential
- C2-C1 are the concentrations of the ions across the membrane

Compare this to the Nernst equation you learned in introductory chemistry courses.

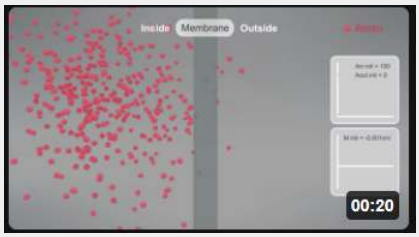
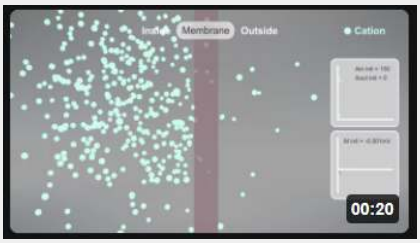
$$E = E^o - \frac{RT}{nF} \ln Q \tag{12.7.7}$$

that relates the reduction potential of an electrochemical reaction to the standard electrode potential, temperature, and concentration, where E is the potential difference.

Now, let's run a Vcell simulation for the diffusion of an anion across a semipermeable membrane. To do so, we must first set the initial transmembrane potential to solve the Goldman equation numerically. At present, this type of simulation **can not** be embedded into this book. So instead, the concentration vs. time graphs for two different simulations, one at an **initial** transmembrane potential (E_2-E_1) = - 0.001 (i.e., 0) and one at -60 mV (a typical cell resting potential), are presented in the Figure below. The reaction is A-in ↔ A-out [A-in]t=0 = 100 for each.

$A^-_{in} \leftrightarrow A^-_{out}$ $[A^-_{in}]_{t=0} = 100$; $[A^-_{in}]_{t=0} = 0$; P = 100; $V_{initial} = 0$ (-0.001 V)	$A^-_{in} \leftrightarrow A^-_{out}$ $[A^-_{in}]_{t=0} = 100$; $[A^-_{in}]_{t=0} = 0$; P = 100; $V_{initial} = -60$ mV
	

Here are animations that show the selective reversible movement of anions (A^-, left panel, red) and cations (C^+, right panel, cyan) as they move across a membrane from the inside to the outside. This is a very simplified simulation as it shows no counterions on either side. Assume they exist, even though having a "container" with just anions or cations would be impossible. The initial transmembrane potential (t=0) is 0 (actually -0.001 to allow the calculations using Vcell).

$A^-_{in} \leftrightarrow A^-_{out}$ $[A^-_{in}]_{t=0} = 100$; $[A^-_{out}]_{t=0} = 0$; P = 100; $V_{initial} = 0$ (-0.001 V)	$C^+_{in} \leftrightarrow C^+_{out}$ $[C^+_{in}]_{t=0} = 100$; $[C^+_{out}]_{t=0} = 0$; P = 100; $V_{initial} = 0$ (-0.001 V)
	

As anions move to the outside (left animation), the inside becomes less negative with respect to the outside, so the membrane potential V becomes more positive. This is indicated by the membrane changing to a blue color. Conversely, as cations move to the outside (right animation), the inside would become more negative with respect to the outside, so the membrane potential V becomes more negative. This is indicated by the membrane changing to a red color. (Animations by Shraddha Nayak and Hui Lui)

12.7.4: Facilitated Diffusion

Now, let's return to the diffusion of a noncharged solute down a concentration gradient (i.e., favored) after binding to a membrane receptor. The answer to that question depends on the biological function of the macromolecule. We can simplify this process by adding one additional step, as reflected in the equilibrium binding expression shown below:

$$M + L \rightleftharpoons ML \rightleftharpoons M + X$$

This expression indicates that the free ligand has changed in some fashion to x. In the next two chapters, we will consider two kinds of transformations:

- L is a ligand on the outside of a biological membrane (Lout) that binds to a membrane protein receptor, R. This undergoes a conformational change (as we studied in the binding of dioxygen to hemoglobin), which leads to the expulsion of the bound ligand to the inside of the membrane (Lin). This can be modeled with the simple equation:

  $$R + L_{out} \rightleftharpoons RL \rightleftharpoons R + L_{in}.$$

  This process is called facilitated diffusion and represents a physical as opposed to chemical process since no covalent bonds are made or broken. This process proceeds down a concentration or chemical potential gradient ($\Delta\mu < 0$) and hence is spontaneous (thermodynamically favored). If the ligand concentration is higher inside the cell, net diffusion moves it to the outside. Passive (non-facilitated) diffusion is kinetically slow without a receptor since membranes present formidable barriers to the passage of polar molecules.
- L is a ligand (or substrate S) that binds to a protein enzyme, E. The bound substrate is chemically altered to produce a new product, P, which dissociates from the enzyme. This can be expressed most simply as:
  E + S <==> ES <==> E + P .

Consider the mechanism illustrated in Figure 12.7.7.

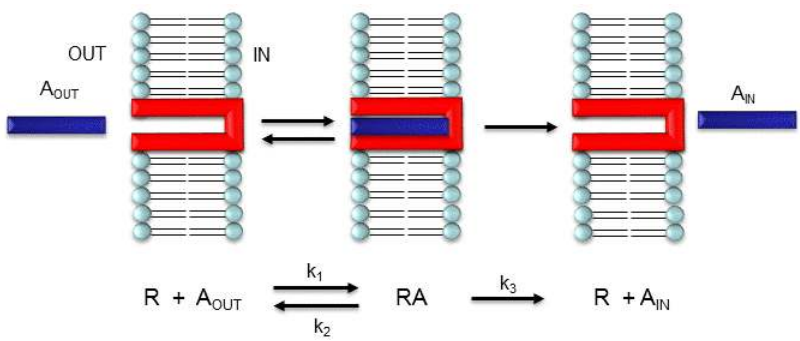


Figure 12.7.7: Mechanism for faciliatated diffusion

Let's assume that the initial flux will be measured for this system. We want to derive equations showing J as a function of A_{out} (assuming that A_{in} is negligible over the time of measuring the initial flux). Also, assume that the $J_{facilitated} >> J_{passive}$. In contrast to passive diffusion, JA is not proportional to A_{out} but rather to $[A_{bound}]$.

Consider this example to help you understand the proportionality. Pretend that the receptor is a truck that can carry one particle across the membrane at a time (i.e., 1/1 stoichiometry). Also, assume that the particle can't get across without being carried by the truck. If no trucks are in the membrane, no load can be delivered. If trucks are in the membrane without particles, no load will be delivered. As the number of particles available to be loaded into the truck increases, the truck will have an increased chance to be loaded (depending, of course, on the affinity of the particle for the truck). If the number of loaded trucks is doubled, the number of particles dumped to the other side will double. Therefore, by analogy,

JA is proportional to [RA] or

$$J_A = \text{const}[RA] = k_3[RA] \tag{12.7.8}$$

How can we calculate RA when we know A and R? Let us assume that A_{total} (A_0) >> R_0, as is the likely biological case, and A_{in} = 0. We can calculate RA using the following equations and the same procedure we used for the derivation of the binding equation

$$[ML] = \frac{[M_0]\,[L]}{K_D + [L]} \tag{12.7.9}$$

The equation for the dissociation constant K_D

$$K_D = \frac{[A]_{eq}[R]_{eq}}{[RA]_{eq}} = \frac{(A)(R)}{RA} \tag{12.7.10}$$

The equation of mass balance of R

$$\mathrm{R_0 = R + RA \text{ so } R = R_0 - RA} \tag{12.7.11}$$

Since we will assume that A_0 is much greater than R_0, we will not need the mass balance for A (which is Ao = A + RA).

Substitute x into x and rearrange to get:

$$\begin{gathered}\mathrm{K_D}(RA) = (A)(\mathrm{R}) = (A)\left((\mathrm{R_0}) - (A)(RA)\right. \\ \mathrm{K_D}(RA) + (A)(RA) = (A)\left(\mathrm{R_0}\right) \\ \left(\mathrm{K_D + A}\right)(RA) = (A)\left((\mathrm{R_0}\right) \\ (RA) = \frac{(\mathrm{R_0})\,A}{\mathrm{K_D + A}}\end{gathered} \tag{12.7.12}$$

Substitute x into z gives the final equation,

$$\mathrm{J_A = k_3[RA]} = \frac{\mathrm{k_3}\left(\mathrm{R_0}\right)A}{\mathrm{K_D + A}} = \frac{\mathrm{J_{max}}A}{\mathrm{K_D + A}} \tag{12.7.13}$$

It should be clear to you from this equation that:

- a plot of J_A vs A is hyperbolic
- $J_A = 0$ when $A = 0$.
- $J_A = J_{max}$ when A is much greater than K_D
- $A = K_D$ when $J_A = J_{max}/2$.

These are the same conditions we detailed for understanding the binding equation.

This derivation is based on the assumption that the K_D can determine the relative concentrations of A, R, and AR for the interactions and the concentrations of each species during the early part of diffusion (i.e., under initial rate conditions). Remember, under these conditions, A_{out} does not change much with time. Is this a valid assumption? Examine the mechanism shown in the above figure. A_{out} binds to R with a second order rate constant k_1. RA has two fates. It can dissociate with a first-order rate constant k_2 to A_{out} + R (to give the original species) or dissociate with a first-order rate constant of k_3 to give A_{in} + R (as A moves across the membrane). If we assume that $k_2 >> k_3$ (i.e., that the complex falls apart much more quickly than A is carried in), then the relative ratios of A, R, and RA can be described by K_D. Alternatively, you can think about it this way. If A binds to R, most of A will dissociate, and a small amount will be carried across the membrane. If this happened, R would be free and quickly bind A_{out} and reequilibrate. This occurs since the most likely fate of bound A is to dissociate, not to be carried across the membrane since $k_3 << k_2$.

12.7.5: "Receptors" in Facilitated Diffusion

Two types of proteins are involved in facilitated diffusion: **carriers and channels**. **Carrier** proteins (also called permeases or transporters), such as the glucose transporter (GLUT1), move solute molecules across a membrane. **Channels/pores** facilitate the diffusion of ions down a concentration gradient by providing a pore in the membrane. In this section, we won't describe the more complicated processes of phagocytosis and endocytosis. These processes are illustrated in Figure 12.7.8.

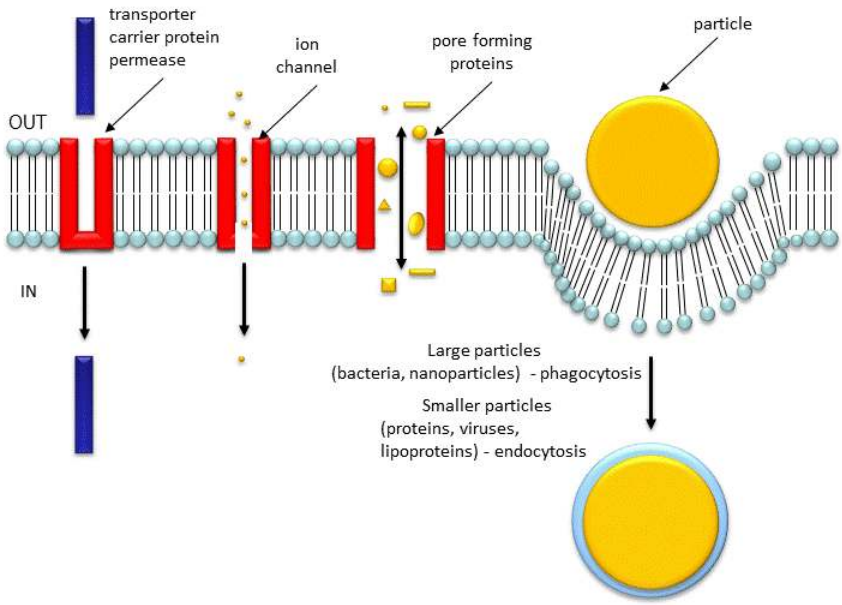


Figure 12.7.8: Movement of particles across a membrane

In the case of permeases and transport proteins, ligands bind and induce a conformational change in the receptor as illustrated in the case of the glucose transport protein shown in Figure 12.7.9.

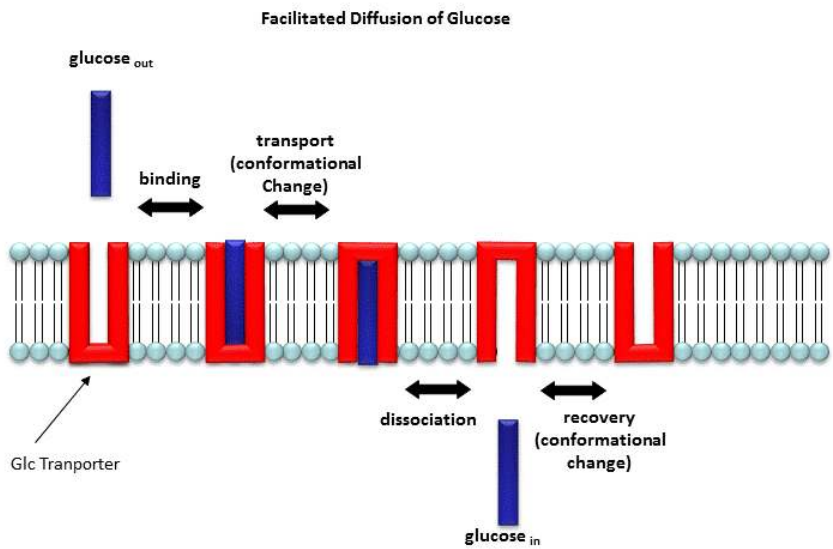


Figure 12.7.9: Model for facilitated diffusion of glucose across a membrane

A ligand can bind to the receptor (channel protein) in channels and pores, which induces a conformational change in the receptor, a "ligand-gated" channel through the membrane. This process would lead to the diffusion of many ions across the membrane (down a concentration gradient) until the channel closes (which can be induced by ligand dissociation or other events).

The mathematics we derived for the carrier proteins does not apply to the channel proteins. In addition, there are other ways to "gate" open a channel protein, which we will discuss later. Also, some transporters can move solute molecules across a membrane against a concentration gradient. These proteins require an external energy source (like ATP or coupling to the favorable collapse of a second transmembrane gradient) to drive this thermodynamically unfavored process. This is called active transport and will be discussed in the next chapter section.

- ribitol diffusion through a glycerol channel using interactive molecular dynamics

Both links above are from the Theoretical and Computational Biophysics group at the Beckman Institute, University of Illinois at Urbana-Champaign. These molecular dynamic simulations were made with VMD/NAMD/BioCoRE/JMV/other software support

developed by the Group with NIH support.

12.7.6: Carrier proteins (permeases or transporters)

Now let's look at some examples of carrier proteins:

Glucose Transport Proteins

Glucose is a key metabolic fuel, so its movement into cells is critical and highly regulated. There are multiple types of glucose transporters. **GLUT 1**, a plasma membrane protein found in most cells, is responsible for constitutive or basal glucose uptake. **GLUT 4** is involved in insulin-regulated uptake in skeletal and heart muscles and adipose cells of glucose after meals. Its official name is **solute carrier family 2,** or **facilitated glucose transporter member 4.** No structure is yet available for GLUT 4, but there is for GLUT1, which is highly expressed in cancer cells with high energy demands.

Figure 12.7.10 shows an interactive iCn3D model of a glucose transporter, GLUT1 (5eqg), bound to an inhibitor, cytochalasin B (spacefill). The inhibitor binds in the inward-open state where glucose binds.

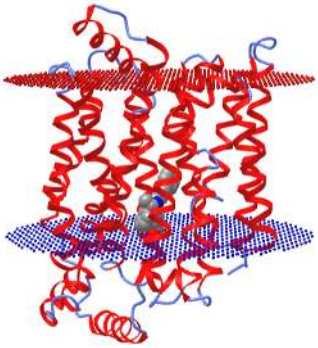


Figure 12.7.10: Glucose transporter, GLUT1 bound to the inhibitor, cytochalasin B (5eqg). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...zUTZEcTWQrQAX6

Mitochondrial ADP-ATP Carrier Protein

We will see in a few chapters that most of the ATP made in cells occurs in the mitochondrial matrix. It won't do cells much good if it stays there since it is needed in the cytoplasm and elsewhere to drive unfavored processes. Likewise, ADP is concomitantly high when ATP is depleted in a cell. What is needed is an inner mitochondrial membrane protein that can shuttle ATP out of the mitochondria and ADP in down concentration gradients. It would not make sense to need to power an uphill movement of ATP into the mitochondria from low to high concentration, driven by ATP cleavage. Let's look at the structure of the bovine ADP-ATP carrier protein, which resides in the inner membrane. Figure 12.7.11 shows an interactive iCn3D model of the bovine mitochondrial ADP-ATP carrier protein (1okc).

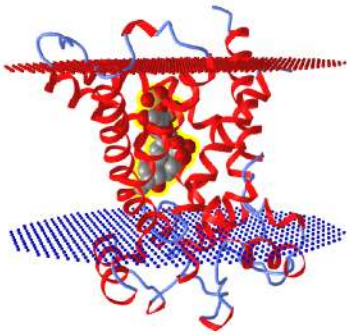


Figure 12.7.11: Bovine mitochondrial ADP-ATP carrier protein (1okc). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Zc9iLpCncYYD7gBc6

The transmembrane domain contains six alpha-helices, forming a depression leading to the inner leaflet. The cyan spacefill amino acids on the bottom of the depression are RRRMM, a motif found in nucleotide carrier protein. A conformational transition must transiently open the depression into a channel. The spacefill molecule in CPK colors represents carboxyatractyloside, a diterpene glycoside that inhibits the carrier protein.

A Special Case: Fatty acid carrier proteins

You might guess that free fatty acids, derived, for example, from lipids after the actions of lipases on triacylglycerol, would **not** need a carrier protein to move across the cell membrane since they are almost completely nonpolar. Hexanoic acid can indeed pass readily, but for solute diffusion across the membrane, size matters as well. A whole family of proteins, **Fatty Acid Transport Proteins (FATPs)** have evolved to help long-chain fatty acids across membranes. Human fatty acid transport proteins are transmembrane proteins. Its mechanism of action is unclear. No crystal structures of these are readily available. Many proteins in this class catalyze the formation of fatty acid-CoASH derivatives, an endergonic reaction powered by ATP. The mechanism for fatty acid movement across the membrane probably involves simple diffusion coupled to processes driven by ATP. However, there is still controversy about the role of passive vs. facilitated diffusion for fatty acids.

First, consider the cell's problems in moving fatty acids across two aqueous environments. Figure 12.7.12 shows a mass balance depiction of the reservoirs of fatty acid in the extracellular and intracellular environment.

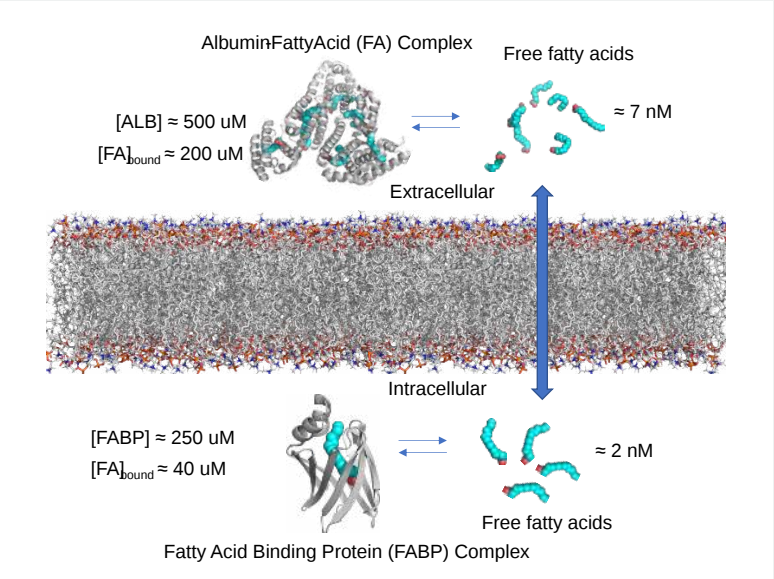


Figure 12.7.12: Mass balance depiction of the reservoirs of fatty acid in the extracellular and intracellular environment

Free fatty acids are very insoluble in aqueous solution, so their concentrations on either side of the membrane are very low, in the low nanomolar range. Hence, no great thermodynamic drive exists to move free fatty acids across the membrane. Assuming that fatty acids can reasonably transverse the membrane without a carrier protein, there would be no huge kinetic barriers to movement except for their low concentrations.

On each side of the membrane, the free fatty acids are in an "equilibrium" with protein-bound amino acids. In the blood and interstitial fluids, albumin, which can bind multiple fatty acids, is in high concentration, so it can act as a buffer to keep free fatty acids in a useful concentration range. Likewise, in the cytoplasm, **fatty acid binding proteins (FABPs)**, which typically bind just one fatty acid, are also relatively high in concentration and buffer the free fatty acids in the cytoplasm. One other note. Free fatty acids are single-chain amphiphiles, making them detergents that could easily lyse cell membranes, so their free concentrations must be kept very low compared to the critical micelle concentration.

Figure 12.7.13 shows an interactive iCn3D model of the human brain fatty acid-binding protein bound to docosahexaenoic acid (1fdq).

Figure 12.7.13: Human brain fatty acid-binding protein bound to docosahexaenoic acid (1fdq). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...M2sTpEuU6hj5RA

How do long-chain fatty acids cross the membrane? We will first examine the role of proteins. Let's look at a couple of players.

Fatty acid transport proteins (FATPs): There are six members of the human FATP family, which is also known as the Solute carrier family 27. **FATP 1 (SLC27A-1)** is found in plasma and endoplasmic reticulum membranes, and based on sequence analysis, it is a single-pass membrane protein. Its highest expression is in muscle and adipose cells. It possesses a C-terminal AMP-binding domain and acyl-CoA synthase activity. The mouse protein has an N-terminal transmembrane domain, and predictions from the human sequence show that there is likely just one transmembrane helix at the N terminus (amino acid 13-35). Given that and the absence of a 3D structure, it would appear that this protein would not bind and transfer the bound lipid across the membrane through a conformational change in the transmembrane domain. FATP 4, located in the ER membrane, is predicted to have 2 transmembrane helices, which would probably be inadequate to serve as a class translocase. It is expressed in the endoplasmic reticulum cell membrane.

Platelet Glycoprotein 4 (CD36): Another candidate is platelet glycoprotein 4, which is also called CD36, Glycoprotein IIIb, fatty acid translocase, or the thrombospondin receptor. The protein has many functions and binds many types of proteins (thrombospondin, fibronectin, collagen, or amyloid-beta) and lipids (oxidized low-density lipoprotein (ox-LDL), anionic phospholipids, long-chain fatty acids, and bacterial diacylated lipopeptides). It is present in plasma and Golgi membranes. Sequence analysis shows that it passes through the membrane twice (amino acids 7-29 and 441-463) and is palmitoylated at both N- and C-terminal ends. A PDB structure (5LGD) for most of the protein, except the putative N- and C-terminal helices, is known. In the structure, it is bound to a malarial protein (shown in grey) and two palmitic acids (spacefill) bound in the nonmembrane domains. Again, from this description, it doesn't appear that the bound fatty acids are translocated via a conformational change in the glucose and ADP/ATP receptor, as described above.

Fatty Acid Binding Proteins (FABPs): These proteins might also participate in the process. In addition to the cytoplasmic form, there is a plasma membrane-associated fatty acid-binding protein ($FABP_{pm}$), also known as FABP-1. It's the same protein as mitochondrial aspartate aminotransferase (UniProtKB - P00505 (AATM_HUMAN). It has many possible functions.

Figure 12.7.14 shows a possible model for how fatty acids may transfer or be handed off from albumin to membrane-bound GP36 or FATP-1, possibly through an intermediary protein like FABPpm (for GP36).

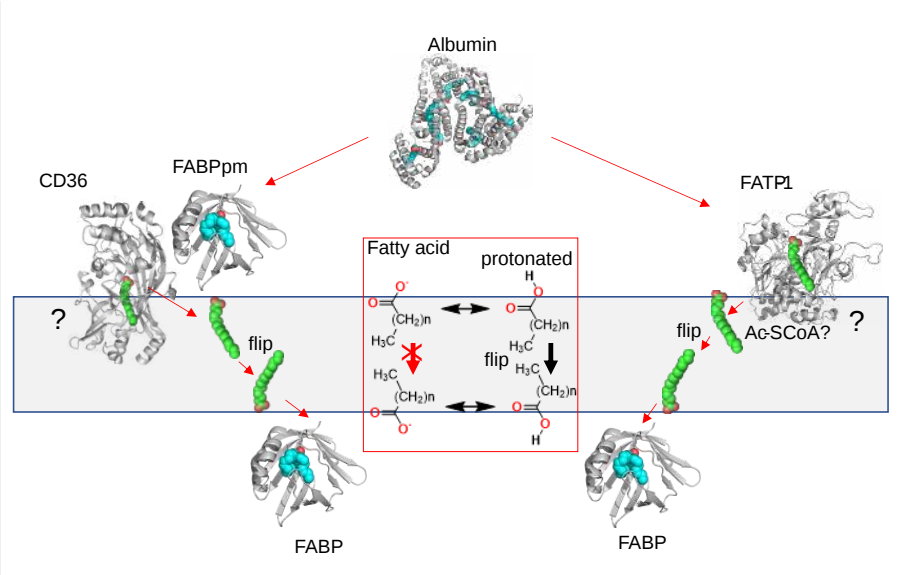


Figure 12.7.14: Possible model for how fatty acids may move from albumin to membrane-bound GP36 or FATP-1

Some proteins (albumin, FABP) deliver fatty acids to the membrane proteins (CD36, FATP-1), which deliver free fatty acids into the outer leaflet, where they flip to the inner leaflet, where they are picked up by membrane-associated cytoplasmic fatty acid binding proteins (FABP). Until structures are known for the transmembrane-bound proteins GP36 and FATP-1, whose full amino acid sequences don't suggest a classic carrier protein, this mechanism is a reasonable one. FATP-1 has acyl-CoA synthase activity, so the transferred fatty acid is likely converted to the acyl-CoASH before moving to the cytoplasm. Fatty acid transporters are also implicated in insulin resistance and type 2 diabetes.

The alternative model proposes that long-chain fatty acids, the preferential energy source for cardiac muscle, can cross the membrane by passive diffusion (red boxed area above) even if the activity of CD36 is inhibited. Diffusion depends on altering the pKa of outer membrane adsorbed free fatty acids (around 7.5) compared to free fatty acids in solution (around 4.5). The protonated fatty acid adsorbed to the membrane would move into the outer leaflet. It would flip to the inner leaflet and be picked up by cytoplasmic and/or peripheral membrane proteins. This process would also be associated with the movement of H+ across the leaflet. This type of diffusion has been observed in protein-free lipid vesicles and cells. Likewise, long-chain fatty amines (instead of carboxylic acids) can diffuse into vesicles and cells, supporting this passive diffusion if fatty acid-binding proteins don't bind the amine forms.

12.7.7: Summary

This chapter delves into the fundamental processes governing the movement of solutes across biological membranes. It begins by distinguishing the various modes of solute transport—passive diffusion, facilitated diffusion, and active transport—each driven by thermodynamic principles and subject to kinetic constraints. The chapter introduces Fick's First Law, which describes how the net flux of a solute is proportional to the concentration gradient and is modulated by the diffusion coefficient and membrane thickness. Through mathematical derivations, the relationship between flux and permeability is established, culminating in the formulation of the permeability coefficient.

The text further explores the concept of electrochemical potential, integrating the effects of both concentration gradients and transmembrane voltage differences. This understanding is extended to charged species, where the Goldman Equation is presented as a more comprehensive model than the Nernst Equation for predicting ion flux across membranes under realistic conditions.

In addition to passive diffusion, the chapter covers facilitated diffusion. Here, membrane proteins such as carrier proteins and ion channels play a crucial role in accelerating solute movement by binding substrates and undergoing conformational changes that enable transport across the membrane. Mathematical models analogous to Michaelis–Menten kinetics are derived to describe the dependence of flux on substrate concentration in receptor-mediated processes.

The chapter also discusses specific examples, including the facilitated transport of glucose via GLUT1 and the movement of ions modulated by ionophores like valinomycin. For long-chain fatty acids, the text contrasts protein-mediated transport mechanisms

involving fatty acid transport proteins and binding proteins with passive diffusion models, emphasizing the challenges imposed by their hydrophobic nature and low aqueous solubility.

Overall, this chapter integrates thermodynamic and kinetic perspectives to explain how solutes traverse membranes. It connects theoretical frameworks, such as Fick's Law and electrochemical potential concepts, with practical applications and experimental simulations, providing a comprehensive understanding of membrane transport processes that are vital for cellular function and metabolic regulation.

This page titled 12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.8: Diffusion Across a Membrane - Channels

Learning Goals (ChatGPT o3-mini)

1. **Distinguish Between Pores and Channels**
   - Explain how the size and structure of a membrane defect (pore vs. channel) determine which chemical species can cross the membrane.
   - Describe the role of gating mechanisms (e.g., ligand binding, voltage, lipid interactions, and mechanical stress) in controlling channel opening.
2. **Understand Ion Gradients and Membrane Potentials**
   - Analyze typical extracellular and intracellular ion concentrations (Na^+, K^+, Cl^-, Ca^{2+}) in neurons.
   - Relate how these gradients, combined with membrane potentials, drive ion movement during depolarization and repolarization.
   - Quantitatively interpret how small changes in ion number can lead to significant changes in transmembrane potential.
3. **Describe the Structure and Function of Pentameric Ligand-Gated Ion Channels (pLGICs)**
   - Outline the overall architecture of pLGICs, including the pentameric assembly and the “Cys-loop” motif.
   - Differentiate between orthosteric and allosteric binding sites, explaining how ligand binding at these sites triggers conformational changes that open or modulate the ion channel.
   - Discuss the physiological significance of pLGICs in fast neuronal communication and their relevance in pharmacology (e.g., analgesics and anesthetics).
4. **Examine Voltage-Gated Ion Channels: Structure and Mechanism**
   - Detail the structural organization of voltage-gated K^+ and Na^+ channels, including the arrangement of transmembrane segments (S1–S6) and the role of the S4 segment as a voltage sensor.
   - Explain the molecular basis of ion selectivity (e.g., the TTVGYG motif in K^+ channels and the DEKA motif in Na^+ channels) and the “knock-on” mechanism of ion permeation.
   - Compare the conformational states (closed, open, and inactivated) of voltage-gated channels and describe how changes in membrane potential lead to transitions between these states.
5. **Explore the Pharmacological and Toxin Interactions with Ion Channels**
   - Identify how neurotoxins (e.g., tetrodotoxin, scorpion and sea anemone toxins) interact with specific sites on Na^+ channels to alter their function.
   - Discuss the impact of anesthetics on channel behavior, including their binding to pLGICs and effects on membrane lipid dynamics (e.g., the Meyer-Overton correlation).
6. **Investigate Lipid-Gated and Mechanosensitive Channels**
   - Explain how specific lipids, such as PIP_2, act as agonists for channels like Kir2.2 to regulate ion flow.
   - Describe the unique features of mechanosensitive channels (e.g., TRAAK) and how physical changes in the bilayer (tension, curvature, or lipid removal) induce channel opening.
   - Assess the significance of these channels in maintaining or restoring the resting membrane potential and in transducing mechanical stimuli.
7. **Apply Quantitative and Structural Reasoning**
   - Use capacitance and charge equations to relate the number of ions moved during neuronal signaling to changes in membrane potential.
   - Interpret molecular models and simulations (such as those provided by iCn3D) to understand key interactions within ion channels, including selectivity filters and voltage sensors.
   - Evaluate how mutations or post-translational modifications affect channel structure, gating kinetics, and overall cellular function.

By achieving these goals, students will be able to integrate structural, biophysical, and pharmacological concepts to explain how ion channels control rapid cellular responses and contribute to complex physiological phenomena such as neural signaling and anesthesia.

12.8.1: Introduction

If you punched a hole or pore in the membrane, depending on its size, multiple types of chemical species could flow through it simultaneously. We'll talk about pores in the next section. Let's focus on channels with much smaller openings that are gated open to allow ion flow through them. They are often called **ionotropic** receptors. Channels can be "gated" open by many mechanisms, including ligand binding, change in membrane potential, lipid interactions, and mechanical stress. Opening a channel to ion flow allows rapid transmission of information (in this case, an electrical signal) into the cell, leading to rapid cellular responses. This is an ideal signaling mechanism for neural cells, which demand quick responses.

We'll show examples of each type of gating mechanism. Before we do, it is helpful to know typical extracellular and intracellular ion concentrations in a mammalian neuron, for example (Table 12.8.1).

Table 12.8.1: Extracellular and intracellular ion concentration in a mammalian neuron

ion	extracellular = $[ion]_{out}$ (mM)	intracellular = $[ion]_{in}$ (mM)
Na^+	145 mM	5-10 mM
K^+	5 mM	140 mM
Cl^-	110 mM	10 mM
Ca^{2+} (free)	1.2 mM	100 nM

When ion channels are opened in neural cell membranes, the direction of favorable thermodynamic flow is down a concentration (chemical potential) gradient. However, the direction is also affected by the transmembrane potential. Typical resting potentials of neural cells are about -60 to -70 mV (negative inside). When a nonspecific **cation** channel is gated open, the kinetic barriers to diffusion are relieved. At that moment, Na^+ ions would flow in due to both the chemical and electrical potential. K+ ions would flow out, but with less driving force since the negative transmembrane potential hinders their efflux. How are such large gradients of these ions formed? We'll answer that in the next section on active transport.

12.8.2: Pentameric ligand-gated ion channels (pLGICs)

These channels play a key role in neuronal signaling. They are ligand-gated channels. In neural systems, the ligands are neurotransmitters. All are comprised of five monomers, forming the functional channel together with the pore formed in the center of the pentameric structure. The subunits can be identical (homopentamer) or different (heteropentamer). All have "Cys-loop" motifs, so they have also been called Cys-loop receptors. Examples include the mammalian nicotinic acetylcholine, serotonin (5-HT), γ-aminobutyric (GABA), glycine, and glutamate receptors.

Figure 12.8.1 shows the generic structure of the pLGICs

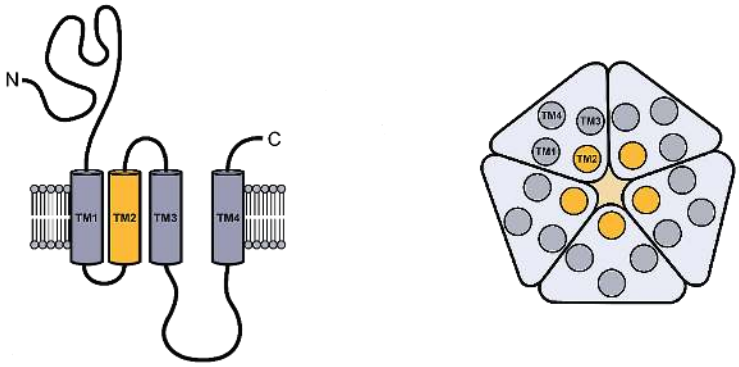


Figure 12.8.1: Generic structures of pentameric ligand-gated ion channels. Smelt, Charles L.C.; (2019) Allosteric modulation of pentameric ligand-gated ion channels. Doctoral thesis (Ph.D.), UCL (University College London). https://discovery.ucl.ac.uk/id/eprint/10075893/. Original content in this thesis is licensed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) Licence (https://creativecommons.org/licenses/by/ 4.0/)

The monomeric structure is shown on the left. Each contains four transmembrane helices (TM1-4). A top-down view of the pentameric structure is shown to the right. The pore surface forms at the interface of the central TM2 helices. The ligand (neurotransmitter) binds to the extracellular domain with contributions from all the subunits. On ligand binding, TM2 and TM3 rearrange to allow the formation of a transient pore and passive diffusion of specific ions.

pLGICs are incredibly interesting and pharmacologically relevant. In general, they have two different types of binding sites.

Orthosteric sites bind ligands in the extracellular domains. A conformational change leads to the rearrangement of helices, opening the pore when bound. The natural ligand is also called the agonist, as it promotes the ion channel's function (either neuron excitation or inhibition). The natural ligand/agonist binding opens the channel to ion flow. This can lead to activation or excitation of the neural cell if positive cations flow into the cell, which depolarizes the cell as the transmembrane potential becomes more positive. Neurotransmitters that lead to this response are excitatory. Alternatively, the binding of inhibitor neurotransmitters in the orthosteric site can lead to inhibition of the neural cell activation if the channel is a ligand-gated anion channel. This hyperpolarizes (makes the transmembrane potential more negative), leading to inhibition of neural cell activation. Inhibitors or antagonists of channel function, whose structure typically resembles at least somewhat the structure of the endogenous ligand or agonist, also bind to the orthosteric site.

Allosteric sites are distal to the orthosteric site. Ligands that bind to allosteric sites also lead to conformational changes that either augment or diminish the effect of normal ligand/agonist binding by modulating ion flow through the pore.

pLGICs interact with analgesics and anesthetics, which makes them even more interesting.

Figure 12.8.2 shows two excitatory pLGICs, the 5-hydroxytryptamine (5HT) (left) and nicotinic acetylcholine (right) receptors. The ligand binds in an orthosteric site in the extracellular domain (ECD), which is composed mostly of beta secondary structure. Allosteric sites are often found in the transmembrane domain.

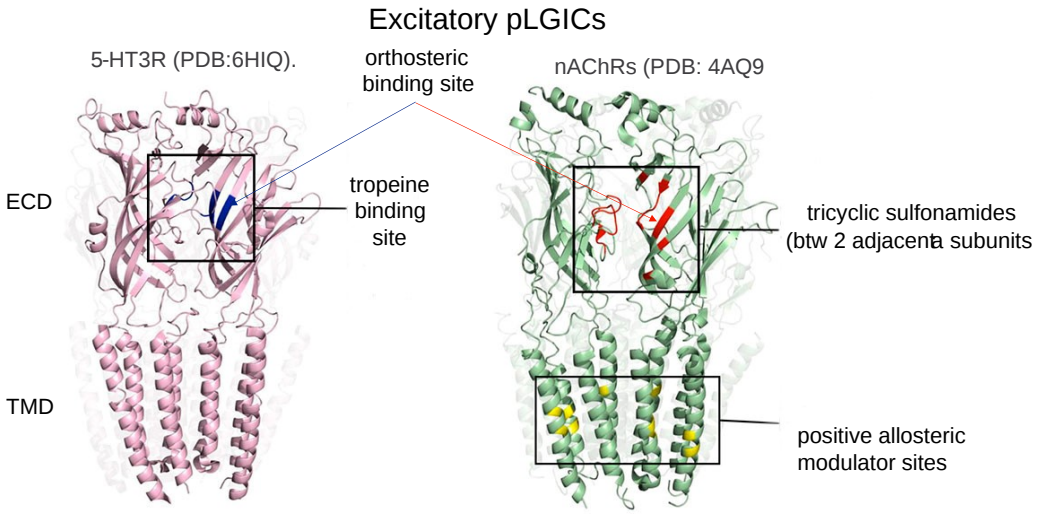


Figure 12.8.2: Two excitatory pLGICs, the 5-hydroxytryptamine (5HT) (left) and nicotinic acetylcholine (right) receptors. Adapted from Lara César O., Burgos Carlos F., Moraga-Cid Gustavo, Carrasco Mónica A., Yévenes Gonzalo E. Pentameric Ligand-Gated Ion Channels as Pharmacological Targets Against Chronic Pain, Frontiers in Pharmacology, 11 (2020), 167. https://www.frontiersin.org/article/...har.2020.00167, **Creative Commons Attribution License (CC BY)**.

Likewise, Figure 12.8.3 shows orthosteric binding sites for GABA and glycine, as well as the binding of modulators that can bind in the TMD or, in the case of benzodiazepines, in the ECD as well.

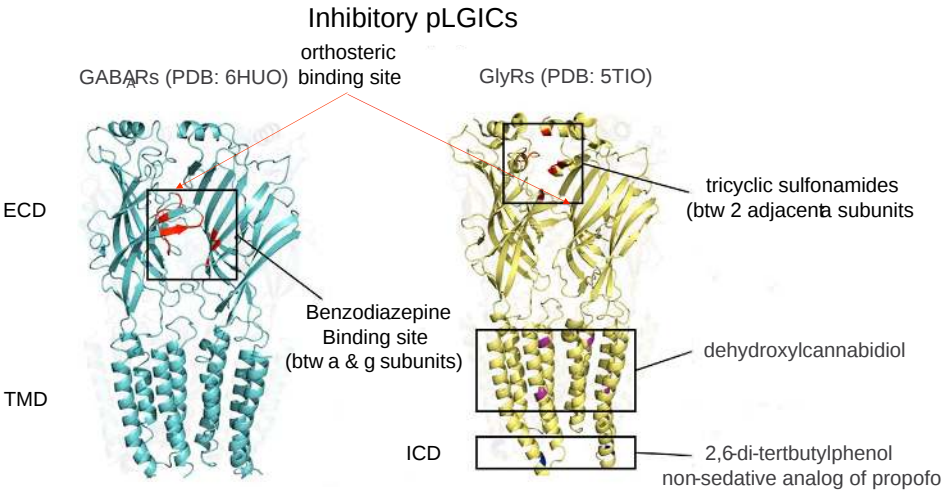


Figure 12.8.3: Orthosteric binding sites for GABA and glycine, adapted from Lara César O., Burgos Carlos F., Moraga-Cid Gustavo, Carrasco Mónica A., Yévenes Gonzalo E. Pentameric Ligand-Gated Ion Channels as Pharmacological Targets Against Chronic Pain, Frontiers in Pharmacology, 11 (2020), 167. https://www.frontiersin.org/article/...har.2020.00167, **Creative Commons Attribution License (CC BY)**.

? How do inhalational anesthetics work?

We have used them for decades, but their mechanism of action is still not clear!

Details!

The actual mechanisms of how anesthetics work are still unclear. These fascinating molecules can alter function in various organisms, including bacteria, yeast, worms, flies, plants, and animals. Of course, their effect on consciousness appears only in animals. Their selective "turning off" of a function (consciousness) of an entire organ (the brain) is stunning! One theory suggests that they exert their effects through bulk changes in the lipid bilayer, as the potency of anesthetics is

generally related to their hydrophobicity. Most are nonpolar and have long been known to work on membranes, presumably altering ion flow through neural membrane ion channels. Typical inhalational and intravenous anesthetics are shown in Figure 12.8.3, along with their date of first use.

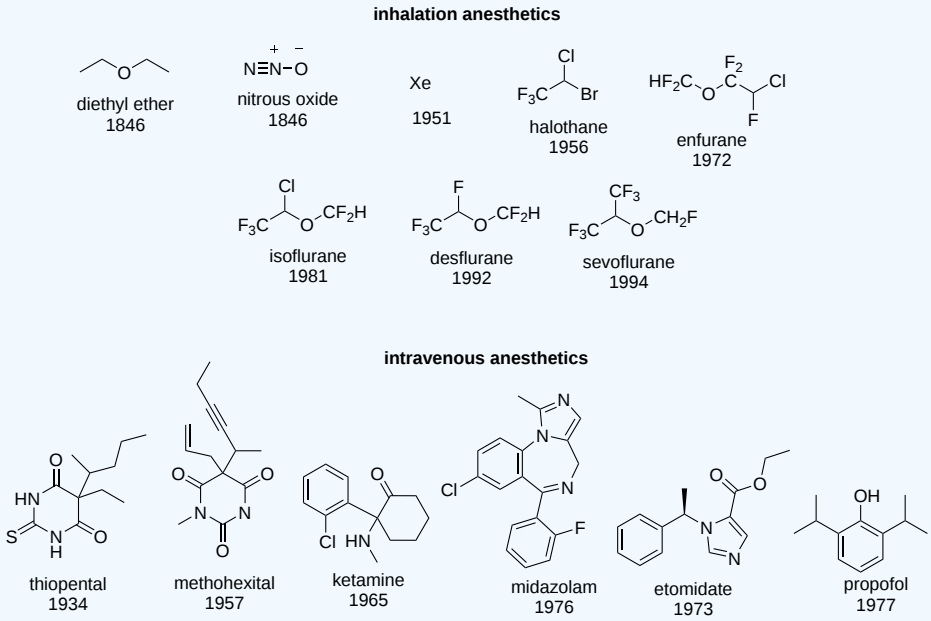


Figure 12.8.3: Structure of common inhalational and intravenous anesthetics. Adapted from Eur J Anaesthesiol. 2009 Oct; 26(10): 807–820. doi: 10.1097/EJA.0b013e32832d6b0f

A very robust correlation is found between the minimum alveolar concentration (MAC, in atmospheres) of inhalational anesthetics and their partition coefficient into olive oil (a measure of their hydrophobicity). This is illustrated in the Meyer-Overton plot shown in Figure 12.8.3.

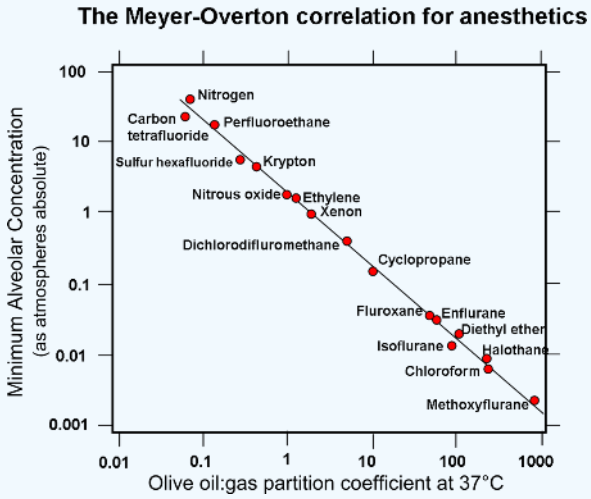


Figure 12.8.3: The Meyer-Overton plot of minimum alveolar concentrations (in atmospheres) vs partition coefficient into olive oil for inhalation anesthetics. https://commons.wikimedia.org/wiki/F...orrelation.png

The MAC is the concentration of inhaled anesthetic within the alveoli at which 50% of people do not move in response to a surgical stimulus (i.e., it is much like an IC_{50} for receptor inhibition). What's so fascinating is the range of molecular species that can act as anesthetics, including N2 and most Noble gases. Consider N_2, a nonpolar and nonreactive molecule, which comprises 80% of the atmosphere. From the graph, it is evident that it takes high pressure for it to exert its anesthetic properties. That can occur when scuba diving using regular air in tanks. Divers can experience nitrogen narcosis (also

called depth intoxication or rapture of the deep) when using just compressed air. Hence, a mixture of 21% oxygen, 35% helium, and 44% nitrogen is often used. Dentists use nitric oxide to alter consciousness and pain perception, but it does not cause general anesthesia except in some who can be hypersensitive to its use.

Additional studies suggest protein:anesthetic interactions are important. For example, their activity affects the water-soluble protein luciferase. Some molecules (like dichlorohexafluorocyclobutane), expected to have anesthetic properties based on their hydrophobicity, don't. The (S) enantiomer of isoflurane is 50% more potent than the (R) enantiomer in rats, which is hard to explain based on nonspecific partitioning into a bilayer.

Most modern theories suggest that they more directly affect specific target proteins and their proximal interacting lipids in neuromembrane bilayers. The main targets of anesthetics appear to be pLGICs. Anesthetics reduce neuron excitability and firing. Hence, you could hypothesize that they inhibit excitatory pLGICs (such as the 5HT and acetylcholine receptors) and/or activate inhibitory ones, such as the GABA and glycine receptors. pLGICs are pharmacological targets of many general anesthetics. However, anesthetic inhibition of certain GABA channels and potentiation of nicotinic acetylcholine channels have also been observed.

Recent elegant studies have shown that the inhaled anesthetics chloroform and isoflurane affect K^+ ion flow through the potassium channel subfamily K member 2, also known as the outward rectifying potassium channel protein TREK-1. The channel converts between a voltage-insensitive potassium leak channel and a phosphorylated voltage-dependent outward rectifying potassium channel. It doesn't affect the channel protein directly but indirectly through alteration in the local membrane, which affects the location of phospholipase D2 (PLD2), a protein anchored to the membrane by covalent attachment of palmitic acid. PLD2 hydrolyzes phosphatidylcholine, with a positively charged choline head group, to choline and phosphatidic acid, with a negatively charged phosphate head group

The effect of these general anesthetics appears to be on lipid rafts in neural membranes. Lipid rafts are enriched in cholesterol and saturated lipids, especially sphingomyelins such as monosialotetrahexosylganglioside1 [GM1]. Rafts are especially important in the brain, where cholesterol can reach up to 45% of plasma membrane lipids. Typical rafts are about 100 nm in diameter. In the presence of hydrophobic anesthetics, the rafts become larger and more dispersed as the anesthetic partitions into them.

Membrane proteins also partition into rafts. One such protein is phospholipase D2 (PLD2), which is targeted to the inner leaflet rafts by post-translational palmitoylation. In the presence of general anesthetics, the PLD1 laterally translates away from the disrupted and enlarged lipid raft and binds to a disordered C-terminal region of the TREK-1 protein. This localizes the PLD2 and helps activate it to produce high local concentrations of phosphatidic acid with its negatively charged head group. That group interacts with a positive region of the TREK-1 protein, inducing a conformational change in the channel, which opens it to K^+ efflux. In effect, PLD2 activates TREK-1 through the local formation of phosphatidic acids.

The opening of the channel hyperpolarizes the cell membrane (making the inside more negative and inhibiting neural activity, a hallmark of anesthesia). These concerted actions are shown in Figure 12.8.a below.

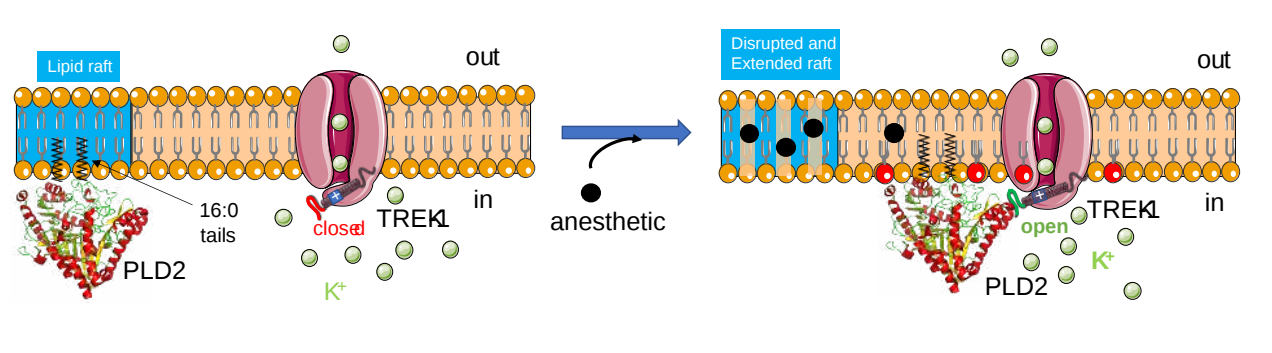


Figure 12.8.a: Effects of general https://www.pnas.org/content/117/24/13757. r Creative Commons Attribution-NonCommercialNoDerivatives License 4.0 (CC BY-NC-ND)

Deletion of the TREK-1 gene decreases the effect of the anesthetics. If a catalytically inactive mutant of PLD2 (K758R) is overexpressed, all effects of chloroform are eliminated.

We will explore two pLGIC, the eukaryotic nicotinic acetylcholine channel and a prokaryotic analog, GLIC.

12.8.2.1: Nicotinic acetylcholine channel (6cnj)

One very interesting channel is the one involved in nicotine addiction. It binds nicotine (an exogenous alkaloid) and the normal endogenous neurotransmitter, acetylcholine. Both compete for the same orthosteric binding site. Since the binding of nicotine gates the channel open, nicotine acts as an agonist. The similarities in their structures are illustrated in Figure **12.8.4**.

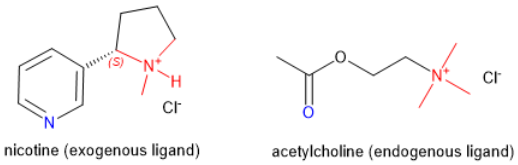


Figure **12.8.4**: Structures of nicotine and acetylcholine

The membrane protein is a ligand (acetylcholine)-gated (open-close) positive ion (Na^+ or K^+) channel involved in fast neural communication (such as at the neuromuscular junction). The quaternary structure of the pentameric receptor consists of two α4 and three β2 subunits - $(\alpha4)_2(\beta2)_3$. This isoform is the most abundant in the human brain and the one involved in nicotine addiction.

The iCn3D model (6CNJ) below has two bound nicotines (spacefill) in the extracellular domain and one Na^+ ion (spacefill) in the transmembrane domain containing the pore. The Na^+ or K^+ ions flow down a concentration gradient across the membrane in a thermodynamically favored process.

Figure **12.8.5** shows an interactive iCn3D model of the nicotinic acetylcholine channel with two bound nicotines (spacefill) in the extracellular domain and one Na^+ ion (spacefill) in the transmembrane domain containing the pore (6CNJ). (long load time)

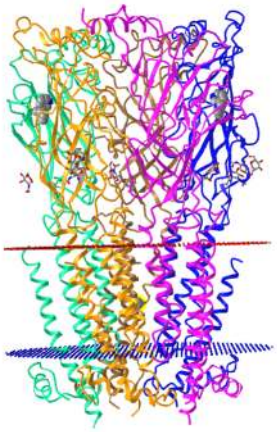


Figure **12.8.5**: Nicotinic acetylcholine channel with two bound nicotines (spacefill) in the extracellular domain and one Na^+ ion (spacefill) in the transmembrane domain containing the pore (6CNJ). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...rbXmpaha6FYNB8

The gold, blue, and brown β2 subunits are glycosylated on Asn 143 and are shown with a Man (β4) GlcNAc (β4) GlcNAc N-linked oligosaccharide. Nicotine is bound between two alpha-beta interfaces. One is shown between the green (alpha) and gold (beta) subunits, and the other is between the magenta (alpha) and blue (beta) subunits.

12.8.2.2: GLIC: A prokaryotic pLGIC

This protein is a **proton**-gated cation channel with specificity for Na^+ and K^+, which diffuse down their electrochemical gradients. In a sense, H^+s in the extracellular side act as "ligands" as the channel is opened with increasing H^+ concentration (decreasing pH) on the outside of prokaryotic cells. The protein is homologous to eukaryotic pLGICs. Structures of the protein from Gloeobacter violaceus bound to **propofol**, an anesthetic, are known. GLICs also interact with ethanol and barbiturates. Hence, they serve as models to elucidate the binding and effects of anesthetics.

In contrast to eukaryotic pLGICs, the "ligand - H^+" does not bind in the orthosteric site in the extracellular domain occupied by traditional ligands. Rather, changes in protonation states of key proton acceptors and donors in the protein lead to conformational

changes analogous to those found on binding ligands to orthosteric sites on classical pLGICs. The external pH associated with half-maximal inward current, pH_{50}, is approximately 5.1 ± 0.2.

Evidence suggests that when pH is lowered from 7 to 4, Glu 35 (distant from the orthosteric site), with a pKa ≈ 5.8, becomes protonated. It connects through other H^+ acceptors and donors in the open form through a hydrogen bond network. These include two triads of amino acids found at the interface between the extracellular (ECD) and transmembrane (TMD) domains. R192-D122-D32 comprises a conserved "electrostatic triad." The second is Y197-Y119-K248. The network allows bridging of the effects, starting with Glu 35 in the ECD into the transmembrane region, where allosteric effectors usually interact with the protein.

Figure 12.8.6 shows an interactive iCn3D model of the open form of GLIC (3P50) with bound propofol (long load time)

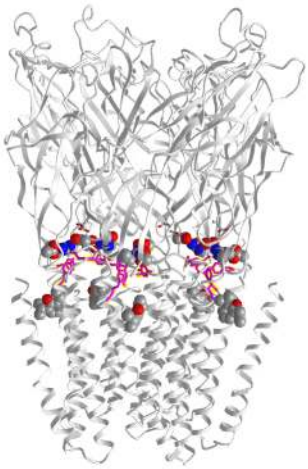


Figure 12.8.6: The open form of the prokaryotic pentameric GLIC ion channel with bound propofol (3P50). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...v92cdeunVY65R8 (long load time)

Orient the iCn3D model below with the extracellular domain (mostly beta structure) at the top and the transmembrane domain (alpha-helical) at the bottom. Key molecular players involved in the interactions described above, from the top down:

- Glu 35 (stick, color CPK)
- R192-D122-D32 electrostatic triad (sphere, CPK color)
- Y197-Y119-K248 triad (stick, color magenta)
- propofol (sphere, color CPK)

Propofol and another anesthetic, desflurane, bind at the same site localized in the upper part of the transmembrane domain of each of the five subunits.

The model below shows the mostly nonpolar (induced dipole-induced dipole) interactions between a bound propofol and side chains in the TMD. Also shown is an interaction between phosphatidylcholine and propofol.

Figure 12.8.7 shows an interactive iCn3D model showing the mostly nonpolar (induced dipole-induced dipole) interactions between one bound propofol and side chains in the TMD. Also shown is an interaction between phosphatidylcholine and propofol. (long load time)

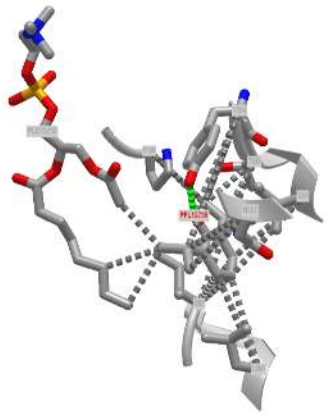


Figure 12.8.7: Propofol interactions in the transmembrane domain of the pentameric GLIC ion channel (3P50). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...s8dJnmX7SFKjLA

12.8.3: Voltage-Gated Ion Channels

In contrast to pentameric ligand-gated ion channels, which require five monomeric subunits to aggregate into a quaternary structure to form a pentameric pore, **voltage-gated ion channels** can form a channel from an aggregate of monomeric proteins, each containing a single six-transmembrane helical unit, or from a longer polypeptide containing multiple repeating 6-transmembrane helical units.

Figure 12.8.8 shows a cartoon of a common voltage-gated K^+ channel. The monomer (top), denoted as the α subunit, contains a transmembrane domain containing six helix segments. Four of these monomers aggregate to form the actual homo- or heterotetrameric channel (bottom).

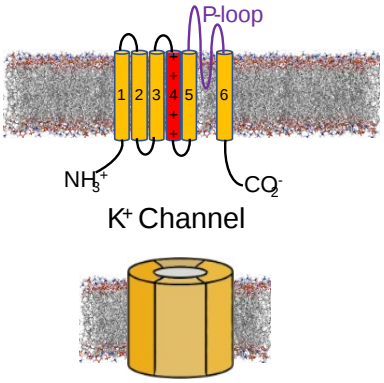


Figure 12.8.8: Cartoon of a common voltage-gated K^+ channel

The genes for the K_v channel family, which facilitate K^+ diffusion across the membrane, encode α subunits of approximately 500 amino acids and a molecular weight of about 57,000. **Four of** these α subunits come together in the membrane to form the functional channel, a tetramer of α subunits, forming a single central pore together. The α subunit can form homo- or heterotetramers since there are different α subunit encoding genes. In addition, the functional channel has smaller regulatory β subunits as well.

A cartoon structure of a typical voltage-gated Na^+ channel is shown in Figure 12.8.9. It is a **single** polypeptide (α) chain that contains four sequential repeats of the six transmembrane helical segments (I-IV). The functional channel (bottom) has just one polypeptide chain.

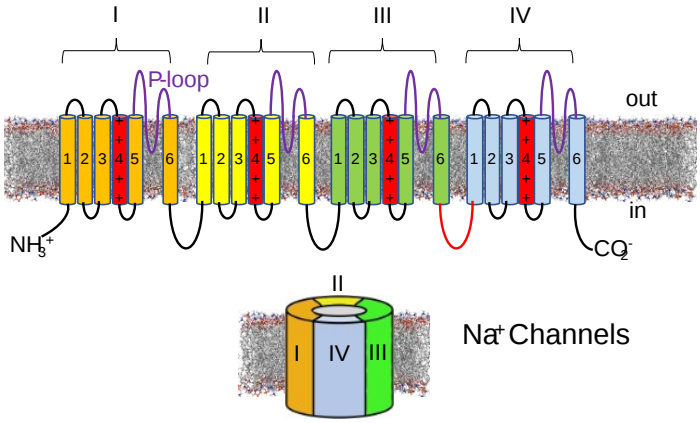


Figure 12.8.9: Cartoon structure of a typical voltage-gated Na^+ channel adapted from https://www.anaesthesiajournal.co.uk...264-4/fulltext

The Na^+ channel has a molecular weight of around 229 kDa and about 2000 amino acids (each 4x that of the K^+ channel α subunit). It is glycosylated and subjected to multiple post-translational modifications. Usually, the protein in the central nervous system is a complex of the α subunit and small additional regulatory β subunits, which modify the kinetics and voltage-dependency of the α subunit channel.

Segment (helix) 4 of each of the four repeat units illustrated above is the conserved "voltage" sensor. It contains multiple charged amino acids whose disposition changes with changes in the transmembrane potential, allowing conformational changes in the protein and gating of ion flow. Each of the four repeating units above also contains an extracellular P-loop (colored purple in segment I) connecting helix 5 and helix 6.

We will focus on K^+ and Na^+ channels below.

K^+ Permeation through Kv1.2 Channel

Voltage-dependent potassium channels (Kv) have four subunits and can be homo- or heterotetramers. They allow the voltage-gated flow of potassium ions through the membrane. Several obvious questions should arise. How can they be selective for K^+ ions? How can they allow the larger K^+ ions to flow passively through and not the smaller Na^+ ions? Secondly, how can a change in the transmembrane potential cause the channel to open or close? That question boils down to how a change in transmembrane potential can change the conformation of proteins. We'll show several iCn3D models of this protein.

Figure 12.8.10 shows a simplified view of the rat Kv1.2 channel (3lut) from top and side views (without parts of the cytoplasmic domain), showing each of the four identical monomeric subunits in a different color. Each monomer has S1-S6 transmembrane segments.

https://www.ncbi.nlm.nih.gov/Structu...cn3d/full.html

Shaker Potassium Channel Kv1.2

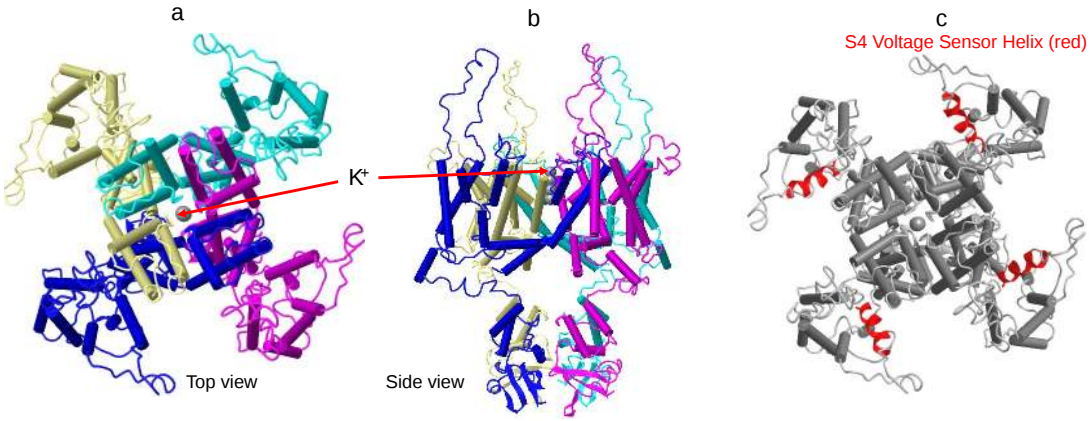


Figure 12.8.10: Rat Kv1.2 channel (3lut) from top and side views (without parts of the cytoplasmic domain), showing each of the 4 identical monomeric subunits in a different color.

The protein is found in the central and peripheral nervous systems and cardiovascular systems. K^+ ions diffuse through the center from the extracellular to intracellular side down a concentration gradient but potentially against the transmembrane potential. The four monomers in the homotetramer form one central pore.

Figure 12.8.11 shows an interactive iCn3D model showing the rat Kv1.2 channel (3lut) described in Figure 12.8.10.

Figure 12.8.11: Rat Kv1.2 channel (3lut) . (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...99VbNxVzCZyr9A

Figure 12.8.12 shows an interactive iCn3D model showing the S4 voltage sensor helix in each monomer of the Kv1.2 potassium channel (3lut) .

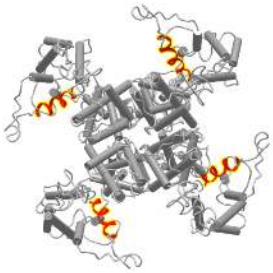


Figure 12.8.12: S4 voltage sensor helix in each monomer of the Kv1.2 potassium channel (3lut) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...nZnNTpqSwA6Ab9

Figure 12.8.13 shows a more detailed structure of the Kv1.2 potassium channel (3lut). The four monomers, which pack together to form the tetramer, are shown in light and dark grey to allow key residues highlighted in color to stand out. The details are explained below.

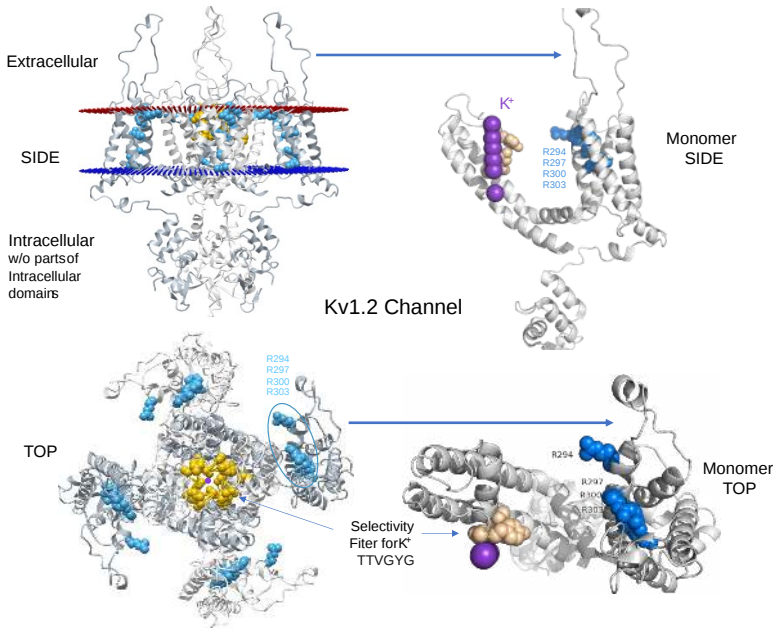


Figure 12.8.13: Detailed structure of the Kv1.2 potassium channel (3lut)

The left images show the channel from the side (top left view) and top (bottom left view). The right images show just one of the monomers with different side chains highlighted.

Figure 12.8.14 shows an interactive iCn3D model detailing key residues in the workings of the Kv1.2 potassium channel (3lut)

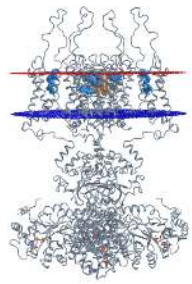


Figure 12.8.14: Key residues in the workings of the Kv1.2 potassium channel (3lut) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wYWMWmHWp1UP6A

K^+ selectivity - Even if not voltage-gated, all potassium ion channels solve the selectivity dilemma similarly. All have in the narrowest part of the pore in the center of the channel this consensus sequence - Thr-Thr-Val-Gly-Try-Gly (TTVGYG) - which is found in the P-loop. These are shown in gold and brown colors in the figure above. The -OHs in the selectivity filter can interact with a dehydrated K ion but not with a dehydrated Na ion, which can not approach close enough to form significant interactions. Surrounding the filter are twelve aromatic amino acids, which constrain the size of the pore opening. The interactions of the filter O's with the K ion make up for the energetically disfavored dehydration of the ion. The filter contains K^+ ions, which repel each other, assisting in the vectorial discharge of the ions through the membrane. These ions must form weak interactions with the selectivity filter. The pore is mostly hydrophobic, facilitating ion flow through the membrane.

Figure 12.8.15: below shows a close-up of the selectivity filter. Four Thr 374s (second Thr in the selectivity filter sequence of TTVGYG) from the four different monomers in the channel are clearly shown interacting with the top K^+ ion (gray sphere).

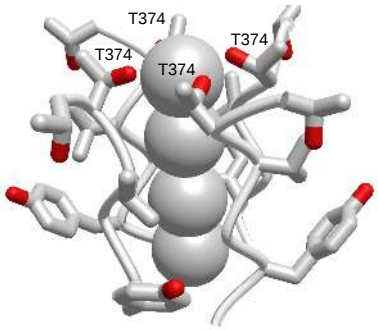


Figure 12.8.15: K^+ Selectivity filter of the Kv1.2 potassium channel (3lut)

Figure 12.8.16 shows an interactive iCn3D model detailing key residues in the K^+ selectivity filter of the Kv1.2 potassium channel (3lut). Hover over the residues to identify them.

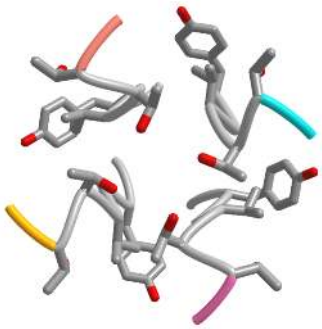


Figure 12.8.16: Key residues in residues in the K^+ selectivity filter of the Kv1.2 potassium channel (3lut) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...FFKJVVHtbpuYj6

Different voltage-gated ion channels alter ion selectivity through changes in these amino acids in the P-loop, as illustrated below in Table 12.8.2. As channels lose their specificity for K^+, they gain specificity for Na^+ and Ca^{2+}. Red highlights denote conserved residues, and yellow residues that are chemically similar.

sequence	specificity
TVGYG	strong K^+ channels
CIGYG	weak K^+, HCN channels
TVGDG	TRP channels
STFEG	ionotropic glutamate receptors
LCGEW	strong Ca^{2+} voltage-gated channels

Table 12.8.2: P-loop specificity side chains in voltage-gated ion channels

Voltage gating -

Helix S4 in each monomer in the transmembrane domain of the complex is the voltage sensor. The sequence of this helix is LAILRVIRLVRVFRIFKLSRH. Note that the arginines and lysine are highlighted in blue. They repeat every three amino acids. The voltage-sensor domain must be shielded from the nonpolar acyl of the bilayer. Four conserved Arg residues on S4, part of the voltage-sensor domains, are shielded from the lipids and coupled to an amphiphilic helix running parallel to the plane of the membrane. The arginines move under the influences of forces arising from changes in the membrane's electric field initiated by ion movement through other ion channels in the membrane. The electric field does mechanical work on the voltage sensor as the charged Arg residues are moved through the electric field. The movement is coupled through the amphiphilic helix to the pore, which changes conformation. In turn, the voltage sensor's S4 and coupled S5 helices do mechanical work on the pore by altering its conformation to open/close the pore, specifically at the activation gate of the pore. This seems quite similar to how iron movement

into the heme plane in hemoglobin on oxygenation pulls the proximal His on the F8 helix, which then transmits a conformation change to other helices in the subunit, leading to cooperative conformational changes in the tightly packed protein. About 12 charges move across the transmembrane potential field.

Channels, once open, must be inactivated. In the case of the voltage-gated potassium channel, inactivation occurs when the amino-terminal cytoplasmic domain binds to the potassium pore on the cytoplasmic side, in interaction likened to the binding of a "ball on a chain" (the cytoplasmic domain) to the pore opening. The chain acts to tether the ball domain so it may swing to bind to the pore opening. The ball domain contains both positively charged and hydrophobic regions. Where is the ball domain in the absence of inhibition? Recent studies (Oliver et al.) have shown that a positive domain can bind to proximal phosphatidyinositol 4,5-bisphosphate (PIP2) lipids on the inner leaflet of the membrane bilayer. When so bound, the channel is prevented from being deactivated. As you will see in the next section, PIP2 can also be cleaved to form diacylglycerol and inositol 1,4,5-trisphosphate when cells are activated by external factors (hormones, growth factors, etc) in the process of signal transduction.

Figure **12.8.17** shows a molecular dynamics simulation showing K^+ interaction with the channel lining and the "knock-on" mechanism showing how an incoming K^+ ion can repel a K^+ ion in the pore through the channel.

Figure **12.8.17**: Molecular dynamics simulation of K^+ movement through a channel

Voltage-gated sodium channels (Na_V)

The eukaryotic voltage-gated sodium channels (NaV) allow inward movement of Na^+ ions, which depolarize a neuron and set off an action potential in nerve and muscle cells. (For more information on neuron signaling, see Chapter 28.9). The NaV has an alpha subunit that forms the pore, plus beta subunits that associate with it and modulate its activity. Nine eukaryotic isoforms exist. NaV has four domains, I-IV, each containing segments 1-6. Each of the S1-S4 segments forms a voltage-sensitive domain, and each of the S5 and S6s forms the pore. In contrast to eukaryotic NaVs with a single chain, bacterial NaVs contain four identical subunits. Post-translational modification of the alpha subunit can regulate its activity.

As important as it is to initiate Na^+ influx to trigger neuron firing, it is equally important to turn it off to control neural signaling. The inward Na^+ current is stopped by fast inactivation within a few milliseconds.

Many neurotoxins bind to the NaV and regulate its activity. Key examples are the α-scorpion and sea anemone toxins, which both inhibit the fast inactivation of the NaV, leading to prolonged or sequential action potentials. These toxins bind to the voltage sensor in domain IV, which is key for the fast inactivation in the absence of the toxin.

The S4 helical segments in each domain are the key voltage sensors. Each of the S4 segments has 4-6 positively charged Arg and Lys side chains. On depolarization of the cell (when the inside becomes less negative and more positive), this helix moves "up"

from the cytoplasmic side (which is increasingly more positive), which opens a voltage-gated channel. Specificity for the small Na^+ ion (over the K^+ and Ca^{2+} ions) is determined mainly by four amino acids, DEKA (Asp-400, Glu-755, Lys-1237, and Ala-1529) found in P loops of domain I-IV, respectively. This selectivity filter is conserved.

Figure 12.8.18 below shows a cartoon of the NaV with two associated regulatory beta chains.

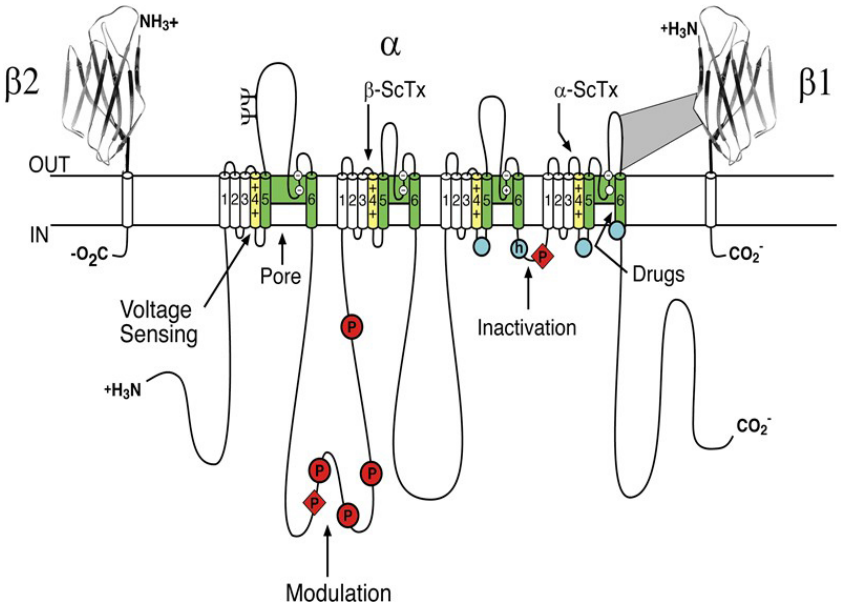


Figure 12.8.18: Cartoon of the eukaryotic voltage-gated sodium channel (Nav) with two associated regulatory beta chains. https://www.guidetopharmacology.org/...rd?familyId=82. **CC BY-SA 4.0**

Note the sites for posttranslational modification by phosphorylation and drug interactions. Cylinders represent probable α-helical segments S1-S6. Bold lines represent the polypeptide chains of the selectivity filter and tetrodotoxin binding site. The yellow S4 segments are the voltage sensors. The "h" in the blue circle is in the inactivation gate loop. Blue circles are sites implicated in forming the inactivation gate receptor. Sites of binding of α- and β-scorpion toxins (ScTX) and a site of interaction between α and β1 subunits are also shown. Tetrodotoxin is a specific blocker of the pore of sodium channels. In contrast, the α- and β-scorpion toxins block fast inactivation and enhance activation, respectively, generating a persistent sodium current that causes depolarization block of nerve conduction.

Table 12.8.3 shows the different types of neurotoxin receptor sites found in NaVs

Neurotoxin Receptor Site #	Toxin or Drug	Domains
1	Tetrodotoxin	IS2–S6, IIS2–S6
	Saxitoxin	IIIS2–S6, IVS2–S6
	μ-Conotoxin	
2	Veratridine	IS6, IVS6
	Batrachotoxin	
	Grayanotoxin	
3	α-Scorpion toxins	IS5–IS6, IVS3–S4
	Sea anemone toxins	IVS5–S6
4	β-Scorpion toxins	IIS1–S2, IIS3–S4
5	Brevetoxins	IS6, IVS5
	Ciguatoxins	
6	δ-Conotoxins	IVS3–S4
local anesthetic drug sites	Local anesthetic drugs	IS6, IIIS6, IVS6
	Antiarrhythmic drugs	
	Antiepileptic drugs	

Table 12.8.3: Neurotoxin receptor sites in NaVs. https://www.guidetopharmacology.org/...rd?familyId=82. **CC BY-SA 4.0**

Figure 12.8.19 shows a view of the effective pore in the bacterial Nav.

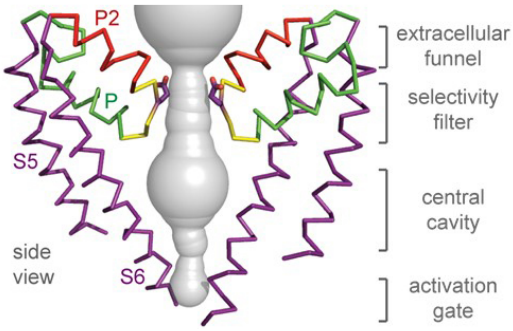


Figure 12.8.19: **Structure of the bacterial sodium channel NavAb pore B.** Architecture of the NavAb pore. Glu177 side chains in the P loop are shown in purple. The pore volume is shown in grey. The P and P2 alpha helices that form the scaffold for the selectivity filter and outer vestibule are shown in green and red, respectively. https://www.guidetopharmacology.org/...rd?familyId=82. **CC BY-SA 4.0**

As mentioned above, the bacteria NaVs have four monomeric subunits. In contrast to the K^+ channel, which requires the K^+ ions to be dehydrated to make sufficient interactions with the pore and to pass through, the Na^+ ions need to be hydrated. Figure 12.8.20 shows an interactive iCn3D model of the A and B chains of the bacterial NaV voltage-gated sodium channel pore and C-terminal domain (5BZB)

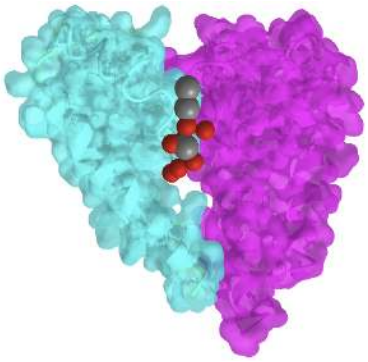


Figure 12.8.20: A and B chains of bacterial NavMs voltage-gated sodium channel pore and C-terminal domain (5BZB) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vRmDqmDHGsYLh8

One subunit is shown in a transparent cyan surface, and the second in magenta. Three Na^+ ions are shown (gray spheres). Eight water molecules are shown interacting with them in the pore. The other two subunits are not shown for clarity.

The voltage-gated sodium channel has three major conformational states:

- a basal **closed** state found at resting cell potentials, in which an activation gate occludes the pore for Na^+
- an **open** state found when a depolarizing potential is reached in the cell
- an **inactivated** state formed within 10 ms of opening of the channel when the inactivation gate with a a Ile-Phe-Met (IFM) sequence motif, found in the intracellular linker between domain III and IV (near the cytoplasmic face of the receptor) closes off the pore to further Na^+ entry.

The protein converts back to the closed state when the transmembrane potential is restored to its initial value (around -70 mV), and the positively charged S4 segments move back towards the cytoplasmic face.

A cartoon of the three-state model for the Na^+ channels (as other voltage-gated ion channels in general) is shown in Figure 12.8.21 below.

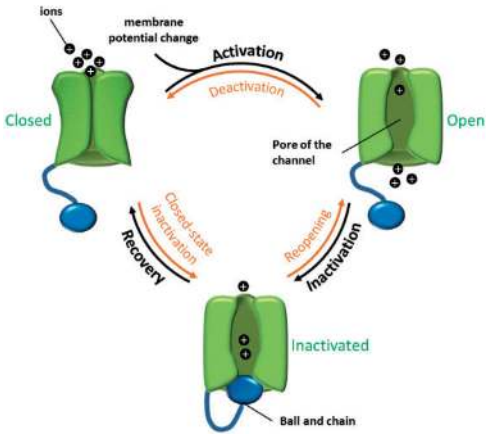


Figure 12.8.21: Three-state model for the voltage-gated ion channels. Hinard, Valerie & Britan, A & Rougier, Jean-Sébastien & Bairoch, Amos & Abriel, Hugues & Gaudet, Pascale. (2016). ICEPO: The ion channel electrophysiology ontology. Database: the journal of biological databases and curation. 2016. 10.1093/database/baw017. Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/)

The YouTube video by Pete Meighan below shows an incredibly clear description of the three conformational states of the channel and the conversion from closed to open to inactivated states.

To understand these conformational states, we need to look at the protein structure in greater detail. Figure 12.8.22 below shows multiple representations of the voltage-gated sodium channel. The pore domain PD is formed from helix segments 4 and 5 on each of the IV domains of the protein. The voltage sensor domain is formed from the S1-S4 segments, of which segment S4, containing multiple positively-charged Arg and Lys, is key.

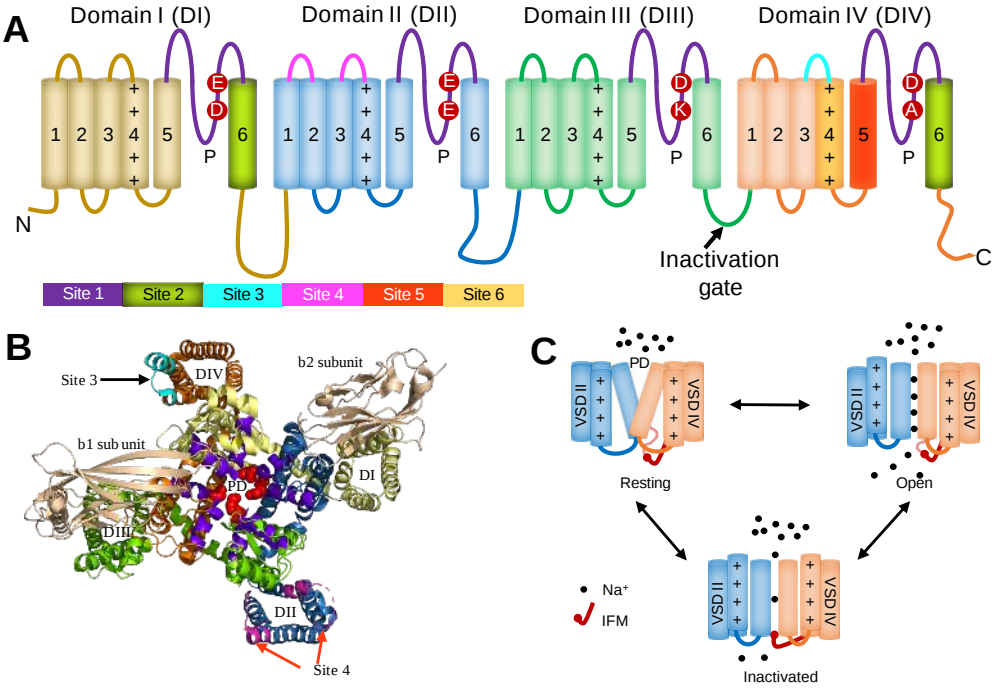


Figure 12.8.22: Structure features of the voltage-gated sodium channel. Dongol, Y.; C. Cardoso, F.; Lewis, R.J. Spider Knottin Pharmacology at Voltage-Gated Sodium Channels and Their Potential to Modulate Pain Pathways. *Toxins* **2019**, *11*, 626. https://doi.org/10.3390/toxins11110626. Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**)

Panel (**A**) shows a schematic representation of the α-subunit of the voltage-gated sodium (Na_V) channel. Four non-identical domains (DI–DIV) feature six neurotoxin receptor sites (Sites 1–6) and key residues contributing to the outer Na^+ ion selectivity filter (EEDD) and inner selectivity filter (DEKA). The connecting S5–S6 linker is called the P-loop (P), which, together with S5 and S6 segments from each domain, contributes to the formation of the Na^+ ion-selective channel pore. Sites 1-6 (colored purple, green, cyan, magenta, etc) are sites where inhibitors such as toxins bind.

Panel (**B**) shows the structure of the $Na_V1.7$ channel (PDB 6J8G). Four voltage-sensing domains (VSDs), DI (yellow), DII (blue), DIII (green), and DIV (orange), are shown with their corresponding pore-forming segments (S5 and S6) arranged to form the pore domain (PD) selective to Na^+ ions. The P-loop that contributes to forming the inner selectivity filter is colored in red spheres (DEKA), and the outer selectivity filter (EEDD) is colored in purple. The S6 segments of all four domains contribute to forming the intracellular region of the pore. Site 3 (cyan) and Site 4 (pink) are the major binding sites for spider knottins (neurotoxins). The β1 and β2 subunits that interact with DIII and DI, respectively, are highlighted in beige color.

Panel (**C**) shows a schematic of the protein's three main conformational states that control the NaV channel gating. At polarized potentials, the DI–DIV S4 segments are drawn towards the intracellular side due to the positive gating charges to render the closed conformation (down state). Depolarization relieves the forces holding the down state, and DI–DIII S4 segments are rapidly released extracellularly to open the S6 channel gate in the open conformation (up state). Note the movement of the S4 helix with its positive charge toward the extracellular side of the membrane. The DIV S4 moves up slowly compared to DI–DIII S4 and drives the fast inactivation, where the channel is occluded intracellularly by the Ile, Phe, and Met (IFM) motif. After cell repolarization, the channel returns to a closed (resting) state.

Figure **12.8.23** shows an interactive iCn3D model of a ternary complex of human Nav1.2 with the beta2 regulatory subunit and conotoxin IIIA (6J8E)

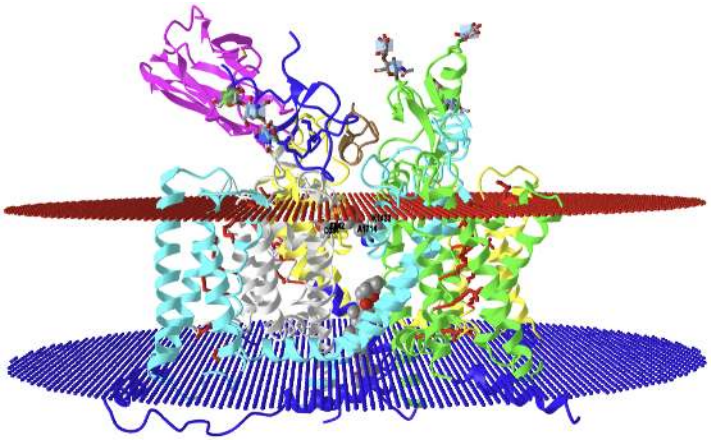


Figure **12.8.23**: Human sodium channel Nav1.2-beta2-KIIIA ternary complex (6J8E) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?aYxomdUqZHqbMmE5A

Domains I-IV are gray, yellow, green, and cyan, respectively. The separate beta-2 regulatory subunit is shown in magenta. The positive side chains in each of the S4 segments of the four domains are shown as red sticks. The side chains of the inner four amino acids (DEKA) comprising the selectivity filter are shown in spacefill, CPK colors, and labeled. The brown peptide cartoon on the extracellular side (red sphere layer for outer leaflet) is the μ-conotoxin KIIIA. The spacefill molecule in the pore near the blue sphere layer representing the inner leaflet is the neurotoxin **veratridine** (VTD), which inhibits channel inactivation and lengthens the action potential with possibly fatal consequences. A single Na^+ ion is shown at the pore's top as an orange sphere labeled Na. The μ-conotoxin blocks the pore. The Ile1488-Phe1489-Met1490 (IFM) motif, found in the intracellular linker between domain III and IV (near the cytoplasmic face of the receptor), responsible for the fast inactivation, is shown in gray spheres and labeled with single-letter codes.

The selectivity filter DEKA differs from the filter in another sodium channel (NavAb), which has four glutamates. Asp and Ala line the wall of the filter region and Glu and Lys can attract/release the Na^+ ion. The IFM motif must interact with an"inactivation gate receptor" within the protein for fast inactivation. Short intracellular loops connecting all the S4 and S5 segments are likely candidates for this. The inactivation gate has three key amino acids, F1651, L1660, and N1662.

Figure **12.8.24** shows another interactive iCn3D model of a ternary complex of human Nav1.2 with the beta2 regulatory subunit and conotoxin IIIA (6J8E) highlighting just the DEKA selectivity filter, the IFM motif, and its receptor inactivation gate F1651, L1660 and N1662.

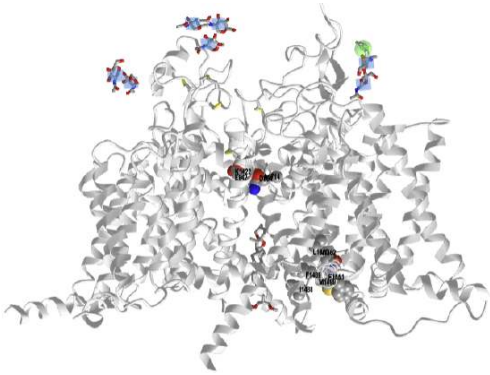


Figure 12.8.24: Human sodium channel Nav1.2-beta2-KIIIA ternary complex highlighting just the DEKA selectivity filter and the IFM and its receptor inactivation gate F1651, L1660 and N1662 (6J8E). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?tq3kXbT44GeCCUcY7

They are all shown in spacefill and labeled. Orient the molecule with the cytoplasmic domain at the top. The bilayer leaflets are omitted for clarity but the cytoplasmic pore inhibitor is still shown in sticks.

Figure 12.8.25 below shows a close-up of the bound Na^+ ion.

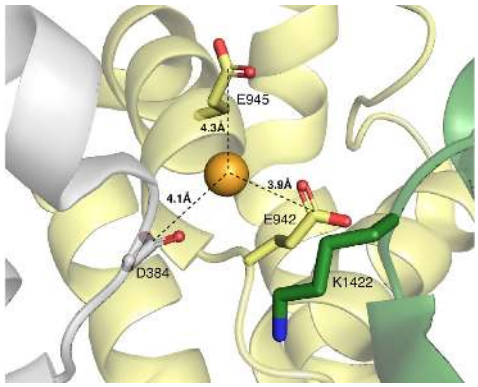


Figure 12.8.25: Sodium ion in the SF of the PDB ID 6J8E $Na_v1.2$ cryo-EM structure. The pore blocker μ-conotoxin KIIIA is not shown. Alberini et al. *J. Chem. Theory Comput.* 2023, 19, 10, 2953–2972. April 28, 2023. https://doi.org/10.1021/acs.jctc.2c00990. Creative Commons.

Structures of the closed, open, and inactivated states of NaV1.5 are now known. Key regions are shown in Figure 12.8.26 below. The open state was the hardest to solve since it closes in milliseconds to form the inactive state, as described above. Mutations in the Ile-Phe-Met (IFM) motif to QQQ prevented inactivation. This would ordinarily be deleterious to an organism, but in the presence of a small molecule inhibitor, propafenone, it was possible to isolate this state.

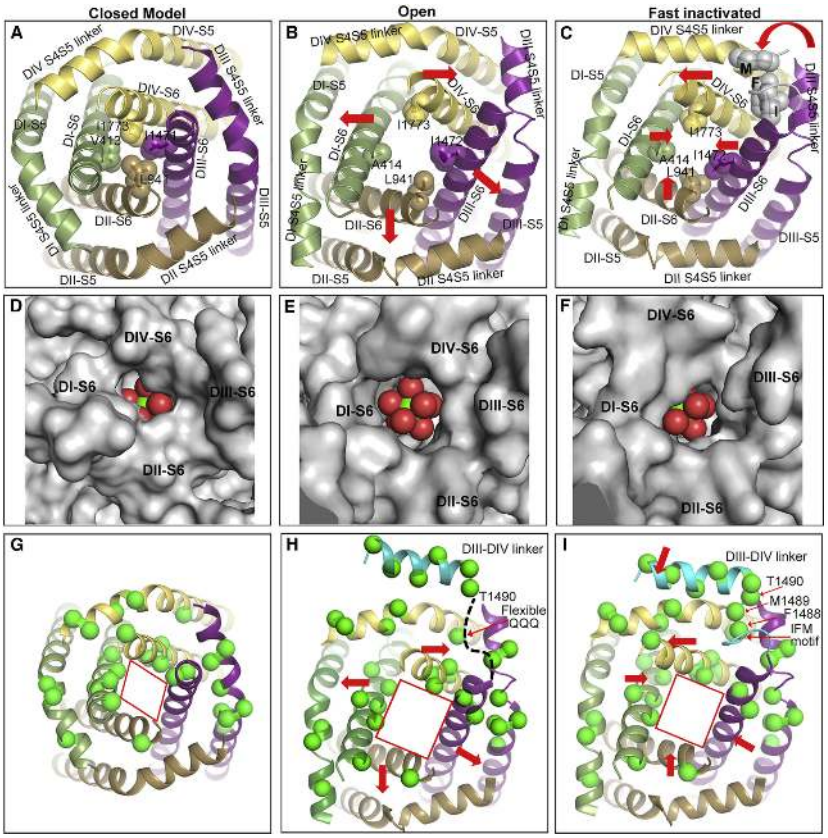


Figure 12.8.26: Closed, open, and inactivated conformations of the activation gate and the locations of arrhythmia mutations. Daohua Jiang, Richard Banh, Tamer M. Gamal El-Din, Lige Tonggu, Michael J. Lenaeus, Régis Pomès, Ning Zheng, William A. Catterall. Open-state structure and pore gating mechanism of the cardiac sodium channel. Cell, Volume 184, Issue 20, 2021, 5151-5162.e11, ISSN 0092-8674,. https://doi.org/10.1016/j.cell.2021.08.021. Reprinted with permission from Elsevier. May not be sublicensed, assigned, or transferred to any other person without the publisher's written permission.

Panel (A) shows the closed activation gate of $Na_V1.5$ generated by MODELER based on the resting-state structure of Na_VAb (PDB: 6P6W), sealed by a square of hydrophobic side chains of hydrophobic residues V413, L941, I1471, and I1773 (spacefill, black) that in the closed state completely seal off the cytoplasmic opening in the pore. These rings of amino acid side chains come together on conformational changes resulting from the engagement of the IFM motif with its internal receptor.

Panel (B) shows the open activation gate of $Na_V1.5$/QQQ triple mutation. Red arrows indicate the directions of movement of S6 segments compared to the resting state.

Panel (C) shows the partially open but nonconductive activation gate of $rNa_V1.5_C$ in the inactivated state. Red arrows indicate the directions of movement of the S6 segments compared to the open state.

Panels (D–F) show structures from (A) to (C) are shown in space-filling surface representation with hydrated Na^+ placed in the central cavity behind the activation gate. Red and green spheres represent water and Na^+, respectively. van der Waals distances measured across the orifice of the activation gate are 4.3 Å (*D*I-*D*III) × 2.8 Å (*D*II-*D*IV) for the resting/closed state, 6.9 Å (*D*I-*D*III) × 5.0 Å (*D*II-*D*IV) for the inactivated state, and 7.3 Å (*D*I-*D*III) × 8.2 Å (*D*II-*D*IV) for the $Na_V1.5$/QQQ open-state structure.

Panels (G–I) show structures from (A) to (C) with the locations of arrhythmia mutations causing LQT-3 overlaid as green spheres.

The spacefill models in D-E clearly show that the cytoplasmic opening is "open" in the open state E (obviously) and occluded to increasing degrees in the inactive state (F) and closed state D.

Finally, figure 12.8.27 shows an interactive iCn3D model of the electrostatic surface potential of the rat $Na_v1.5$ channel (6UZ3)

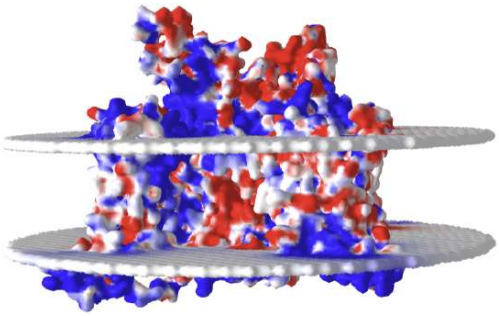


Figure 12.8.27: Electrostatic surface potential of the rat $Na_v1.5$ channel (6UZ3) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UZdQeUUPtYGgR6

The red indicates negative potential, and the blue indicates positive potential. The white is neutral. Note that the bilayer representations are essentially neutral with respect to their effect on the overall surface potential. Tilt the model to view the top and bottom entrances to the pore and the overall negative charge density expected in the net attraction of the positive Na^+ ions.

12.8.4: Lipid-Gated Ion Channels

Membrane receptors are in the lipid bilayer, so it should not be surprising that specific lipids might bind and trigger conformational changes in the receptor, mediating a specific biological activity. The specific lipid might form during an upstream event and then bind to the receptor, triggering a function in signal transduction, which we will explore in the next chapter. Look at receptors gated by binding the lipid phosphatidylinositol bisphosphate (PIP_2). It is found in small concentrations in membranes and can be cleaved in by phospholipase C to form a small polar intracellular signaling molecule, IP_3 (discussed in Chapter 12). PIP_2 can also recruit proteins to the membrane and participate in signaling events. For this section, we will discuss how it binds to a K^+ ion channel protein and opens so K^+ can flow into cells (the opposite direction of the normal efflux), which helps control transmembrane potential.

Kir2.2 -Inward rectifier potassium channel Kir2.2

The names of this channel can best be understood if you understand the meaning of the word rectify. The verb has several meanings, but for us, the best definition is to **correct**. What does this channel correct? When is it functional? We'll explore the details more in the next chapter, but first, we need a better understanding of what happens to the actual Na^+ and K^+ ion concentrations outside and inside the cell during neural activation. How much do they change? We need this to understand the driving force for the Kir2.2 channels that move K^+ into the cell, which is opposite the usual direction. In that sense, it is "rectifying" the K^+ concentrations.

? How many ions actually move on opening a channel?

It's less than you think!

Click for the details!

When a neural cell is activated or fires, the cell membrane potential goes from a resting potential of around -60 to -70 mV (inside negative) to a more positive potential as Na^+ ions enter the cell through voltage-gated Na^+ channels (after an early neurotransmitter-gated ion channel is open after ligand binding) with the ions flowing down a **chemical and electric potential** gradient. At a certain membrane potential (about +30 mV), K^+ channels open, allowing an efflux of K^+, also down both a **chemical and electric potential** gradient, returning the cell potential close to its equilibrium value potential of around -60 to -70 mV (inside negative). But how much do the actual K^+ and Na^+ ion concentrations change in this process? The somewhat counterintuitive answer is hardly any at all!

We need to understand first how the membrane acts as a capacitor. The charge Q on the surface of a plate or side of a membrane is proportional to the voltage across the plate or membrane. Figure 12.8.28 shows how a membrane with a transmembrane potential acts as a capacitor. The dielectric medium in the capacitor will determine how quickly the charges

on the plate will dissipate. When the medium is an insulator resistant to charge flow, the plates remain charged longer. The same is true for the membrane. The hydrophobic bilayers act as an insulator and resist the discharge of the membrane potential. The bilayer offers high resistance (low conductance) to charge flow. Only when channels are open, and the ions become reasonably permeable to flow does the membrane potential change over short periods. The following derivation is adapted.

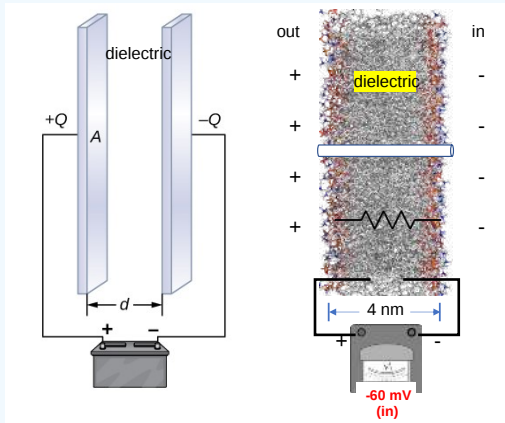


Figure 12.8.28: A membrane as a capacitor. After https://openpress.usask.ca/physics15...d-capacitance/. **License**: CC BY: Attribution

It makes sense that the stored charge (Q) on either side of the membrane is proportional to the membrane voltage. We can write the following simple equations:

$$Q \propto V$$
$$Q = CV \tag{12.8.1}$$

where C, the proportionality constant, is the capacitance with units of the Faraday (which you remember from introductory chemistry). Let's normalize this equation for an area of 1 cm^2. The measured capacitance of lipid bilayers is about 10^{-6} F/cm^2. Let's assume a voltage change from -70 mV to + 30 mV for a total of 0.1 V. Hence

$$Q = \frac{10^{-6}\ \text{F}}{\text{cm}^2}(0.1\ \text{V}) = \frac{10^{-7}\text{Coul}}{\text{cm}^2} \tag{12.8.2}$$

Let's change that into the number of elementary charges on the surface of the membrane per μm^2 surface area.

$$\left(\frac{10^{-7}\text{Coul}}{1\ \text{cm}^2}\right) x \left(\frac{1\ \text{ion}}{1.6 \times 10^{-19}\text{Coul}}\right) = \frac{6.25 \times 10^{11}\ \text{ions}}{\text{cm}^2} \times \frac{1\ \text{cm}^2}{10^8 \mu\text{m}^2} = \frac{6,250\ \text{ions}}{\mu\text{m}^2} \tag{12.8.3}$$

What does this mean for an ordinary eukaryotic cell? Let's model the cell as a sphere with a diameter of 10 μm.

Knowing the equations for the volume ($V = (4/3)\pi r^3$) and surface area ($4\pi r^2$), the volume of 10 μm cell is about 524 μm^3, and the surface area is 314 μm^2. The table below shows the cell's resting ion concentrations and the number of ions in the 524 μm^3 volume of the cell (column 3).

Using the calculated value of 6250 ions moved/μm^2, the total number of K^+ and, accordingly, Na^+ ions that move across 314 μm^2 of total cell membrane surface area is about 2 million. The results are shown in Table 12.8.3 below.

1. ion	2. $[ion]_{intracellular}$ (mM)	3. # $ions_{intracellular}$	4. total ions moved during neuron response	5. change in $[ion]_{intracellular}$
sodium	10 mM	3.2×10^9 Na^+	~ 2,000,000 in on depolarization	~6.3×10^{-2} mM (~0.6% change)
potassium	150 mM	4.7×10^{10} K^+	~ 2,000,000 out on repolarization	~6.3×10^{-2} mM (~0.6% change)

Table 12.8.3: How many ions move across a membrane

$$\frac{2 \times 10^6 \mathrm{Na}^+ x \left(\frac{1 \text{ mol}\mathrm{Na}^+}{6.022 \times 10^{23} \mathrm{Na}^+}\right)}{524 \mu M^3 \times \left(\frac{1 \text{ L}}{10^{15} \mu M^3}\right)} = 6.3 \times 10^{-6} \mathrm{M} = 6.3 \times 10^{-3} \mathrm{mM} \tag{12.8.4}$$

The actual change in intracellular Na^+ ion concentration on excitation is only about 0.6% of the initial $[Na^+]_{intracellular}$ on excitation and depolarization of the cell. We can also assume that the K^+ ion changes to the same degree in repolarization. So when the Na^+ and K^+ channels open, the "flood gates" are **not** opened. Permeability increases, but this leads to a tiny influx of Na^+ ions, **in**sufficient to change the intracellular Na^+ concentration significantly. It is, however, enough to change the transmembrane potential! It's a misconception that there are significant changes in ion concentration across the membrane on depolarization and repolarization of the cell. The Goldman-Hodgkins-Katz equation (previous section) shows that the membrane potential is a function of both concentrations and permeability coefficients.

Now, we can present the inward rectifying potassium channels, which facilitate the movement of K^+ **into** the cell from the outside. This is in the opposite direction of the usual flow of the ions. Since the ions are moving towards higher levels of K^+ inside the cell, it would appear this would be an example of either passive diffusion favored only by the electrical potential or a case of active transport. None of these is true. K^+ ions diffuse from outside to inside the cell since outward diffusion is substantially blocked by molecules such as polyamines and Mg^{2+}! We don't have to invoke active transport or a violation of the basic rules of thermodynamics. The channels thus display strong inward currents and weak outward ones.

The channels allow large conduction of K^+ ions if the membrane potential is more negative compared to the resting K^+ ion equilibrium potential, but less if it is more positive. Hence, the net effect is to maintain the resting K^+ potential.

There are several subfamilies of the Kir channel. They are always potentially active (open) except those gated by G-proteins (see next chapter) or by ATP binding. (these are involved in metabolism). Phospholipids and proteins regulate Kir activity. Now, we will explore the Kir 2.2 channel, which is gated-open by PIP_2, a membrane lipid, not by changes in the transmembrane potential. Yet by opening this channel and allowing the inward flow of K^+ ions, PIP_2 regulates the transmembrane potential. It is the agonist for the Kir 2.2 channel.

Instead of having monomeric units with S1-S6 transmembrane helical segments with a voltage sensor (S4), P loop, and S5-S6 selectivity filter, it has just two transmembrane helices. The N- and C-termini are both in the cytoplasm, connected by an extracellular loop (H5), which helps form the prototypical K^+ selectivity filter with the consensus sequence T-X-G-Y(F)-G. Similar to the voltage-gated K^+ channel, four of these aggregate to form homo- or heterotetramers in the membrane. Figure 12.8.29 illustrates these points.

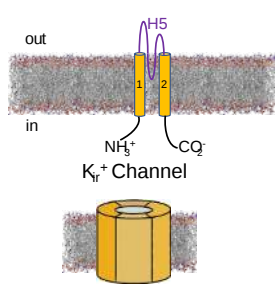


Figure 12.8.29: The Kir potassium channel

The proteins have large intracellular domains (ICD). On binding of PIP_2 to the region between the transmembrane and ICD, which produces a large conformational change, allowing K^+influx. The model below shows the R186A mutant tetrameric Kir 2.2 channel (3SPG) with four bound PIP_2 analogs containing two short fatty acids (octanol) esterified to the glycerol backbone. There appear to be two binding sites for lipids: a nonspecific site in the TMD and a specific one for PIP_2 in the ICD.

Figure 12.8.30 shows an interactive iCn3D model of the inward rectifying R186A mutant tetrameric Kir 2.2 channel (3SPG) with four bound PIP_2 analogs containing two short fatty acids (octanol) esterified to the glycerol backbone. Hover over the residues to identify them.

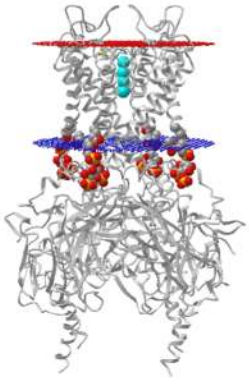


Figure 12.8.30: Inward rectifying R186A mutant tetrameric Kir 2.2 channel (3SPG). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3...S5vuBsdbSSqwV8 (long load time)

Figure 12.8.31 shows an animation that shows the monomeric Kir protein morphing from the apo state (without PIP_2, 3JYC) to the PIP_2-bound state (3SPI).

Figure 12.8.31: Monomeric Kir protein morphing from the apo state (without PIP_2, 3JYC) to the PIP_2-bound state (3SPI)

Another PIP_2-gated ion channel is the transient receptor protein (TRP) channel. These channels play a role in vascular tone. Smooth muscle cells have TRPC3 and TRPC6 channels, Na^+ or Ca^{2+} channels that cause depolarization, leading to smooth muscle contraction and vasoconstriction.

12.8.5: Mechanosensitive channels

We saw in the previous chapter that some pores (not channels) are gated open not by voltage, specific agonists such as neurotransmitters, or even specific lipids like PIP_2, but by more general changes in membrane bilayer properties (membrane tension, curvature, pressure). Likewise, specific channel proteins can be gated by changes in the physical properties of the bilayer. We will consider one mechanically-active channel, TRAAK.

TRAAK - Potassium channel subfamily K member 4

The protein is a K^+ ion channel that is not sensitive to voltage changes but is open in response to the mechanical deformation of the bilayer. The channel can be opened by making the cytoplasm basic and raising the temperature. It is involved in pain sensation and pressure transduction. The TRAAK 4 opening is blocked by lipids from the inner leaflet, which occludes the pore. If physical changes remove the lipid, transmembrane helix 4 rotates, which prevents the lipid from blocking the channel, thereby opening it. Further changes in the coupled transmembrane helices 2 and 3 stabilize the opening.

Figure 12.8.32 shows an interactive iCn3D model of the TRAAK channel protein (4wff) in the **closed** state. Hover over the residues to identify them.

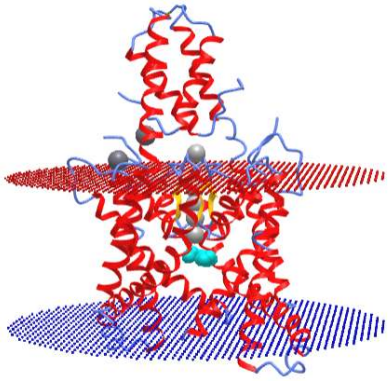


Figure 12.8.32: **Potassium channel subfamily K member 4** TRAAK channel protein in the **closed** (4wff) state. Hover over the residues to identify them. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...UzX6p5d36VicdA

The K^+ ions are aligned in the channel. Decane, a nonpolar molecule, is shown in spacefill and colored cyan. The decane is probably decanoic acid, and the carboxyl group was not defined in the structure due to its high flexibility. This suggests that the "decanoic acid" is not tightly bound. It binds through the cytoplasmic side through an opening in the membrane protein. In the open state, transmembrane helix 4 rotates, blocking access to the cavity. Hence, lipid binding closes the channel.

12.8.6: Summary

12.8.6.1: Chapter Summary

This chapter provides an in-depth exploration of the structure, function, and regulation of various ion channels in biological membranes—a central topic in cellular biochemistry and neurobiology. Key points include:

- **Channels vs. Pores and Gating Mechanisms:**
  - The chapter begins by distinguishing between large membrane pores and the much smaller ion channels.
  - It emphasizes the importance of gating mechanisms—controlled by ligand binding, voltage changes, lipid interactions, or mechanical stress—that regulate ion flow through these channels.
  - These gating events enable rapid signaling, such as the quick electrical responses in neural cells.
- **Ion Gradients and Membrane Potential:**
  - A discussion of typical extracellular and intracellular ion concentrations (Na^+, K^+, Cl^-, Ca^{2+}) sets the stage for understanding how ions move according to both concentration gradients and electrical potentials.
  - The chapter illustrates how even small changes in ion numbers, quantified using capacitance principles, can significantly alter the membrane potential without greatly changing overall ion concentrations.
- **Pentameric Ligand-Gated Ion Channels (pLGICs):**
  - pLGICs, also known as Cys-loop receptors, are highlighted as crucial components in neuronal signaling, mediating fast synaptic transmission through their assembly of five subunits.
  - The structural layout includes extracellular ligand-binding sites (orthosteric) and additional regulatory sites (allosteric) that modulate channel function.
  - Pharmacological relevance is underscored by the interactions with various anesthetics and neurotoxins.
- **Voltage-Gated Ion Channels:**
  - The chapter details the structure and function of voltage-gated K^+ and Na^+ channels.
  - It describes the critical role of the S4 segment as a voltage sensor that responds to changes in membrane potential by undergoing conformational changes to open or close the channel.
  - Ion selectivity is explained through specific sequences (e.g., the TTVGYG motif in K^+ channels and the DEKA motif in Na^+ channels) that ensure proper ion discrimination and transport.

- Multiple conformational states—closed, open, and inactivated—are discussed, illustrating how channels transition between these states during neuronal signaling.

- **Pharmacological and Toxin Interactions:**
  - Various neurotoxins and drugs are examined, showing how they bind to specific receptor sites on ion channels to modulate activity.
  - The relationship between anesthetic potency and hydrophobicity (Meyer-Overton correlation) is explored, along with direct interactions between anesthetics and target proteins.

- **Lipid-Gated and Mechanosensitive Channels:**
  - Beyond ligand and voltage gating, the chapter introduces channels regulated by lipids such as PIP_2, using Kir2.2 as an example of inward rectifying potassium channels that modulate the resting membrane potential.
  - Mechanosensitive channels, like TRAAK, are presented as sensors of physical changes in the membrane, illustrating how mechanical forces can lead to conformational changes and channel opening.

- **Integration of Structural and Quantitative Analysis:**
  - The use of molecular models and simulations (e.g., interactive iCn3D models) helps visualize key features such as selectivity filters, voltage sensors, and binding sites.
  - Quantitative reasoning, including calculations of ion movement and membrane capacitance, ties together the physical principles that underlie rapid electrical signaling in neurons.

Overall, the chapter equips biochemistry majors with a comprehensive understanding of how ion channels contribute to rapid cellular communication, the molecular basis of ion selectivity and gating, and the broader implications for pharmacology and physiology.

This page titled 12.8: Diffusion Across a Membrane - Channels is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.3: Diffusion Across a Membrane - Channels** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.9: Diffusion Across a Membrane - Pores

Learning Goals (ChatGPT o3-mini)

- **Describe the mechanisms of pore formation:**
  - Explain how pores in lipid bilayers form through both lipid rearrangements (e.g., toroidal pores) and protein assembly.
  - Compare pore formation via electroporation versus protein-mediated pore formation.
- **Identify key pore-forming proteins and their functions:**
  - Recognize various pore-forming proteins (PFPs) including bacterial toxins (e.g., ClyA/HlyE), membrane attack complex components, and porins like OMF and VDAC.
  - Discuss how these proteins regulate the passage of ions and molecules based on their structure and gating mechanisms.
- **Analyze the role of membrane lipid dynamics in pore formation:**
  - Evaluate how changes in lipid packing can lead to different pore environments (hydrophilic vs. hydrophobic linings).
  - Understand how electroporation alters lipid bilayer properties to form pores.
- **Examine the biophysical principles underlying ion and molecule transport:**
  - Apply Ohm's law and related concepts (e.g., current density and conductivity) to explain how ions move through open and closed pores.
  - Discuss the influence of electrochemical gradients, electrophoresis, and electroosmosis on solute movement.
- **Compare specialized channels and transport systems:**
  - Contrast the structure and function of aquaporins with other pore-forming proteins, emphasizing the mechanisms that prevent unwanted ion flow (e.g., proton exclusion).
  - Differentiate between mechanosensitive ion channels (MscS and MscL) and voltage-gated channels regarding their structural responses to physical forces and membrane tension.
- **Understand the architecture and function of gap junctions and the nuclear pore complex:**
  - Describe how connexins assemble into gap junction channels and the physiological implications of their mutations.
  - Detail the structure of the nuclear pore complex (NPC), including the roles of nucleoporins, FG-repeat domains, and the overall transport mechanism for macromolecules.
- **Integrate structural and dynamic data in pore biology:**
  - Interpret molecular dynamics simulations and structural models (e.g., cryoEM, iCn3D models) to understand the conformational changes during pore formation and gating.
  - Assess how the interplay between protein conformation and lipid environment determines pore functionality in various cellular contexts.

These goals aim to deepen students' understanding of both the structural and functional aspects of pore formation, emphasizing the integration of biophysical principles with molecular mechanisms in a cellular context.

12.9.1: Pores and Pore-Forming Proteins (PFPs)

If you form a pore in a cell bilayer, molecules of all sizes could move either way based on their electrochemical potential. They will move from regions of a higher to lower electrochemical potential in a thermodynamically favorable process. Hence, movement through pores represents a special case of facilitated diffusion when part of the driving force is a concentration gradient and part an electrical potential. Several questions might come to mind.

- What proteins are involved in pore formation?
- How is the specificity of solute movement through the pore regulated?
- What is the mechanism of pore formation?

Pore formation can lead to cell death, which is the function of some pore-forming proteins (PFPs), including the toxin Hemolysin E (also known as HlyE, ClyA, SheA) secreted from E. Coli and S. Aureus. Human proteins also form a membrane attack complex (examples include the membrane attack complex-perforin/cholesterol-dependent cytolysin (MACPF/CDC) superfamily and the membrane attack complex (MAC). The MAC consists of an assembly of proteins involved in the complement system (part of the effector branch in the innate immune system), which leads to cell death of Gram-negative bacteria like E. Coli. Figure 12.9.1 illustrates the assembly process of the membrane attack complex and the complexity of interactions required to form a lethal pore in a cell.

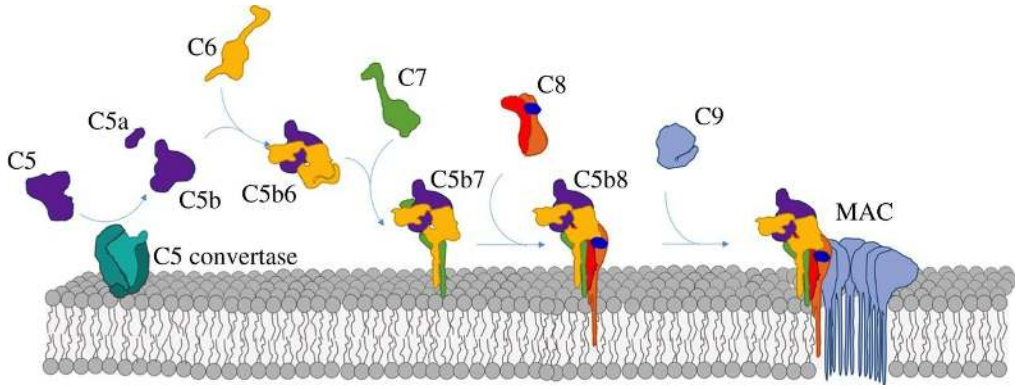


Figure 12.9.1: Human membrane attack complex. Charles Bayly-Jones et al. Philosophical Transactions of the Royal Society B: Biological Sciences, Volume 372, Issue 1726. June 2017 https://doi.org/10.1098/rstb.2016.0221. The Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/

Complement protein C9 can adopt a soluble form or membrane form, which, in aggregate, forms the pore, leading to cell death. This is a common feature of PFPs.

PFPs could create a pore by altering membrane lipid packing to form a toroid-like hole (Figure 12.9.2) and/or by inserting in a membrane and forming a pore within the protein complex itself. In either mechanism, lipid packing is altered.

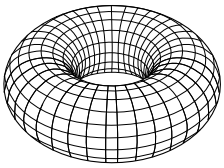


Figure 12.9.2: By GYassineMrabetTalk - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/inde...?curid=3134923

Biophysical evidence shows some support for the formation of "toroidal pores."

Lipid rearrangements in the membrane could lead to a hydrophilic or hydrophobic pore lining, as shown in Figure 12.9.3.

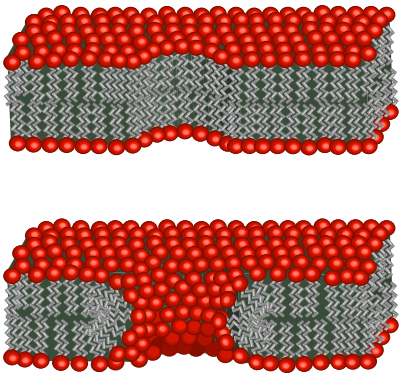


Figure 12.9.3: Diagram of possible lipid arrangements at the edge of a pore through a lipid bilayer. (above) In the absence of any rearrangement, the pore walls will be hydrophobic since the alkane tails are exposed. (below) Some researchers believe that the lipids at the edge re-orient to line the pore wall with hydrophilic head groups. Lipid bilayer. (2023, January 15). In *Wikipedia*. https://en.Wikipedia.org/wiki/Lipid_bilayer

We started our study of lipid bilayers with pure lipid systems and then added membrane proteins. Let's do the same with pore formation. Electroporation is a common technique to form a pore in pure lipid bilayers and cells. This technique is used to move a DNA with a target gene into either a prokaryotic cell (transformation) or eukaryotic cell (transfection) for exogenous gene expression. Without PFPs, this requires altering surface tension by applying an electrical potential. This forms depressions in the membrane, altering the nonpolar acyl chain packing. Small wire-like water columns appear in the process, which, like hydrophobic pores, ultimately rearrange into hydrophilic toroidal pores. Figure **12.9.4** shows snapshots of molecular dynamics simulations as a function of the time of pore formation in electroporation.

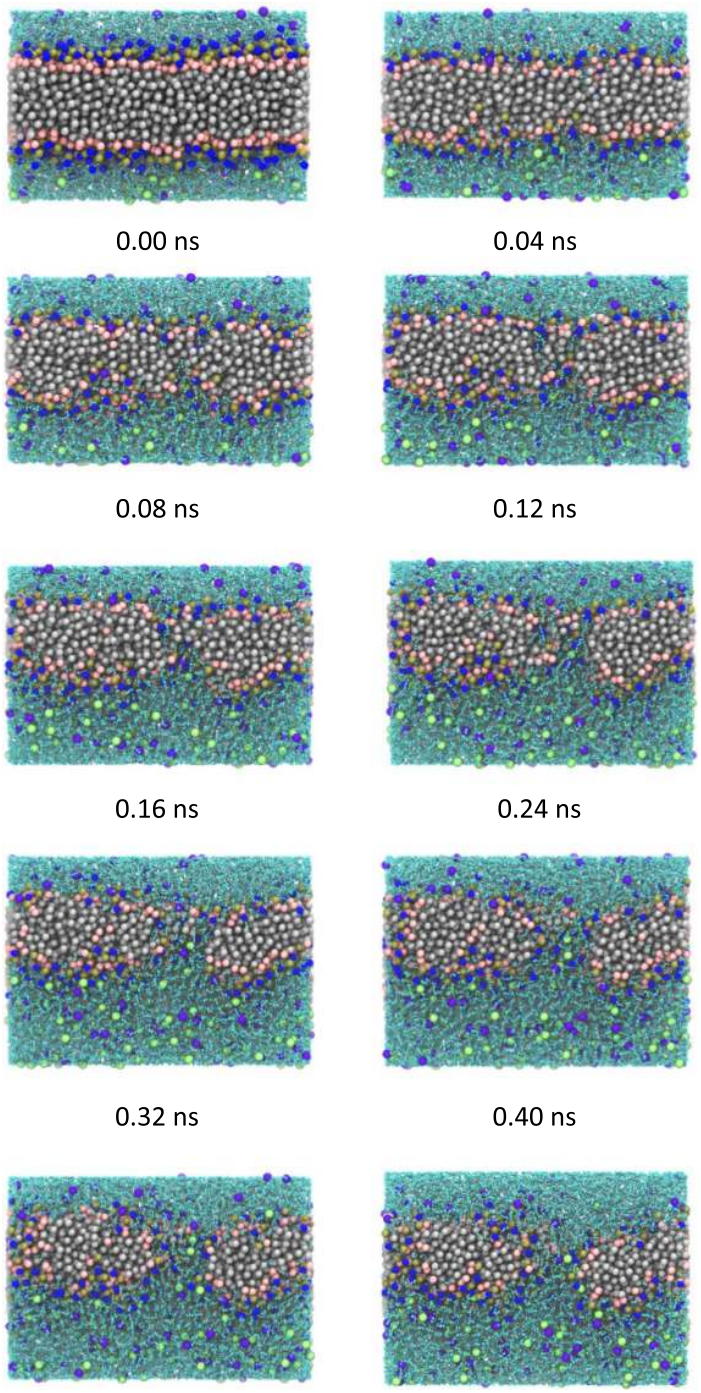


Figure **12.9.4**: Zhou, C., Liu, K. Molecular dynamics simulation of reversible electroporation with Martini force field. BioMed Eng OnLine 18, 123 (2019). https://doi.org/10.1186/s12938-019-0743-1. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

How does DNA pass through the pores in the bilayer? In pure lipid vesicles, it appears to pass through by electrophoresis. Most students are familiar with the movement of DNA fragments through pores during electrophoresis in agarose gels. In actual living cells, small nucleic acids like small interfering RNA (siRNA) and antisense DNA molecules appear to pass through the bilayer by electrophoresis. Large DNAs, like plasmids containing a gene for expression, bind to the cell and form cell surface aggregates, which appear to be endocytosed into the cell. Electroosmosis, the movement of liquids under the influence of an electric field, also plays a role.

12.9.2: Pores - Outer Membrane Factor (OMF) and Voltage-Dependent Anion Channel (VDAC)

Now, let's consider pores made of PFPs. We have already discussed two types of beta-barrel transmembrane proteins: the outer membrane factor (OMF) of Gram-negative bacteria and the voltage-dependent anion channel (VDAC). Both are examples of proteins called **porins** with typical beta-barrel topology.

VDAC: At low membrane potentials, VDAC (also known as mitochondrial porin), the most abundant protein in the mitochondrial outer membrane, moves metabolites and Ca^{2+} ions across the outer membrane of mitochondria. VDAC exist in an open state at zero or very low transmembrane potentials that allow for the transfer of key metabolic anions (pyruvate, oxaloacetate, malate, succinate, ATP, ADP, and P_i, which are involved in metabolism) and Ca^{2+}, and in a closed state (above or below $\pm$ 30 mV), which is not completely closed as it allows for the transfer of ions with a preference for cations. The closed state does not allow for the transfer of ATP. The transition to the closed state is promoted by tubulin and actin (cytoskeletal proteins), negatively charged lipids such as phosphatidyl ethanolamine and cardiolipin, and covalent phosphorylation by protein kinases. Bcl2, proteins that regulate programmed cell death (apoptosis), also interact with VDAC. In contrast to ligand-gated channels, which require ligand binding to open the channel, pore complexes are unusually open. Millions of ATPs/second move across the membrane through the open pore but none through the closed pore.

Figure 12.9.5 shows an interactive iCn3D model of mouse VDAC1 (4c69) with bound ATP loosely held in the site. An alpha helix partly occludes the central pore of this β-barrel protein.

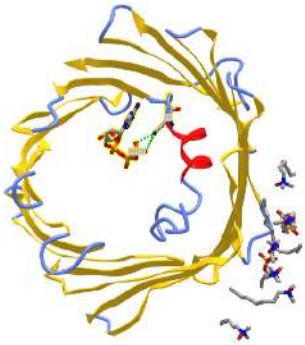


Figure 12.9.5: Mouse voltage-dependent anion channel - VDAC1 (4c69) with bound ATP loosely held in the site. (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?Ueb4wg26f5oEVL4Y9

One ATP is bound in the barrel and interacts with Lys 12 and 20 at each end of the cavity-bound helix. The alpha-helix narrows the pore opening and presumably changes orientation in a voltage-sensitive fashion to gate the pore open and closed, hence regulating the conductance of ions through the pore. That the orientation of the charged arginine side chains would depend on the transmembrane potential should be somewhat obvious.

Aquaporins

Aquaporins can move a billion water molecules per second across membranes and exclude ions, including protons. Waters proceed in a single file through the pore. Instead of moving waters through, it might simply move "H^+ ions" through by aligning subsequent H bond donors and acceptors in a "wire" of water molecules. This is prevented by two conserved asparagine residues in the center of the channel, which disrupt the water-to-water hydrogen bond network in the channel waters and could facilitate proton transfer. Instead, the central waters form hydrogen bonds to the central asparagines. This, along with local membrane potentials, causes opposite water orientations in different leaflet sides of the membrane, precluding H^+ transfer.

Here is a movie of a molecular dynamics simulation of water moving through a porin called aquaporin, GlpF, a protein in E. Coli.

Science magazine (**Tajkhorshid et al., Science Apr 2002, 296:525**). Used with permission from the **Theoretical and Computational Biophysics Group**, the **National Institutes of Health** (NIH) Resource for Macromolecular Modeling and Bioinformatics, at the Beckman Institute, University of Illinois at Urbana-Champaign.

OMF: The outer membrane factor (see previous section) is one member of a class of bacterial porins, the most abundant proteins in the outer membrane of Gram-negative bacteria. They are classified as non-specific or specific (with respect to the solute that passes through) or monomeric, dimeric, or trimeric based on their structure. In Gram-negative bacteria with two lipid bilayers, solute movement from inside to outside includes at least three sets of proteins. Active transport (discussed in the next section) needs an energy source. It is used by inner membrane transport proteins, including ATP-binding cassette (ABC)-type, resistance nodulation division (RND)-type, and major facilitator superfamily (MFS)-type transporters. These are connected to membrane fusion proteins (MFP) that span the periplasm, which then interact with at least 21 types of porins. Molecules with molecular weights greater than 600 generally can not get through the nuclear envelope of Gram-negative bacteria, limiting the size of potential antibiotics, which must enter by passive diffusion.

12.9.3: Mechanosensitive ion channels - Mscs (which are pores!)

As the name applies, these ion channels (with openings large enough to be called pores) are gated open/closed by mechanical (physical) changes in the properties of the membrane, not extracellular/intracellular ligands or voltage changes. Certain bilayer lipids also activate the Mscs. There are two types, small (MscS) and large (MscL) mechanosensitive ion channels. Changes in local (boundary layer) and nonlocal lipids are involved in the channel's gating (in the next section, we will discuss lipid-gated ion channels). They are found in prokaryotes, archaea, and eukaryotes. They are also called stretch-gated ion channels.

Mcss transduces a physical force (stretching and change in turgor pressure) into an electric signal - a flow of ions across the membrane. Turgor pressure is the internal pressure that "presses" the cytoplasm and cell membrane towards the cell wall in bacteria and plant cells. It arises mostly from the osmotic flow of water into the cell. If bacterial cells are placed in a high salt concentration solution (hypertonic), water flows out of the cell, and the cell membrane shrinks to the inside of the volume confined by the cell wall. When placed in a hypotonic solution, water flows into the cell, and the cell membrane presses to the cell wall. The response of these channels is fast, in the millisecond range, which is about as quick as a cellular response can be. Some variants of these channels are called **piezochannels**, based on the piezoelectric effect that describes how a voltage is produced when some materials are deformed by mechanical stress, which causes a redistribution of charges.

Bacteria normally have high concentrations of both K^+ and negatively charged anions, especially glutamate, which leads to a high turgor pressure from the inward osmotic flow of water. At low external osmolarity, turgor pressure in the cell could be as high as 4 atm. If placed in external solutions of high osmolarity, bacterial cells respond by increasing the movement of solutes into the cell.

Mscs are particularly important when bacteria are subjected to sudden osmotic shock. If they are placed in pure water, for example, water would flow down a concentration gradient into the cell and cause it to swell and lyse, killing it. This is done in the lab to prepare almost pure hemoglobin from ruptured red blood cells. The Mscs are opened under these conditions, and small species

from the cytoplasm flow out, helping to keep the cell viable. Their openings must be regulated to prevent too much outward flow, which would kill the cell.

In other organisms, they are also involved in touch (stretch), hearing (vibration, sound waves), and responses to gravity. Stimuli to activate them include fluid shear stress (relevant to endothelial cells that line blood vessels), membrane stretch (relevant to skeletal and cardiac muscle cells), or even indentation of a bilayer with a pipette. Changes in transmembrane turgor or other mechanical pressures cause membrane tension. Yet even in the absence of these changes, Mscs can be activated by anesthetics, phospholipids missing one fatty acid (called lysophospholipids), and certain polyunsaturated lipids. These stimuli also perturb the membrane bilayers.

Given the pore size, these proteins are less selective to ion flow than voltage-gate channels (see next section). Depending on the amino acids that line the pore, some Mscs would allow the preferential flow of anions, while some allow cation flow.

Some somatosensory channels (i.e. not pore) proteins also respond to pressure. When open, these can be selective to specific ions like Na^+ or K^+. Examples include some Transient Receptor Potential (TRP) ion channel variants. Other membrane proteins can also be activated by physical force. But true Mscs have some key characteristics. If mutated or deleted, the mechanosensory response is removed. If added to a cell, a mechanosensory response results.

a. Small-conductance mechanosensitive channel

This protein is a homoheptamer with three helices from each monomer involved in the overall structure. Two helices (1 and 2) interact more with the lipid components of the bilayer. Interactions of specific lipids with the helices seem to promote closing, but changes under high pressure lead to the opening of the pore. Half of each helix three forms the pore, while the other half is more parallel to the membrane and interacts with a large cytoplasmic domain.

Figure 12.9.6 shows the differences between the closed form of E. Coli MscS (2oau) and the open form (2vv5) viewing down the pore axis. The heptameric protein is shown in gray. Two key valines (105 and 109) on each chain are shown in spacefill and colored cyan. These hydrophobic amino acids act like gatekeepers, helping to keep water out, acting like a "vapor seal." In the closed state, the pore is sealed by closing the leucine "rings" as one half of helix 3 packs more closely. Many members of the MscS family vary significantly in size and can have between 3-11 transmembrane regions. The closed pore has a diameter of about 4.8 Å, while the open pore is 13 Å across.

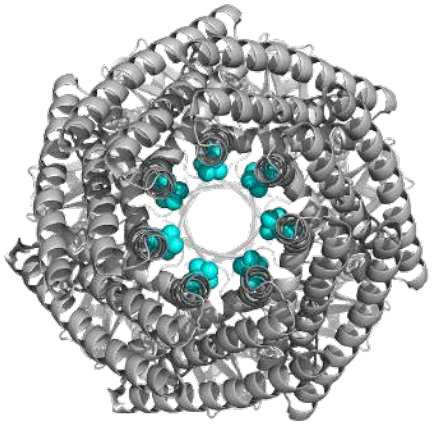


Figure 12.9.6: Differences between the closed form of E. Coli MscS (2oau) and the open form (2vv5) viewing down the pore axis.

Most of you would have studied Ohm's law, given by

$$\mathrm{V} = \mathrm{IR} \text{ or } \mathrm{I} = \frac{\mathrm{V}}{\mathrm{R}} \tag{12.9.1}$$

where I (amps) is the current, R (ohms) is the resistance and V (volts) is the voltage. A more general variant of this law used in physics is

$$\mathrm{J} = \sigma \mathrm{E} \tag{12.9.2}$$

where J is the current density, E is the electric field, and σ is the conductivity (inverse of resistance), which depends on the material. The unit of sigma σ is $ohms^{-1}$ or mhos (ohm written backward). That unit has been renamed the Siemens (S). Mscs channels have a small conductance of approximately 1 nS in 400 mM salt solution.

b. Large conductance mechanosensitive ion channel (MscL):

The channels have large conductances (3 nS) and concomitantly larger pore sizes, allowing the flow of water, ions, and even small proteins. Again, they are involved in diffusion down an electrochemical gradient through pores, so they are not involved in active transporter but rather gated diffusion.

Figure 12.9.7 shows an interactive iCn3D model of the pentameric MscL from Mycobacterium tuberculosis (2oar), a Gram-positive bacterium that causes tuberculosis. About 23% of the world's population is affected by this pathogen. It causes about 1.5 million deaths each year. Compare this to 7 million confirmed deaths during four years of the COVID pandemic (with central estimates of 27 million excess deaths). It has killed over 1 billion people throughout human history (but not as many as malaria). The pore diameter is about 30 Å across.

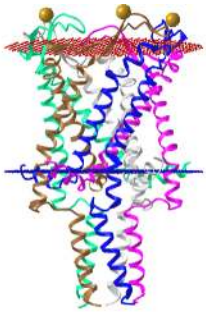


Figure 12.9.7: Closed state of the pentameric MscL from Mycobacterium tuberculosis (2oar). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LsRZZbmFy5NRF6

12.9.4: Pore-forming alpha-helical toxins

We started this chapter by discussing the innate immune system's major attack complex (MAC). Pathogens also employ pore formation to kill host cells. Many secrete soluble proteins that aggregate in the membrane to form alpha-helical or beta-barrel pores. The proteins are called pore-forming toxins (PFTs). Killing occurs when cytoplasmic components leak out or bacterial toxins, such as diphtheria and anthrax toxins, move into the cell.

Figure 12.9.8 shows an interactive iCn3D model of the pore formed by cytolysin A (ClyA, also known as HlyE), an alpha-PFT used by some E. Coli and Salmonella enterica strains. The pore is a large dodecamer that forms from soluble monomeric ClyA (2wcd). It has a pore diameter of about 40 Å,

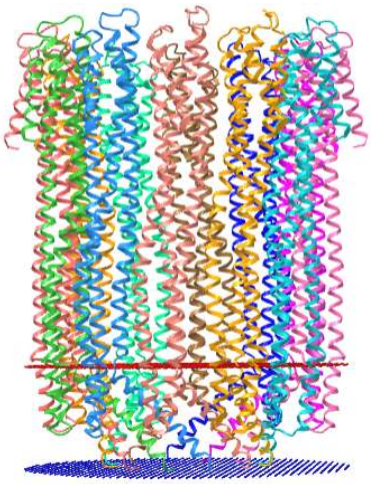


Figure 12.9.8: Pore formed by cytolysin A (ClyA, also known as HlyE). (2wcd) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...RXUsE24d9d8NT6

Figure 12.9.9 shows the soluble monomeric form of cytolysin A (ClyA, HlyE) (1QOY). Nonpolar side chains are shown in cyan. The transparent surface is mostly polar, making the monomer soluble.

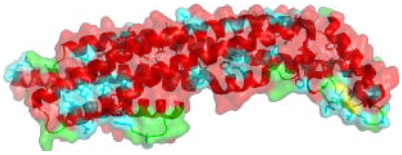


Figure 12.9.9: Soluble monomeric form of cytolysin A (ClyA, HlyE) (1QOY)

Figure 12.9.10 shows the changes in conformation between the single-chain soluble form (shown in the figure above) and a single chain of the membrane oligomeric channel.

Figure 12.9.10: Changes in conformation between the single chain soluble form and the same chain in the channel

We started this chapter section by exploring electroporation and the formation of a toroidal-like hole in the lipid bilayer. In the case of ClyA, the protein aggregate itself forms the pore, not the lipids themselves, although lipid rearrangements are necessary to form the protein complex.

12.9.5: Gap Junctions

Connexins are voltage-gated channels that allow ions, metabolites, nucleotides, and small peptides to flow between cells. A connexin has four transmembrane helices and two extracellular loops. Six protomers combine in a single cell to form a channel complex called a **connexon** or hemichannel. Beta structures in the connexin hemichannel of one cell dock with a similar channel on an adjoining cell to form a full channel passing through the membranes of both cells, forming a gap junction between the cells. This is shown in Figure 12.9.11. There are 20 different connexins encoded in the human genome.

The left figure below shows the six protomers, each in a different color, and a gray rectangle representing the bilayer of a single connexon or hemichannel. The right side of the figure shows a full gap junction channel between two cells, with the membranes represented by gray rectangles. The connexin 26 monomer was used to create the diagram. Mutations in this protein are associated with hearing loss.

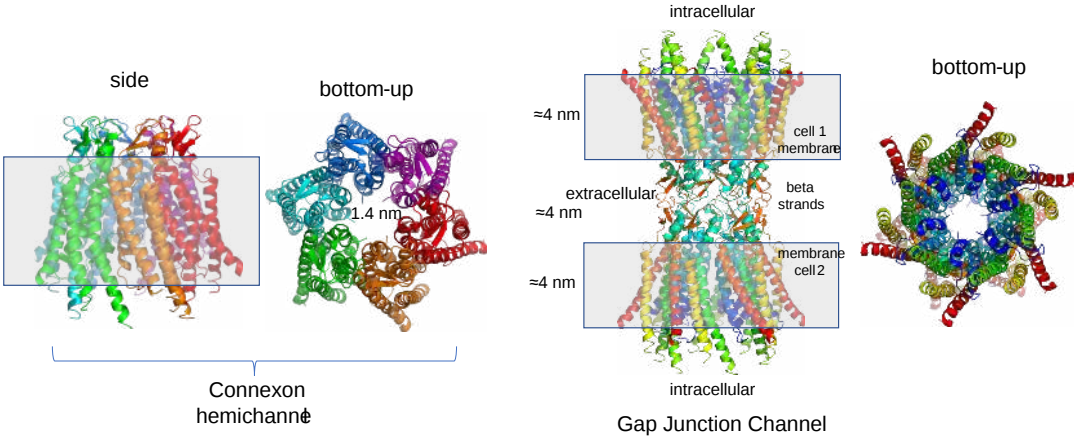


Figure 12.9.11: Connexon channel and the gap junction

The left figure above shows the six protomers, each in a different color, and a gray rectangle representing the bilayer of a single connexon or hemichannel. The right side of the figure shows a full gap junction channel between two cells, with the membranes represented by gray rectangles. The connexin 26 monomer was used to create the diagram. Mutations in this protein are associated with hearing loss.

The cytoplasmic entrance of the channel is positively charged. It forms a funnel of six amino-terminal helices, leading to a negatively charged transmembrane lining. The entrance diameter is 14 Å.

Figure 12.9.12 shows an interactive iCn3D model of a full gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). Only the bottom membrane bilayer is represented by red and blue dummy atoms.

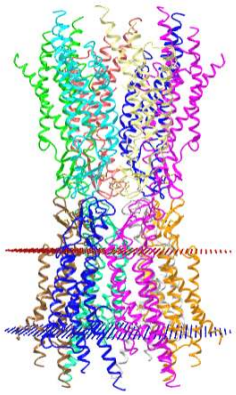


Figure 12.9.12: Gap junction channel connecting two membranes using the human connexin 26 monomer (2zw3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?TpuQ6Agk8XWuApsWA

12.9.6: The Nuclear Pore Complex (NPC)

Channels have pores that can be gated open and allow the selective flow of ions. Pore-forming proteins have larger entrances that allow small and large molecules to pass through the bilayer. The pore opening in even large mechanically sensitive channels (about pale in size compared to the nuclear pore complex, which has a combined molecular mass of around 125,000,000! Its outer diameter is ≈1,200 Å, and its inner one is ≈ 425-Å. Figure 12.9.13 shows the relative size of nuclear pore compared to other molecular structures, including the eukaryotic ribosome, nucleosome, a soluble tetrameric protein (rubisco, 270K), and MscL (shown as a circle, which represents the pore diameter).

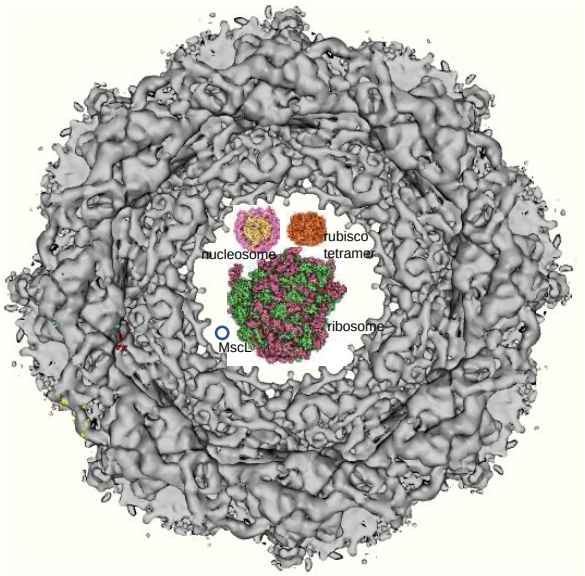


Figure 12.9.13: Relative size of nuclear pore compared to other molecular structures including the eukaryotic ribosome, nucleosome, a soluble tetrameric protein (rubisco, 70K) , and MscL (shown as a circle which represents the pore diameter)

Its job is to shuttle small molecules by passive diffusion down a concentration gradient through the pore. In addition, it moves large molecules and molecular structures (proteins, RNA, and perhaps ribosomes) across the nuclear membrane in a process that requires energy. The proteins that comprise this complex are called **nucleoporins (nups)**, which appear to number around 34 in humans. Each NPC complex contains around 1000 **nucleoporins**. The complex fuses the inner and outer nuclear membranes.

We have focused so much on single bilayer membranes that comprise the plasma membrane and membranes of organelles like the Golgi complex and lysosomes it might come as a surprise (perhaps not to biology students) that the nuclear membrane or envelope appears to consist of two bilayers. Most know that the mitochondria have two bilayers, an inner and outer membrane, similar to Gram-negative bacteria. Mitochondria are believed to have arisen from bacteria, so the double bilayer there makes sense. Figure 12.9.14 shows the nuclear membrane or envelope of two bilayers (1) with an outer ring (2), spokes (3), a basket (4), and filaments (5). The NPC spans both bilayers.

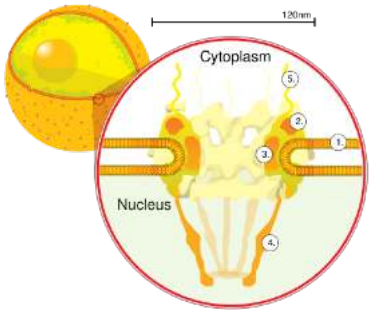


Figure \(\PageIndex{14\): Nuclear membrane or envelope of two bilayers (1) with an outer ring (2), spokes (3), a basket (4) , and filaments (5) https://en.Wikipedia.org/wiki/File:N...e_crop.svg.png; Creative Commons Attribution-Share Alike 2.5 Generic

The outer bilayer of the nuclear envelope is continuous with the endoplasmic reticulum, as shown in Figure \(\PageIndex{15\) below. The dots on the ER membrane are ribosomes, making this the rough ER (as opposed to smooth ER, which has no attached ribosomes).

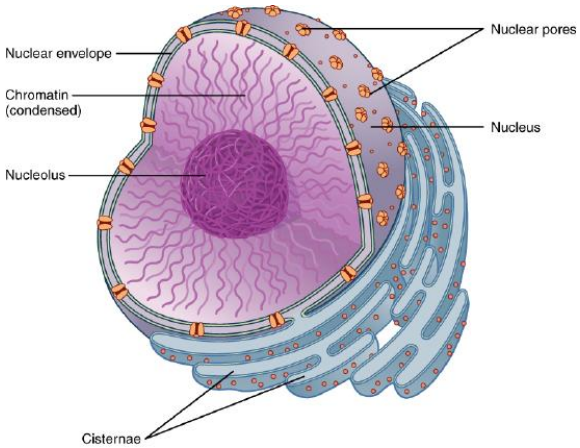


Figure \(\PageIndex{15\): The outer bilayer of the nuclear envelope is continuous with the endoplasmic reticulum https://commons.wikimedia.org/wiki/F...18_Nucleus.jpg; OpenStax, CC BY 4.0 <https://creativecommons.org/licenses/by/4.0>, via Wikimedia Commons

Figure \(\PageIndex{16\) below shows a model of the basket-like structure of the nuclear pore complex (NPC). It shows that instead of two separate bilayers, there is just one bilayer, with each leaflet bending around at the NPC and reversing directions! Think of the interesting lipids and protein components that enable the bend! Alternatively, you could say that two different membranes fuse at the NPC.

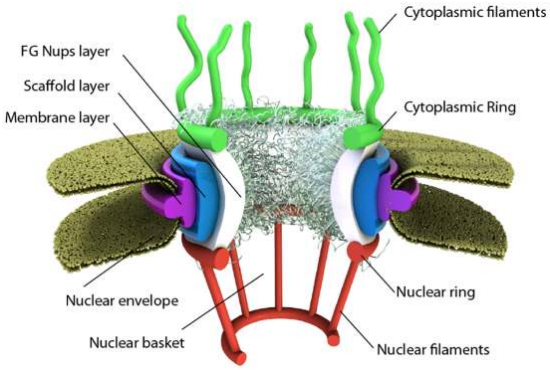


Figure \(\PageIndex{16\): A model of the basket-like structure of the nuclear pore complex **Citation:** Azimi M, Mofrad MRK (2013) Higher Nucleoporin-Importinβ Affinity at the Nuclear Basket Increases Nucleocytoplasmic Import. PLoS ONE 8(11): e81741. https://doi.org/10.1371/journal.pone.0081741. Creative Commons Attribution License,

The NPC consists of 32 copies of each specific nucleoporin (Nup) except two. One has 48 copies and the other 16 (even these sum to 2x32 Nups). Three rings form and surround the pore. A 16-membered ring of Nups faces the cytoplasm (**cytoplasmic ring**), and another 16-membered ring of Nups faces the nucleoplasm (**nuclear ring**). There is 8-fold rotational symmetry in each ring, suggesting a dimeric repeat of Nups in the rings. Eight Nups in the cytoplasmic ring have a disordered end that sticks out into the cytoplasm as filaments. In contrast, the disordered ends of eight Nups in the nuclear ring form filaments that bind together to form a ring at the bottom of the nuclear basket.

The **inner ring** (the **FG Nups layer** in the figure above), between the cytoplasmic and nucleoplasmic rings, also consists of Nups with ordered domains and disordered parts. The disordered parts of the inner ring Nups have repetitive sequences enriched in phenylalanine and glycine (hence the name FG Nups) and stick out into the central pore. These disordered regions act as a filter, allowing certain molecules to pass and excluding those with molecular weights greater than 40K. Large molecules need transport proteins called karyopherin transport factors to move through the pore.

The core structure of the NPC, obtained through cryoelectron microscopy, is shown in Figure \(\PageIndex{17\) below.

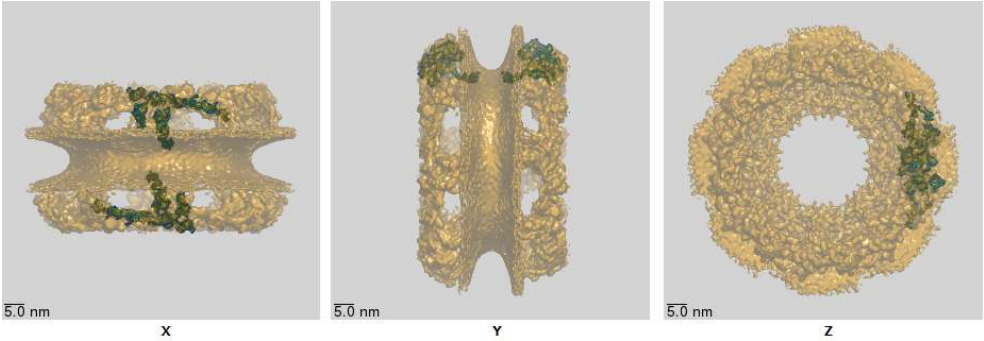


Figure \(\PageIndex{17\): The core structure of the Nuclear Pore Complex https://www.ebi.ac.uk/pdbe/entry/emd...-3103/analysis

The double bilayers are very evident. The nuclear and cytoplasmic filaments and the disordered FG repetitive sequences that stick into the inner ring pore are not seen since they are very flexible and don't adopt single conformations using standard structure determination methods.

Figure 12.9.18 shows an interactive iCn3D model of the nuclear pore complex (NPC) core (5a9q), whose structure was obtained using cryoelectron microscopy. (load slowly)

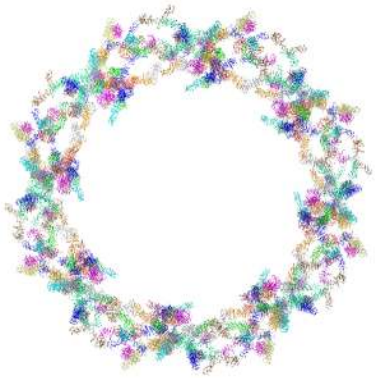


Figure 12.9.18: Nuclear pore complex (NPC) core (5a9q). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?NP7oQb235BCgPARA8

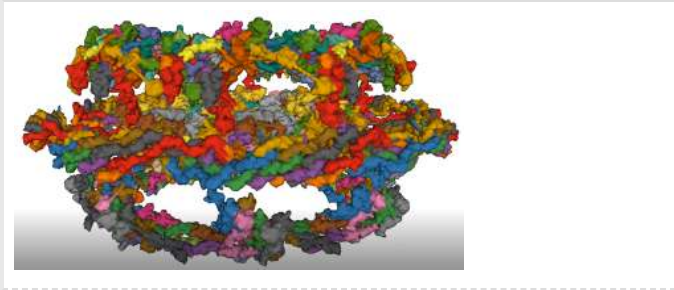
A cartoon figure showing the variety of Nups in the NPC is shown in Figure 12.9.19.

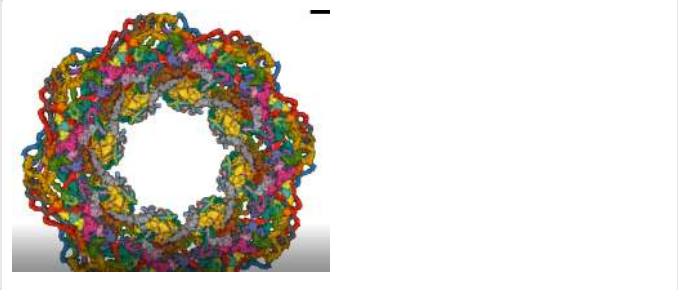
Figure 12.9.19: A cartoon version of the various NUPS in the nuclear pore complex. Donnaloja F., Jacchetti E., Soncini M., Raimondi M. T.Front. Physiol., 12 July 2019 | https://doi.org/10.3389/fphys.2019.00896**Creative Commons Attribution License (CC BY)**.

In this figure, the cytoplasmic and nucleoplasmic rings are shown in green, each formed mainly by 16 copies of the Y-complex arranged in two eight-membered rings. The inner ring, predominantly formed by 32 copies of the Nup93 complex, is shown in red. Transmembrane nucleoporins are depicted in violet, and the cytoplasmic filaments and nuclear basket structure are in orange. Attached to the inner ring are Nup62 complexes (depicted in blue), which form a cohesive meshwork within the central channel through their FG-repeat domains. The position of Nup98 is not indicated, a FG-repeat-containing nucleoporin important for the transport and exclusion function of NPCs; its position in the NPC is less defined, but it might be part of the inner ring. Similarly, Aladin (also known as AAAS), Gle1, Rae1, and Npl1 (also known as hCG1 and NUPL2) have been omitted.

Large proteins and RNA that pass through the pore must first be bound to a cargo receptor, which can move the "cargo" across the pore with concomitant GTP hydrolysis. This process is closer to active transport, which we will discuss in Chapter 11.3.

The entire nuclear pore complex was solved in 2022 using cryoEM. Here are two videos showing the dilated complex (7R5J). Click on the images to download mp4 animations of the complex.





12.9.7: Summary

This chapter explores the diverse mechanisms by which pores form in lipid bilayers and the specialized proteins that facilitate or regulate this process. At its core, pore formation enables the movement of molecules across membranes, driven not only by concentration gradients but also by electrical potentials, making it a unique example of facilitated diffusion.

Key Concepts and Mechanisms

- **Membrane Permeability and Facilitated Diffusion:**
  Pore formation creates channels that allow the passage of molecules and ions. Unlike simple diffusion, this process can be influenced by both the concentration and the electrical potential across the membrane.
- **Electroporation and Lipid Rearrangement:**
  Electroporation demonstrates how applying an electrical potential can transiently disrupt lipid packing, leading to the formation of toroidal pores. These pores initially present as hydrophobic structures but can rearrange into hydrophilic channels, mirroring the behavior seen in some protein-mediated pores.
- **Pore-Forming Proteins (PFPs):**
  Proteins such as bacterial toxins (e.g., Hemolysin E/ClyA), the membrane attack complex (MAC) in the innate immune system, and various porins (e.g., OMF and VDAC) form pores through different mechanisms. Some insert directly into the membrane, while others induce lipid rearrangements to form the channel.

Structural Diversity of Pore-Forming Systems

- **Aquaporins and Ion Channels:**
  Aquaporins showcase selective water transport, preventing proton leakage through specific structural features, while ion channels like VDAC adjust their open and closed states in response to changes in membrane potential. These channels are critical for maintaining metabolic flux and ion homeostasis.
- **Mechanosensitive Ion Channels (MscS and MscL):**
  These channels respond to physical forces such as membrane stretch and changes in turgor pressure. They play a vital role in protecting cells under osmotic stress by providing a safety valve mechanism that prevents cell lysis.
- **Gap Junctions:**
  Formed by connexins, gap junctions create direct intercellular channels that allow ions, metabolites, and small peptides to pass between adjacent cells. Their structure—a hexameric assembly called a connexon—highlights the importance of protein-protein interactions in cell communication.
- **The Nuclear Pore Complex (NPC):**
  The NPC represents a highly complex structure that regulates the bidirectional transport of macromolecules between the nucleus and cytoplasm. Composed of multiple nucleoporins with both structured and intrinsically disordered regions, the NPC functions as a dynamic gate, employing energy-dependent processes for the transport of large molecules.

Integrative Biophysical Principles

Throughout the chapter, classical biophysical concepts such as Ohm's law and the principles of electrophoresis and electroosmosis

are revisited to explain how ions and molecules traverse these pores. The discussion emphasizes that pore formation—whether through lipid rearrangement, protein insertion, or a combination of both—requires a delicate interplay between membrane composition, protein structure, and the physical forces at work in the cellular environment.

By integrating molecular dynamics simulations, structural models (via cryoEM and interactive tools), and biophysical theory, this chapter provides a comprehensive overview of how cells regulate transport across their membranes. This not only deepens our understanding of cellular physiology but also sheds light on the mechanisms by which pathogens and immune cells use pore formation as a strategy to influence cell viability.

This page titled 12.9: Diffusion Across a Membrane - Pores is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.4: Diffusion Across a Membrane - Pores** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

12.10: Active Transport

Learning Goals (ChatGPT o3-mini)

1. **Differentiate Transport Mechanisms**
   - Distinguish between passive transport (facilitated diffusion) and active transport, including the thermodynamic principles (chemical/electrochemical gradients versus energy-coupled uphill transport).
2. **Energy Sources and Their Utilization**
   - Explain how various energy sources (ATP hydrolysis, oxidation, light, and ion gradient collapse) drive active transport.
   - Analyze specific examples, such as the ATP-dependent Na^+/K^+ pump and light-driven proton pumps, to illustrate how energy input is coupled to transport.
3. **Classification of Active Transporters**
   - Classify transporters into primary (directly using ATP or photons) and secondary (using coupled solute gradients).
   - Differentiate between uniporters, symporters, and antiporters, with examples (e.g., GLUT1 for facilitated diffusion versus Na^+/K^+ ATPase as an active uniporter, SGLT1 as a symporter, and the sodium-calcium exchanger as an antiporter).
4. **Structure–Function Relationships in Membrane Transport**
   - Describe the structural features of major transporter families such as the Major Facilitator Superfamily (MFS) and how their two-domain, 12-transmembrane helix architecture underpins the “clamp-switch” mechanism.
   - Discuss how conformational changes facilitate the binding and release of substrates in different transporter states.
5. **Mechanistic Details of Specific Transporters**
   - Analyze the mechanism of proton-linked monocarboxylate transporters (MCTs) and lactose permease, emphasizing key residues involved in substrate binding and proton coupling.
   - Compare the structural and functional aspects of transporters like LeuT, illustrating how substrate and ion binding induce conformational shifts.
6. **Thermodynamics of Active Transport**
   - Interpret the energy landscapes and free energy changes associated with active transport cycles, including how ATP hydrolysis is coupled to transport and the concept of energy dissipation.
   - Understand models such as the Post-Albers cycle for P-type ATPases, emphasizing the roles of E1 and E2 states.
7. **Membrane ATPases and Ion Pumps**
   - Explain the structure and kinetic mechanisms of P-type ATPases, using the Ca^{2+} pump and Na^+/K^+ ATPase as models to illustrate ion transport against steep concentration gradients.
   - Discuss the regulatory mechanisms and inhibition (e.g., vanadate inhibition) of these ATPases.
8. **Understanding ABC Transporters**
   - Outline the structural domains of ABC transporters (transmembrane domains and nucleotide-binding domains) and discuss their roles in exporting toxins and importing essential nutrients.
   - Evaluate the clinical relevance of proteins like P-glycoprotein in multidrug resistance.
9. **Integration of Structural and Kinetic Concepts**
   - Relate the detailed structural information (e.g., transmembrane helices, key amino acid residues) to the kinetic and thermodynamic aspects of transporter function.
   - Interpret data from structural models and molecular dynamics simulations to predict how conformational changes drive substrate movement.
10. **Nuclear Pore Complex and Selective Transport**

- Describe the unique transport mechanism of the nuclear pore complex, emphasizing the role of FG-nucleoporins and the regulation by Ran GTPase.
- Compare and contrast nuclear transport with classical active transport mechanisms, focusing on the specificity provided by nuclear localization (NLS) and export signals (NES).

Each of these learning goals is designed to guide students in integrating structural, kinetic, and thermodynamic principles with real-world examples of membrane transport mechanisms in biological systems.

In the previous sections, we explored facilitated diffusion, diffusion through channels, and diffusion through larger pores. In each case, once a carrier/permease protein was available, or a channel (gated by ligand binding, change in membrane potential, lipid binding, or mechanical forces) or a pore formed, solute flows down a chemical gradient (facilitated diffusion) or electrochemical gradient in a thermodynamically favored process. But what about moving solutes from low to high concentration against a concentration gradient, which would be necessary to capture the last bit of a vital nutrient or energy source? Active transport does just that, but it requires an energy source. Many different types of chemical species are actively transported across the cell and organelle membranes, including sugars, amino acids, (deoxy)nucleotides, metabolites (like carboxylic acids),

12.10.1: Type of Active Transport

A membrane receptor is required for active transport to occur, which recognizes the ligand to be transported. However, the energy source used to drive transport against a concentration gradient is of major interest to us. The biological world has adapted to use almost any source of energy available.

ATP hydrolysis: One would expect this ubiquitous carrier of free energy to drive active transport. This is one of the predominant roles of ATP in the biological world. 70% of all ATP turnover in the brain is used to create and maintain a Na and K ion gradient across nerve cell membranes using the membrane protein Na^+/K^+ ATPase.

Energy released by oxidation: You may have encountered this in previous biology courses. Active transport of protons driven by oxidative processes is exergonic. In electron transport in respiring mitochondria, NADH is oxidized as it passes electrons to a series of mobile electron carriers (ubiquinone, cytochrome C, and eventually dioxygen) using protein complexes in the inner membrane of the mitochondria. The energy lost in this thermodynamically favored process is coupled to conformational changes in the complex, which cause protons to be ejected from the matrix into the inner membrane space. One can imagine a series of conformation-sensitive pKa changes in various side chains in the complexes, which lead to the vectorial discharge of protons.

Light: Photosynthetic bacteria have a membrane protein called **bacteriorhodopsin,** which contains **retinal**, a conjugated polyene derived from beta-carotene. It is analogous to the visual pigment protein rhodopsin in retinal cells. Light absorption by retinal induces conformation changes in the retinal and bacteriorhodopsin, which leads to the vectorial discharge of protons.

The collapse of an ion gradient: The favorable collapse of an ion gradient can be used to drive the transport of a different species against a concentration gradient. We have already observed that the collapse of a proton gradient across the inner mitochondrial membrane (through F_0F_1ATPase) can drive the thermodynamically unfavored synthesis of ATP. The collapse of a proton gradient provides a proton-motive force that can drive the active transport of sugars. Likewise, a sodium-motive force can drive the active transport of metal ions. Since the energy to make the initial ion gradients usually comes from ATP hydrolysis, ATP indirectly powers the transport of the other species against a gradient.

Active transport can be divided into two classes:

- primary - driven by "primary" energy sources, which for us means ATP or photons
- secondary - driven by coupled transport (also called cotransport) of solutes down a concentration gradient, which provides the thermodynamic force to move a solute "uphill."

Active transporters can also be divided into classes based on the direction of movement of the solute and any cotransported solute into uniport (no cotransported solute), symport (solute and cosolute transported in the same direction), and antiport (solute and cosolute transported in the opposite direction) as shown in Figure 12.10.1.

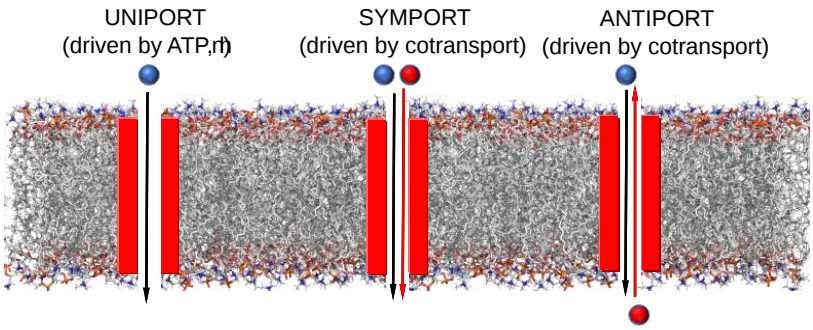


Figure 12.10.1: Uniport (no cotransported solute), symport (solute and cosolute transported in the same direction), and antiport (solute and cosolute transported in the opposite direction) transporters

Examples of these different mechanisms can be found in facilitated diffusion and active transport. Examples of these types include:

- uniport: An often-used example is GLUT 1, which catalyzes glucose's FACILITATED DIFFUSION (not active transport) down a concentration gradient. A somewhat complicated example of a uniport active transporter is the Na^+/K^+ ATPase, which moves Na^+ ions out of a cell and K^+ into a cell against concentration gradients. The movement of one ion does not provide the chemical potential thermodynamic driving force to move the other. Rather, the hydrolysis of ATP is required. So, this protein can be called a uniport active transporter for two different ions.
- symport: Another glucose transporter, the glucose symporter (SGLT1), found in the small intestines, heart, and brain, cotransports one glucose or galactose for every two Na^+ ions that move into the cells down a concentration gradient.
- antiport: The sodium-calcium exchanger pumps out one Ca^{2+} ion (low to high concentration) driven by the influx of three Na^+ into the cell. This keeps Ca^{2+} low in the cytoplasm.

If the species moved is charged, two other terms are used:

- electrogenic - a net electrical imbalance is generated across the membrane by symport or antiport of charged species
- electroneutral - no net electrical imbalance is generated across the membrane by symport or antiport of charged species

In this chapter, we will explore examples of several active transporters. All are polytopic with alpha-helical transmembrane domains, which, through a series of conformational transitions, can move chemical species that are bound to the receptor or channel protein through the membrane by creating transient openings in the protein. We will end with another look at the nuclear pore complex and see how it allows the movement of large proteins and RNA molecules through its pore through a very different type of mechanism. In that case, it is not active transport since the relative concentration of the transported species on either side of the pore is less important than the mechanism by which certain species are targeted to move into and out of the nucleus.

12.10.2: Major Facilitator Superfamily (MFS) Transporters

There are about 74 different families of active transporters in the Major Facilitator Superfamily (MFS), which is the largest of secondary transporters using symport or antiport mechanisms. There are also uniporters within this family. Different members of the MFS move a diverse array of solutes, including sugars, nucleotides, peptides, and drugs, across membranes against a concentration gradient. All have two membrane domains consisting of six transmembrane helix bundles that pack to give an apparent two two-fold rotational axis of symmetry (also called pseudo C_2 axis). The ligand to be transported binds between the two domains in a central cavity, potentially open to either side of the membrane. Figure 12.10.2 panel A (top) below shows a cartoon of the 12 helices and 6-helix bundles. Helices 1 and 4 in domain one and the analogous ones in domains 7 and 10 are involved in ligand movement through the membrane. Correspondingly, helices 2, 5, 8, and 11 are involved in domain-domain interactions. The bottom figure (B) shows two views of the protein with the ligand in spacefill rendering.

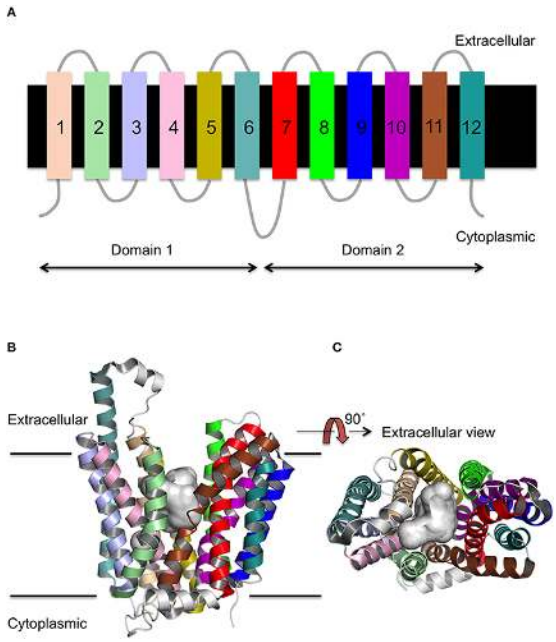


Figure 12.10.2: Structure of the Major Facilitator Superfamily (MFS). Images from Lee J, Sands ZA and Biggin PC (2016) A Numbering System for MFS Transporter Proteins. *Front. Mol. Biosci*. 3:21. doi: 10.3389/fmol.2016.00021 **Creative Commons Attribution License (CC BY)**

As structure mediates function, they all display a similar "clamp and switch" mechanism, in which solute binds to the protein in an open (to binding) conformation, leading to conformation changes forming a closed form, followed by further conformational changes that cause the complex to open to the other side of the membrane. Figure 12.10.3 shows hands opening, closing, and reopening in the other direction as an analogy for this "clamp-switch" mechanism.

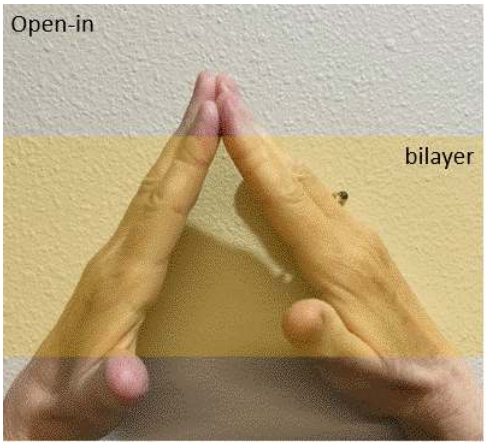


Figure 12.10.3: Hands opening, closing, and reopening in the other direction as an analogy for this "clamp-switch" mechanism of the Major Facilitator Superfamily (MFS)

We will describe two different transporters in this family.

Human monocarboxylate transporter (7BP3)

These proton-linked monocarboxylate transporters (MCTs) move many monocarboxylates (including lactate, pyruvate, and the ketone bodies acetoacetate, beta-hydroxybutyrate, and acetate) out of the cell, driven by the flow of protons across the membrane in the same direction (symport). In cells engaging in glycolysis, a fundamental energy utilization pathway, it catalyzes the high-affinity transport of pyruvate, the end product of glycolysis, and lactate, a reduced derivative of pyruvate, out of the cell. Their activity is effectively turned off in cells not engaged in glycolysis. The flux of pyruvate/lactate has a steep dependence on their concentrations, allowing the protein to act as an on-off switch for transport. A step dependency of activity on concentration is a hallmark of cooperative interactions. Structural analyses show conformational changes in the interface of the dimeric form of the transporter.

We will see when we study glycolysis that the conversion of the oxidized pyruvate to its reduced form, lactate, requires a reducing agent, in this case, NADH, which is oxidized back to NAD+, an oxidizing agent required for glycolysis to continue. When pyruvate is high, NADH is also high, allowing for the conversion of pyruvate to lactate, which can be exported out of the cell. This reaction describes the interconversion:

$$\text{Pyruvate} + \text{NADH} \leftrightarrow \text{Lactate} + \text{NAD}^+$$

Using MCT 2 to remove lactate from the cell pulls the above reaction to the right, regenerating NAD^+ and allowing glycolysis to continue.

Figure **12.10.4** shows an interactive iCn3D model of the inward open human monocarboxylate transporter 2 (7BP3).

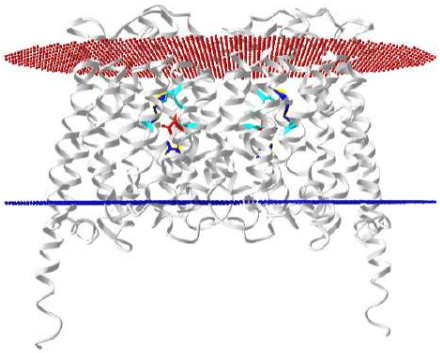


Figure \(\PageIndex{4\): Inward open human monocarboxylate transporter 2 (7BP3). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...yq5U8v4BvdDA57

Several key residues are shown. Aspartic acid 293 (side chains in red sticks) is conserved in all MCTs studied and is likely involved in H^+ transport. It is surrounded by Val 156, Met 289, Ala 290, and Phe 351 (side chains in cyan sticks), all nonpolar amino acids, which would elevate the pKa of Asp 293, allowing it to stay protonated until a conformational change would alter its environment, leading to proton release. Two positively charged residues, Lys 38 and Arg 297 (side chains in blue sticks), are involved in substrate interactions. Lys 38 faces outside the cell in the inward open conformation. Figure \(\PageIndex{5\) below shows the same amino acids in a zoomed view. The amino acids are labeled.

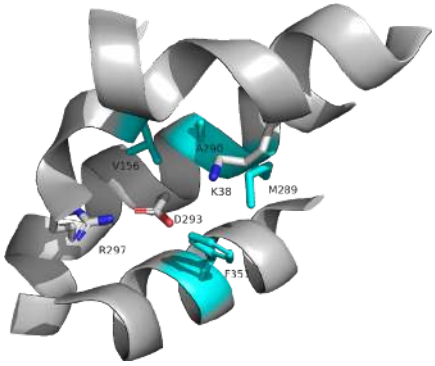


Figure \(\PageIndex{5\): Key amino acid side chains in inward open human monocarboxylate transporter 2 (7BP3)

Lactose Permease (Lactose-proton symport)

This E. Coli protein is also a symporter, which uses the movement of H^+ down a concentration gradient to drive lactose into E. Coli to capture as much lactose - an energy source - as possible. Deprotonation of a protonated Glu 269 leads to ligand binding, as the protein has no binding site in the protonated state. Lactose induces the formation of its binding site. Figure \(\PageIndex{6\) below shows the change in conformation going from a more acidic pH (5.6, PDB 2CFP) to a more neutral pH (6.5, PDB 2CFQ). One lactose is transported from one H^+, which moves down a concentration gradient. Lactose moves from the outside periplasm to the cytoplasm.

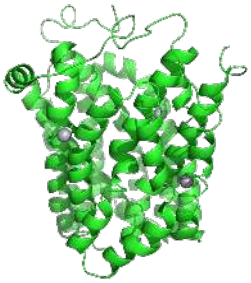


Figure \(\PageIndex{6\): Conformation changes in lactose permease from a more acidic pH (5.6, PDB 2CFP) to a more neutral pH (6.5, PDB 2CFQ)

Figure 12.10.7 shows an interactive iCn3D model of lac permease (1pv7), which highlights three residues essential for lactose binding (Glu 126, Arg 144, and Glu 269 in stick) and three involved in proton transfer (Arg 302, His 322, and Glu 325 in spacefill).

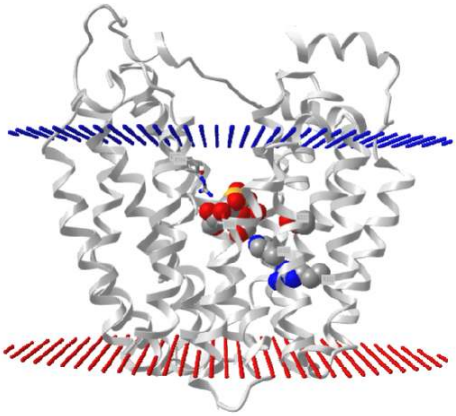


Figure 12.10.7: Lac permease (1pv7). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9mt2juToECQPg6

It has a bound nonhydrolyzable lactose analog (beta-D-galactopyranose-(1-1)-1-thio-beta-D-galactopyranose), shown in spacefill, in the center hydrophilic cavity. The two 6-helix bundles, which give pseudo 2-fold symmetry, are evident. The protein is in the open-to-inward conformation. (The blue dummy atoms represent the inner leaflet.)

Figure 12.10.8 is a video of a molecular dynamics simulation showing lactose moving through lactose permease.

Figure 12.10.8: Molecular dynamics simulation showing lactose moving through lactose permease. https://www.ks.uiuc.edu/Gallery/Movi...annelProteins/

As with other members of the MFS transporters, lactose permease has 12 transmembrane helices that form two bundles. The lactose binding site is between the two bundles. When the protein is in the open-to-out conformation, it can bind protons, likely at an exposed His. This enables lactose binding, which is followed by a conformational change to the open-to-in form. When the protonated His loses a proton, lactose dissociates into the cytoplasm. Movement of both H^+ and lactose together must occur for transport.

Structural work by Singh et al. on the Leu Transporter (LeuT), a member of the solute carrier 6 or sodium-coupled transporters, which is an active transporter requiring movement of Na^+ ions into the cells to power the uptake of Leu, shows an "open-to-out" and occluded binding state for ligand (Leu). Tryptophan, a competitive nontransportable inhibitor, binds to the open-to-out state but is too large for the obligate occluded state, so it is not transported. Figure 12.10.9 shows inward open, closed, and outward conformations of different major facilitator superfamily transporters.

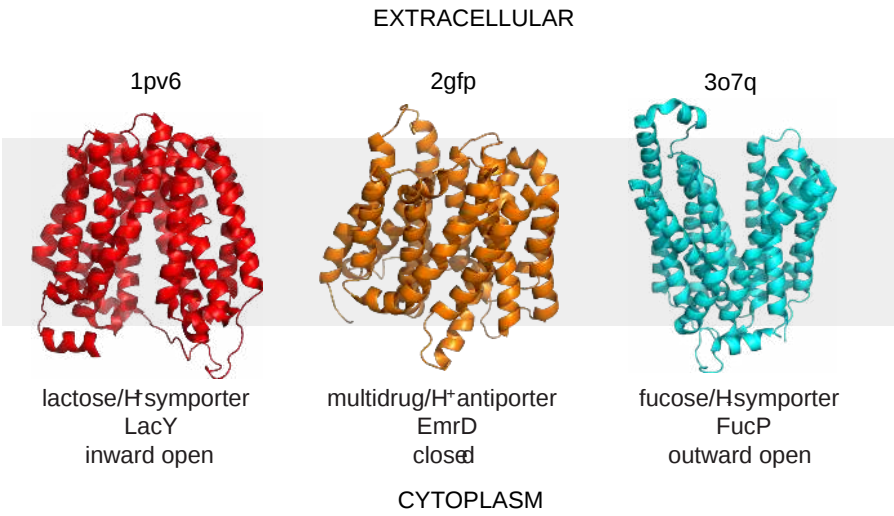


Figure 12.10.9: Inward open, closed, and outward conformations of different major facilitator superfamily transporters

Another member of the MFS Transporter family is LeuT, the leucine transport, which is similar to the neurotransmitter sodium symporter. Figure 12.10.10 shows the transition between two conformational states. The first is the outward open leucine transporter (LeuT). This initial state has bound tryptophan, a competitive inhibitor too big to transport and traps the bound state in the open conformation. The second state has bound Leu and is closed. The conformational changes are subtle but sufficient to allow transient binding, occlusion, and expulsion to the other side of the membrane (not shown). Active transport through this transporter is driven by Na^+ ions (purple). The ions are farther apart in the open state, allowing access to the amino acid for transport.

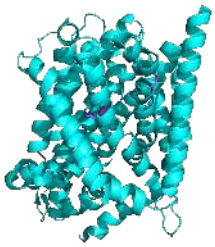


Figure 12.10.10: Transition between two conformational states of the Leu transporter

12.10.3: Membrane ATPases

Many active transporters power the uphill transport of solute through coupled exergonic cleavage of ATP. The largest family is the **P-type ATPase**, which is used to transport a host of different solutes, including ions (Cu, Zn, Mn, Mg, Ca, Na, K, Cd, Co, Pb, Ni,

and H) and phospholipids (from one leaflet to another by translocases like flippases). They convert chemical energy (from cleavage of phosphoanhydride bonds) to electrochemical potential energy. Over 160,000 P-type ATPases from prokaryotes and eukaryotes are known, so they are highly prevalent and evolved early. They should be understood at both a structural and thermodynamic level.

How is energetic coupling achieved? An early model proposed by Albers and Post suggested that energetic coupling occurred by the covalent transfer of a terminal phosphate (P_i) from ATP to an Asp in the protein channel to form a covalent Asp-Pi intermediate, which is an example of a mixed anhydride. This is illustrated in Figure 12.10.11.

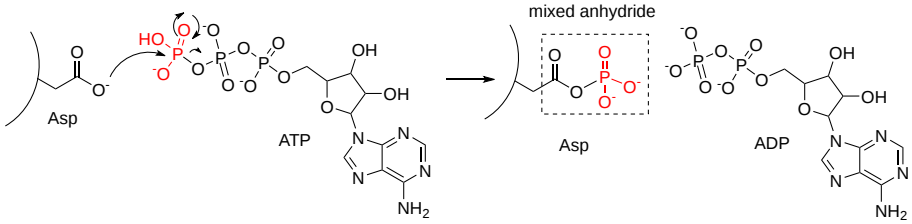


Figure 12.10.11: Formation of a high energy (compared to its hydrolysis products) mixed anhydride of Asp in membrane ATPase. (**Remember, there is no such thing as a "high energy" bond.**)

This reaction is disfavored thermodynamically since hydrolysis of ATP to form ADP proceeds with a **less negative** ΔG^0 than the same reaction with a mixed anhydride. Ultimately, hydrolysis of the mixed anhydride in the next step would favor the overall coupled reactions.

Another feature of the Post-Albers model is that the phosphorylated channel, P-ATPase, has two conformations:

- a high-affinity cation inward (cytoplasm) facing cation binding site called E_1
- a low-affinity cation outward (extracellular or luminal) facing cation binding site called E_2

Figure 12.10.12 shows a more detailed (A) and simpler (B) set of chemical equations to represent the different states of E_1 and E_2 in the Post-Albers model.

The next three figures are from the following reference: Zhang, X.C., Zhang, H. P-type ATPases use a domain-association mechanism to couple ATP hydrolysis to conformational change. *Biophys Rep* **5,** 167–175 (2019). https://doi.org/10.1007/s41048-019-0087-1. Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/

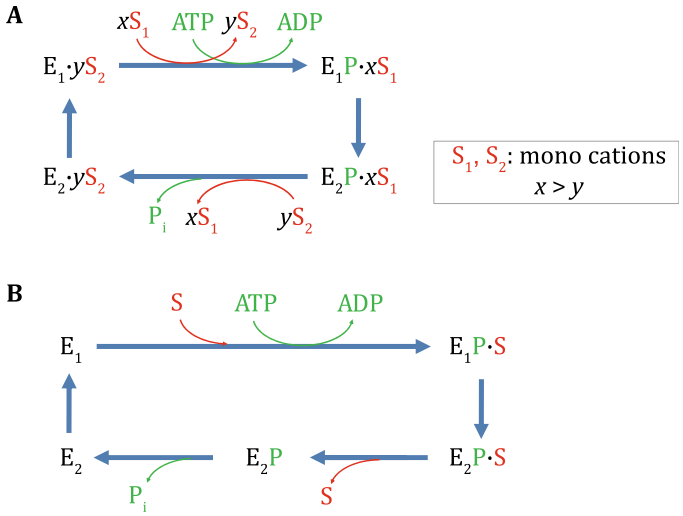


Figure 12.10.12: Chemical equations for conversion of states E_1 and E_2 for membrane ATPases in the Post-Albers mode

- The top line in both shows the reactions of E_1 (high affinity) with the substrate (ions) and ATP. Since E_1 is (auto)phosphorylated by ATP, it is considered a **kinase.**
- The bottom line in both shows the reactions of E_2 (low affinity) with the substrate (ions). Since E_2 is dephosphorylated, it is considered a **phosphatase**.

S_1 and S_2 differ for specific P-type ATPase.

Detailed Structures

Now let's look at the structures of a specific P-type ATPase, the sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (or more simply a Ca^{2+} pump), which catalyzes the active uphill transport of Ca^{2+} from the cytoplasm back into the lumen of the sarcoplasmic reticulum (SR) in muscle cells and acts in the regulation of striated muscle contraction. The actual free concentration of Ca^{2+} in the SR is about 390 μM compared to 0.1-0.2 μM in the cytoplasm. On muscle excitation, Ca^{2+} floods into the cytoplasm but must be actively transported back into the SR by the Ca^{2+} pump.

Here is the simplified chemical reaction: $ATP + Ca^{2+}(cyto) + H_2O \rightarrow ADP + Ca^{2+}(s.\ rect) + H^+ + P_i$.

We'll stick with a cartoon representation of each of the various structure states of this calcium pump, as illustrated in Figure **12.10.13** (also from Zhang et al.).

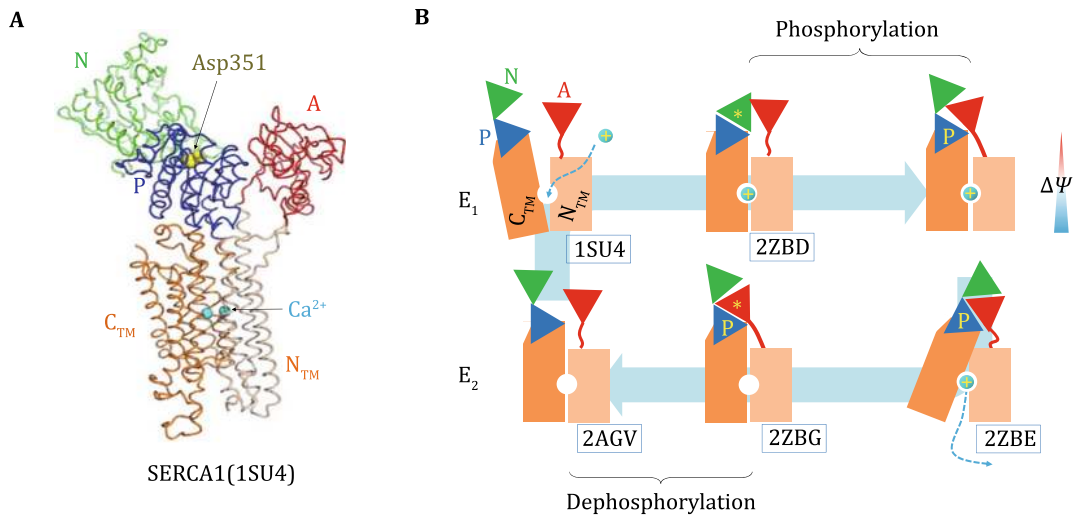


Figure **12.10.13**: States of this calcium pump

Again, different states of E_1, the kinase state, are shown in the top line, while E_2, the phosphatase state, is shown in the bottom line. PDB IDs are shown for each state. Ca^{2+} is shown as a cyan sphere. Note how the cytoplasmic N (nucleotide-binding), P (phosphorylation), and A (actuator) domains and the single transmembrane domain, colored to show the C-terminal (C_{TM}) and N-terminal (N_{TM}). are represented in the cartoon version to the right. P represents the phosphorylated Asp 351 mixed anhydride. The asterisk * denotes a transition state for the phosphorylation of Asp 351 (top line) and its dephosphorylation (bottom line).

Ca^{2+} moves from the cytoplasm to the lumen of the SR. Note that the first E_1 state is open inward (to the cytoplasm) while the first E_2 state is open outward (to the lumen of the Sr, where the bulk of the Ca^{2+} ions are stored. This mimics the open-in and open-out states of the Major Facilitator Superfamily (MFS) channels discussed above.

Detailed Thermodynamics

Zhang et al. provide two different views of the thermodynamics of Ca^{2+} ion pumping into the sarcoplasmic reticulum, as shown in A (domain dissociation) and B (classic) in Figure **12.10.14**. A quick view of the two "energy landscapes" show:

- As expected, both go from higher to lower free energy since ATP hydrolysis powers the process.
- They differ mostly in the energetic coupling steps

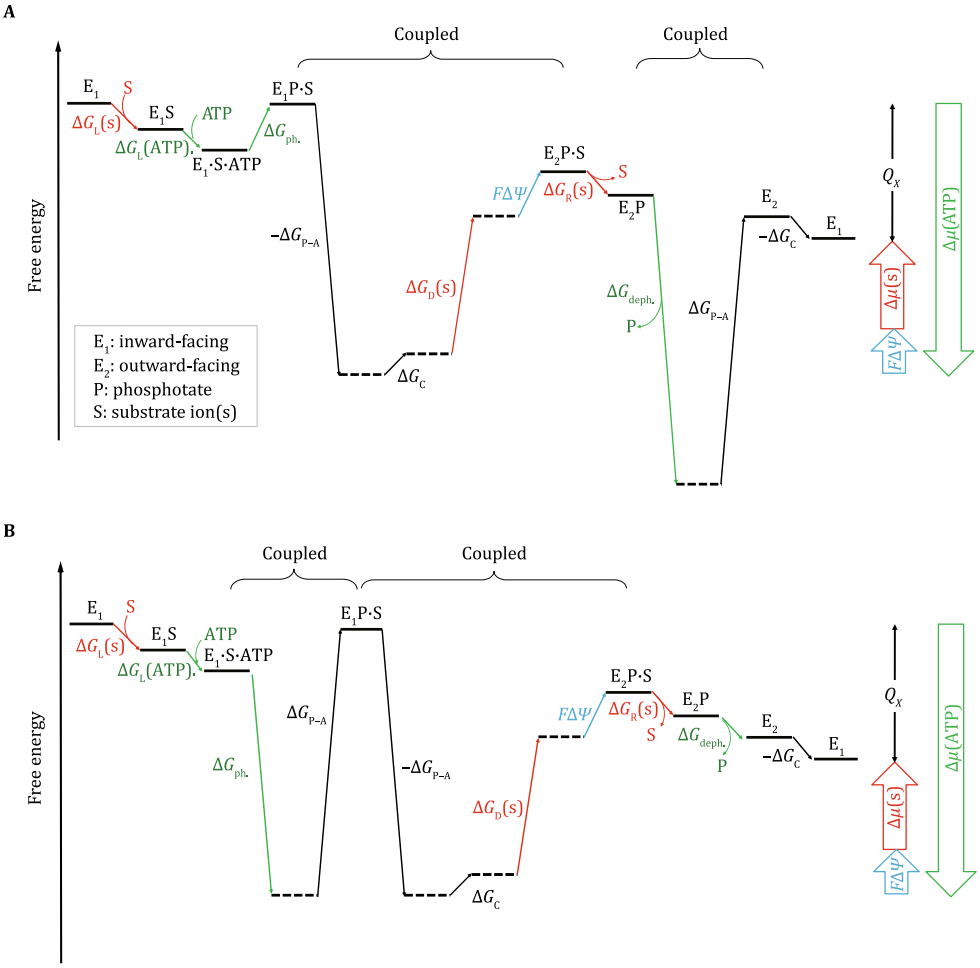


Figure **12.10.14**: Thermodynamics of Ca^{2+} ion pumping into the SR as shown in A (domain dissociation) and B (classic)

In A, the energy released in ATP hydrolysis occurs when the Asp 351-P_i mixed anhydride is dephosphorylated (ΔG_{deph}), which is coupled to the dissociation of the P and A (* actuator) domains. This is shown in the structural cartoon near the last step in the complete cycle.

In the cartoon diagram, six states are represented. Any mechanism can be broken down into more and more states on a more complete description of the actual mechanism. For example, ATP cleavage to ATP can be represented by three states, ATP, ADP, and P_i, or many more if intermediates and transition states are included. The complete Ca^{2+} transport cycle in the thermodynamics diagram is broken down into 12 states, represented by 8 horizontal solid lines —— and 4 dashed ------ lines. Vertical lines with arrows show transitions between states.

- Green arrows → show changes in free energy due to the relative stability (chemical potential) of ATP.
- Red arrows → show changes in free energy due to the relative stability (chemical potential) of the substrate ions.
- Cyan arrows → show changes in free energy due to the electrical potential of the substrate ions.

ΔG_L is for steps involving the loading of reactants, while ΔG_R is for the release. In both diagrams, the conversion of state E_1·S·ATP to E_1P·S is uphill, as we predicted for the formation of the

On the far right of each graph are identical sets of large unfilled arrows (since ΔG is a state function), as the beginning and ending free energies of free E_1 must be independent of the pathway (i.e., mechanism). The two sets of upward unfilled arrows (red for the change in the chemical potential of the substrate ions and cyan for the change in the electric potential of the substrates are both positive as the ions move uphill from a lower to high electrochemical potential. The unfilled downward green arrow shows the change in free energy (or chemical potential) for ATP hydrolysis. Since energy cannot be created or destroyed, the narrow black arrow (Q_x) represents energy that is dissipated in the reaction cycle. The starting and ending states are identical (*i.e.*, E_1), only differing by the energy dissipation (denoted as Q_X) of the P-ATPase transporter during one functional cycle.

There are other membrane ATPase types whose structure and function differ from the P-type described above. Some run in reverse to synthesize ATP. They also vary in substrate ions. Here are several different types.

- F-ATPases (ATP synthases, F_1F_0-ATPases): These are used to synthesize ATP and are powered by the collapse of a H^+ gradient. They are found in mitochondria, chloroplasts, and bacterial cell membranes. In chapters on mitochondrial oxidative phosphorylation and photosynthesis in chloroplasts, we will discuss these more.
- V-ATPases (V_1V_0-ATPases): These pump protons into organelles to acidify them (ex. lysosome) and are also found in bacteria.
- A-ATPases (A_1A_0-ATPases): These are used to synthesize ATP in Archaea.
- E-ATPases: These are found on the cell surface and hydrolyze extracellular nucleotide triphosphate.

Let's look at some special features of one P-type ATPase that transports Na^+ and K^+ ions. It is important in establishing their intracellular and extracellular concentration in neurons, so they are critical for neuron function.

Na^+/K^+ ATPase

This protein keeps the K^+_{in} and Na^+_{out} high compared to their respective concentrations on the other side of the neural cell membranes. ATP and **3** Na^+ ions bind to the cytoplasmic domain of the enzyme in the E_1 conformation. As described more generally above, in the presence of Na ions, the bound ATP is cleaved in a nucleophilic attack by an Asp side chain of the protein. Hence, the protein is a Na^+-activated ATPase or kinase. The phosphorylated enzyme changes conformation to the E_2 form in which Na^+ ions are now on the outside of the cell membrane, from which they dissociate. The phosphorylated protein in conformation E_2 now binds 2 $\mathbf{K}^+$ ions on the outside, which activates hydrolysis of the Asp-PO_3 mixed anhydride link. The dephosphorylated protein is more stable in the E1 conformation, to which it changes as it brings K^+ ions into the cell. Hence, this protein is an electrogenic antiporter. P-type transporters are inhibited by vanadate (VO_4^{3-}), a transition state analog of phosphate. In the lab, transport mediated by P-type membrane proteins can drive ATP synthesis.

A detailed kinetic analysis of ATP and VO_4^{-3} interactions shows low-affinity and high-affinity sites for each on Na/K ATPase. The high-affinity vanadate site appears to be the same as the low-affinity ATP site, which suggests that vanadate binds tightly to the E2 form of the enzyme. The low-affinity vanadate site appears to be the same (based on competition assays) as the ATP site, which is probably the E_1 form. Hence, vanadate binds preferentially to the E2 form, which would inhibit the transition to the E1 form. Vanadate also inhibits phosphatases, enzymes that cleave phosphorylated Ser, Thr, and Tyr phosphoesters in proteins. This supports the notion that vanadate binds preferentially to the E2 form, which has a phosphoanhydride link (Asp-O-phosphate) that is hydrolyzed, promoting the conversion of E_2 back to E_1. Vanadate is probably a transition-state analog inhibitor. It can readily adopt a trigonal bipyramidal structure, mimicking the transition state for cleavage of the tetrahedral anhydride bonds of ATP and Asp-O-PO_3.

Figure **12.10.15** is a YouTube animation of the Na^+/K^+ ATPase from xx. Permission Question?

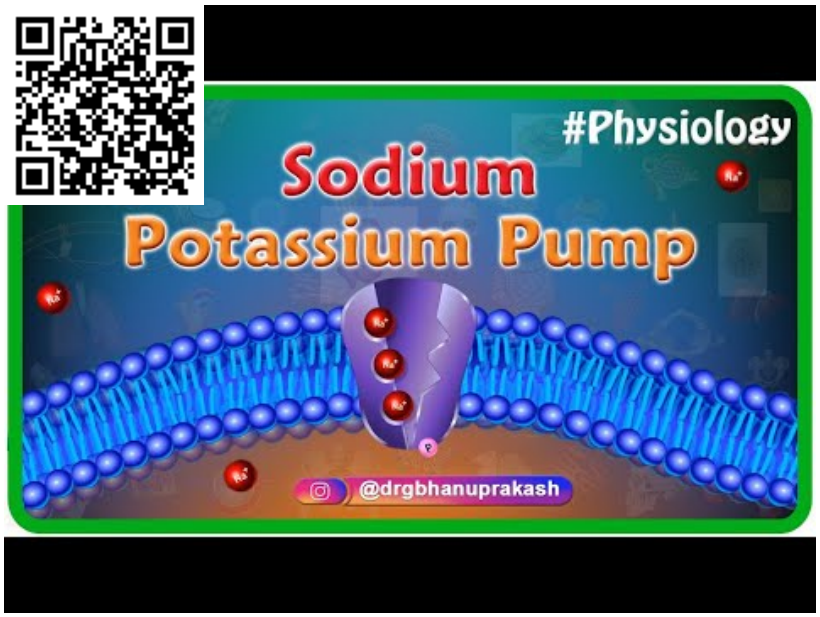


Figure 12.10.15: Animation of the Na^+/K^+ ATPase pump

These interactions are depicted in Figure 12.10.16below (after Stryer 4th ed)

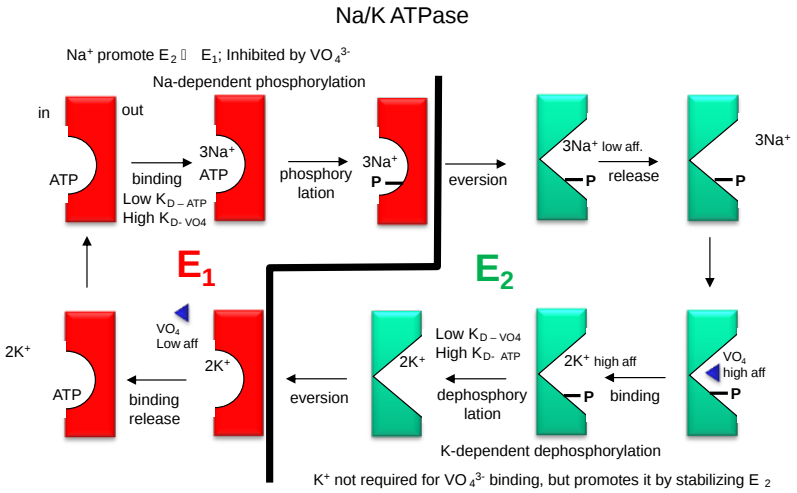


Figure 12.10.16: Steps of the Na^+/K^+ ATPase pump

ABC Transporters

The proteins comprise one of the largest families of membrane proteins with seven main families (ABCA to ABCG). Up to 3% of all bacterial proteins encode proteins associated with the ABC transporters. Different ABC transporters move a variety of required chemical species, from small (ions, sugars, amino acids, nucleosides, vitamins) to large (peptides, lipids, oligonucleotides, and polysaccharides) into the cell. They also remove toxic (to the cell) molecules such as xenobiotics (molecules foreign to the cell like drugs, toxins, etc) and potentially toxic metabolites. All of these require ATP hydrolysis. Moving toxic molecules out of the cell is beneficial to the health of the cell, but in the case of a tumor cell, not to the benefit of the organism. TAs in other active transporters using ATP, its hydrolysis leads to two different conformations, open-outward and open-inward. All ABC transporters have a LSGGQ amino acid sequence in the NBD. They also have a phosphate-binding loop (P-loop or Walker A motif). Most eukaryotic ABC transporters move solutes from the inside to the outside of the cell.

Figure 12.10.17 shows the various structures within the ABC transporter superfamily. The transmembrane domain (TMD, in green) and the nucleotide-binding domain (NBD, blue) are present in most gene versions. PK represents prokaryotic and EK eukaryotic organisms (shown at the very right for each).

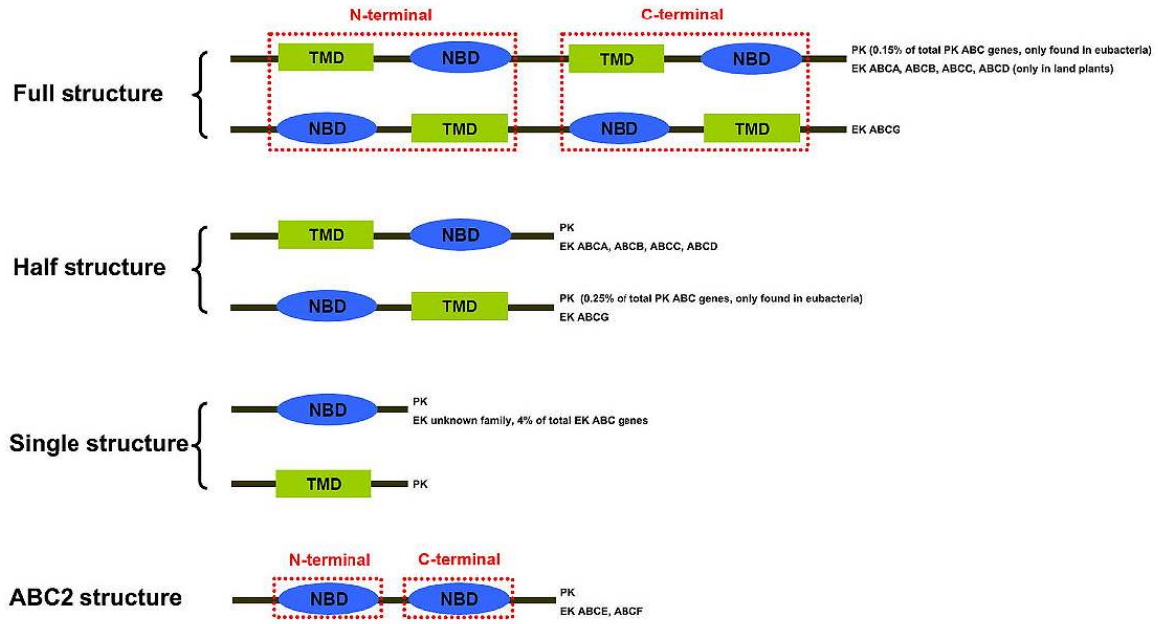


Figure 12.10.17: Variations in the structures within the ABC transporter superfamily ng, J., Feng, J., Yuan, D., *et al.* Tracing the structural evolution of eukaryotic ATP-binding cassette transporter superfamily. *Sci Rep* **5,** 16724 (2015). https://doi.org/10.1038/srep16724. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/

The ABC transporter genes denoted as full structures have 2 TMDs and 2 NMDs, while half structures have one of each. Some genes encode either a single NBD or a single TMD (prevalent in prokaryotes), along with half structures). The ABC2 structure has only two NBDs. The single structure represents the ABC transporter gene with a single TMD or NBD; ABC2 structure represents the ABC transporter gene with only two NBDs. The families possessing certain structures are listed on the right. TMDs typically have 6-10 transmembrane α-helices. Those involved in the export of chemical species have six.

Structural cartoons representing domain and ligand binding for various ABC transporters are shown in Figure 12.10.18 below.

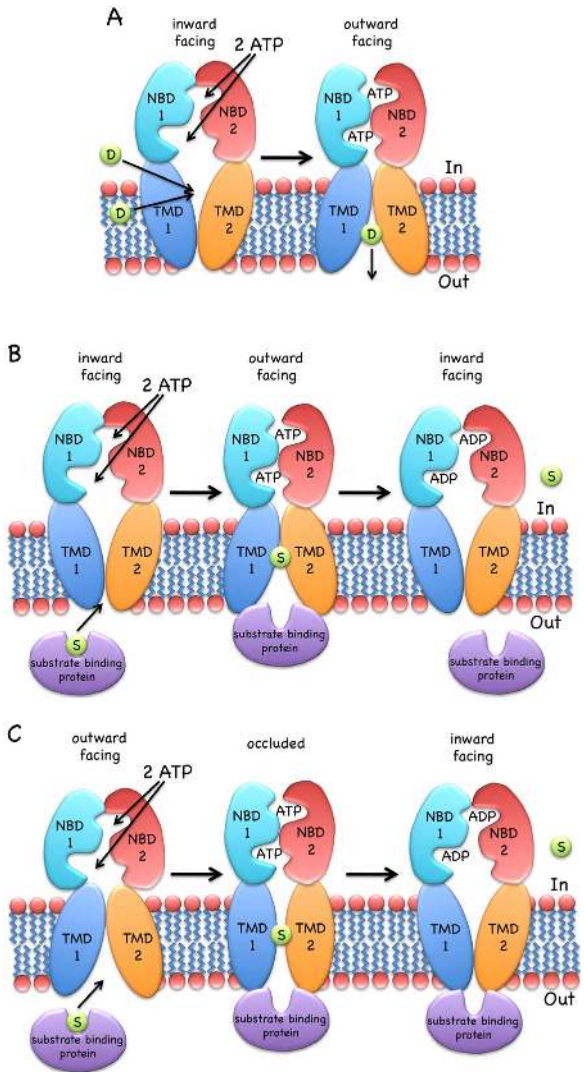


Figure 12.10.18: Domain and ligand binding for a variety of ABC transporters Structure and mechanism of ABC transporters. Stephan Wilkens. F1000Prime Reports 2015, 7:14 (doi:10.12703/P7-14). Creative Commons Attribution-Non Commercial License (http://creativecommons.org/licenses/.../3.0/legalcode).

(A) shows an inward-open drug (D) exporter. The binding of 2 ATPs causes dimer formation between the two NBD domains, resulting in a conformation change to outward facing, allowing the drug to dissociate. Hydrolysis of ATP enables the release of ADP/P_i and dissociation of the 2 NBD domain, which results in a conformational change to the inward-open form.

(B) shows the delivery of a substrate (S) from its binding protein (ex., part of the ABC transporter complex MalEFGK involved in maltose/maltodextrin import) in the periplasm of E. Coli for delivery into the cell.

(C) shows an outward-open variant of (B) (ex., part of the Vitamin B12 import system permease protein BtuC)

The mammalian protein multi-drug resistance (MDR), also known as P-glycoprotein, Phospholipid transporter ABCB1, or ATP-dependent translocase ABCB1**,** is an example of an ABC Transporter. It moves both drugs across the membrane and phospholipids, including phosphatidylcholine, phosphatidylethanolamine, ceramides, and sphingomyelins, from the inner to the outer leaflet of the membrane. The gene for this protein is often mutated in tumor cells, which moves chemotherapeutic drugs from the cell.

Figure 12.10.19 shows an interactive iCn3D model of the structure of the mouse P-glycoprotein (4M2T) bound to a cyclic peptide inhibitor in the inward open conformation is shown below. Most proteins bind substrates specifically, but P-glycoprotein binds them quite indiscriminately, which makes this protein so useful in pumping drugs out of the cell.

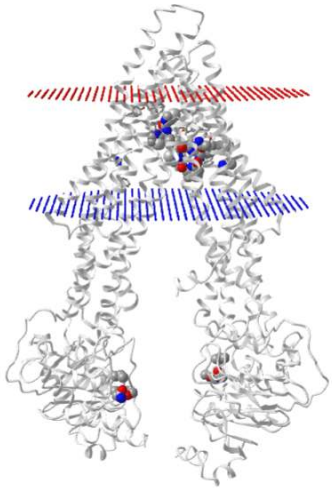


Figure 12.10.19: Mouse P-glycoprotein bound to a cyclic peptide inhibitor in the inward open conformation (4M2T). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...1bHUEmryErK5PA

The iCn3D model below shows the conserved (mouse and human) transmembrane domain aromatic amino acids (H60, F71, T114, F299, Y303, Y306, F332, F339, F724, F728, F766, F938, Y949, F953, F974, F979) involved in the transport pathway in colored sticks. The inhibitor (cyclic-tris-(S)-valineselenazole; QZ59-SSS) is shown in spaceiflll. The consensus LSGGQ (527-531 and 1172-1176) sequences in the nucleotide-binding domain are shown in colored spheres.

Here is a view showing again the transmembrane domain with the conserved aromatic amino acids.

Figure 12.10.20 shows an interactive iCn3D model of just the transmembrane domain with the conserved aromatic amino acids of the mouse P-glycoprotein bound to a cyclic peptide inhibitor in the inward open conformation (4M2T).

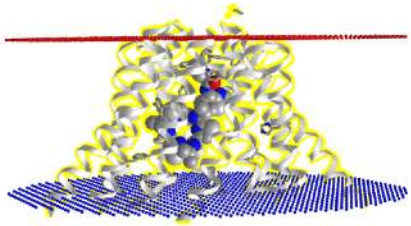


Figure 12.10.20: Transmembrane domain of the mouse P-glycoprotein bound to a cyclic peptide inhibitor in the inward open conformation (4M2T). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PdHtggbLFWuCo7

12.10.4: Another look at the Nuclear Pore Complex

In the previous section, we examined the incredibly complicated structure of the nuclear pore. Both small molecules and large proteins (synthesized in the cytoplasm) and RNAs (synthesized in the nucleus) move through their large pores. The movement of solutes through its gaping pore does not require ATP cleavage in a primary active transport or the collapse of a chemical gradient in a secondary active transport process. Then why study it in this chapter section? It turns out that the regulation of the process requires **GTP** cleavage and a gradient of a particular protein called **Ran**.

What's fascinating about the nuclear pore is its selectivity for protein transfer across the pore. What proteins are allowed in and out? What is the origin of the specificity? The specificity is partly determined by a protein family called the karyopherin-βs (22) of receptors, which bind and transport nuclear proteins. There are two types of these nuclear **transport receptors or transport factors**:

- **Importins** facilitate the movement of proteins into the nucleus (ex., nuclear proteins like histones, DNA and RNA polymerases, etc).
- **Exportins** facilitate the movement of RNA (except mRNA) out of the nucleus.

Large proteins destined to be moved through the nuclear pore are called **cargo** proteins. They have a molecular signal that differentiates them from molecules that should stay in the cytoplasm, move into organelles, or be secreted from the cell. The signals are called:

- **NLS** or nuclear localization sequences
- **NES** or nuclear export sequences

There is no obvious NLS consequence, and different cargos use different motifs. A classical NLS is enriched in Lys and Arg residues. Others appear enriched in Pro-Tyr or Ile-Lys. Large domain structures might also participate in the NLSs. Prediction programs are used to identify clusters of lysines and arginines with gaps between the clusters. Classical NESs appear rich in leucines in an 8-15 amino acid sequence with regularly spaced conserved hydrophobic amino acids.

A third player, **RAN** (Ras-related nuclear protein), determines which direction the transport receptor:cargo complex moves. Ran is a small protein that binds and can hydrolyze GTP to GDP. It is a member of the small G proteins that are GTPases.

Ran itself can run (a pun) or move across the nuclear membrane and is found in the cytoplasm and nucleus. What determines which way a Cargo:receptor complex goes? It depends on whether GTP or GDP is bound to RAN! Higher concentrations of RanGTP are found in the nucleus, and lower concentrations in the cytoplasm.

- Importins bind cargo proteins with a NLS in the cytoplasm where RanGTP is low, move into the nucleus, and release the cargo protein when the abundant RanGTP displaces the bound cargo protein.
- Exportins bind both a cargo protein with a NES and Ran:GTP to form a ternary Exportin:Cargo:RanGTP complex in the nucleus. This moves into the cytoplasm, where the Ran bound GTP is hydrolyzed to GDP, causing the complex to dissociate, freeing the exported cargo protein from the complex.

Importins are dimeric structures consisting of an alpha and a beta subunit. The alpha subunit binds to the cargo protein through the NLS sequence on the cargo protein. The beta subunit interacts with the nucleoporin proteins (Nups) in the

These interactions are shown in Figure 12.10.21.

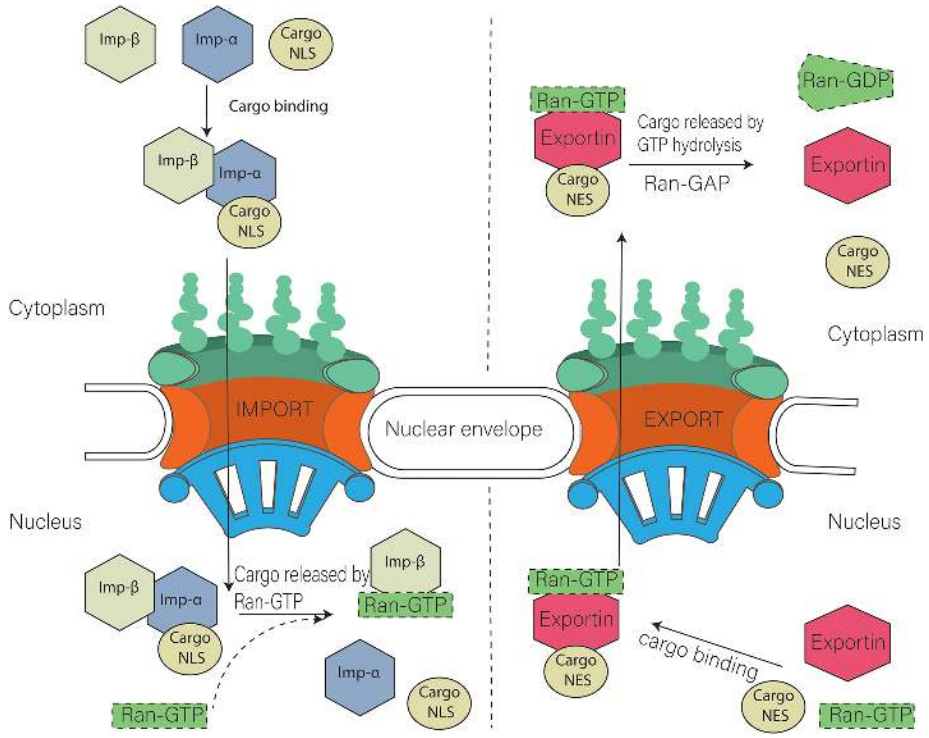


Figure 12.10.21: Model of nuclear import and export. Cargo containing NLS (Nuclear localization signal) is imported with the help of Importin α and importin β heterodimer. Nuclear export of cargo having NES (nuclear export signal) is carried out with the help of exportins. Ran GTP is also required during that process. Khan Asmat Ullah, Qu Rongmei, Ouyang Jun, Dai Jingxing. Front. Physiol., 03 April 2020 | https://doi.org/10.3389/fphys.2020.00239. **Creative Commons Attribution License (CC BY)**.

Figure 12.10.22 shows an interactive iCn3D model of the Delpi electrostatic surface potential map of the putative NLS from the carboxy-terminal of a cargo protein, the W protein of the Hendra virus (4M2T), which is imported into the nucleus.

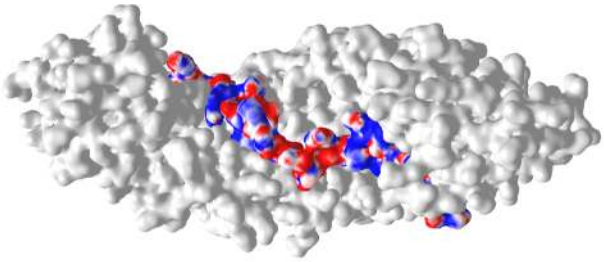


Figure 12.10.22: Delphi electrostatic surface potential map of the putative NLS from the carboxy-terminal of a cargo protein, the W protein of the Hendra virus (6BW9) which is imported into the nucleus (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...mUkDXG1KoiieTA

The virus derives from bats and has recently emerged. The peptide sequence containing the NLS has the following sequence: 419 CLGRRVVQPGMFADYPPTKKARVLLR 444. The red surface indicates the negative and blue positive potential, which is associated with the Lys and Arg side chains in the consensus sequence. It is bound to the human importin-α3 subunit (molecular surface shown in white), which binds the NLS sequence with high affinity.

An astute reader might ask why GTP stays bound and is not cleaved into GDP in the nucleus due to Ran's intrinsic GTPase activity. There is yet another protein found only in the cytoplasm (i.e., it doesn't have a nuclear import signal), which binds to Ran:GTP and promotes GTP → GDP exchange. So oddly, hydrolysis of a nucleotide triphosphate (GTP, not ATP, as in the case of most transporters discussed above) is required for nuclear import. A different but very fascinating mechanism!

The cargo complexes described above must still pass through the nuclear pore. In section 11.3, we discussed the structure of the nuclear pore, which consists of many nucleoporin proteins (NUPs). The **inner ring** is called the **FG Nups layer.** The disordered parts of the inner ring Nups have repetitive sequences enriched in phenylalanine and glycine (hence the name FG Nups) and stick out into the central pore. The FG repeats can take multiple forms, including Phe-Gly (FG), Gly-Leu-Phe-Gly (GLFG), or Phe-any-Phe-Gly (FXFG). These are intrinsically disordered proteins, and the disordered regions in the pore act as a filter, allowing certain molecules to pass and excluding those with molecular weights greater than 40K. Large molecules complexed to importins/exportins move through the pore. This disordered mesh prevents passive diffusion of molecules but allows protein complexes with a cargo:exportin/importin complex through. It's a bit like electrophoresis of small protein complexes through the pores of an acrylamide or agarose gel polymerized matrix, only without the "push" of an electric field. Presumably, the FG-Nups make transient hydrophobic (induced-dipole:induced dipole) interactions between the FGs on the Nups and the nuclear transport receptors.

Figure 12.10.23 shows an **export** complex (5XOJ) from yeast of Ran (cyan with bound GTP in spacefill) bound to an **exportin** (Xpo1p, white surface) and 3 Nup42p peptides containing SxFG/PxFG repeats (spacefill, side chains, which are mostly nonpolar). To bind 3 Nup peptides, the exportin Xpo1 also contains repeating binding sites on domains called HEAT repeats 14–20 of Xpo1p. The exportin Xpo1p is shaped like a toroid with 21 HEAT repeat domains that have two antiparallel sheets with connecting loops of different sizes.

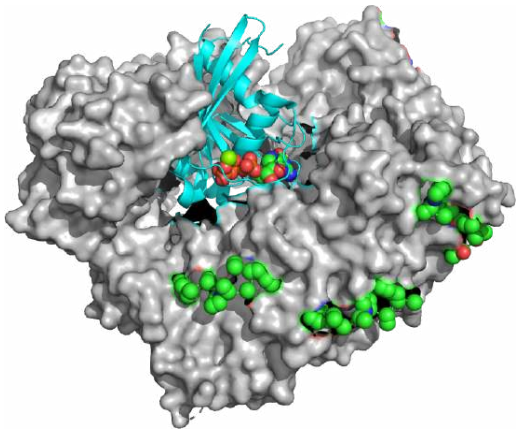


Figure 12.10.23: Yeast export complex (5XOJ) with Ran (cyan with bound GTP in spacefill) bound to an **exportin** (Xpo1p, gray surface) and three Nup42p peptides containing SxFG/PxFG repeats (spacefill, side chains))

Some believe the unfolded mesh of FGs in the core might condense on the binding of motifs on importins/exportins or through FG:FG domain associations. But what is the basis for importins/exportins interactions with the FG structures? To study it in more detail, Fragasso et al. designed and made an artificial FG-Nup that binds to an importin transport receptor, **Kap95,** which interacts with a cargo protein with a nuclear localization sequence (NLS). They attached them to the inside of solid-state nanopores and demonstrated the fast movement of Kap95 through the derivatized nanopore while a control "non cargo" without a NLS, BSA) was blocked. Underivatized nanometer-sized pores are made in a silicon nitride synthetic membrane (SiN) by using an ion or electron beam to tune the size of the hole.

Figure 12.10.24 (top image) shows colored code structures of the inner ring (top of figure) snapshots of different yeast GLFG NUPs. These are named after their Gly-Leu-Phe-Gly repeating motifs and are especially cohesive. The GLFG-Nups, shown in red, are mostly found in the inner ring compared to FxFG/FG-Nups in green.

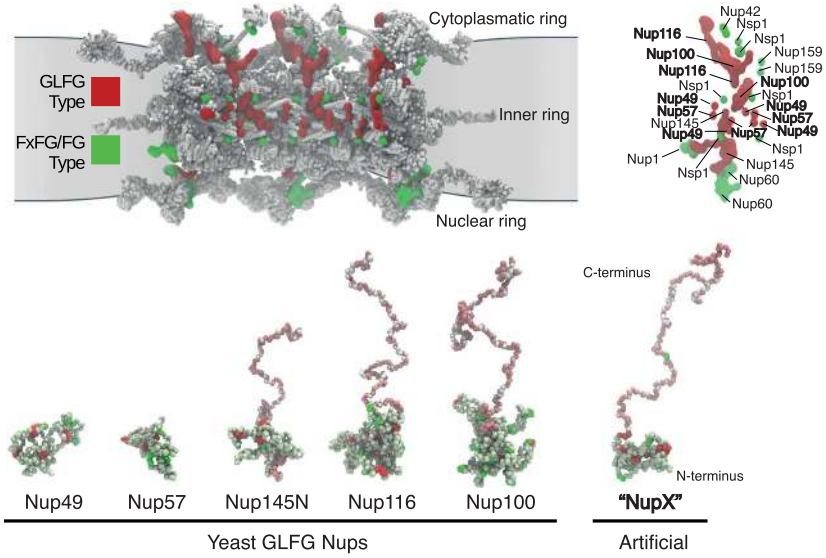


Figure 12.10.24: Inner ring (top of figure) snapshots of different yeast GLFG nups. Fragasso, A., de Vries, H.W., Andersson, J., *et al.* A designer FG-Nup that reconstitutes the selective transport barrier of the nuclear pore complex. *Nat Commun* **12,** 2010 (2021). https://doi.org/10.1038/s41467-021-22293-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The bottom part of Figure 12.10.24 shows snapshots of molecular dynamics simulations of various yeast GLFG-Nups compared to the NupX synthetic one. Three yeast GLFG Nups and the synthetic NupX show a compact and extended domain. Each has a collapsed cohesive domain at one end (characterized by a low charge to hydrophobic amino acid residue ratio (C/H), lots of alternating FG and GLFG repeats, with light green showing non-FG/GLFG/charged residues) and an extended domain at the other

end (high ratio of charged/hydrophobic amino acids, no FG repeats, pink-red showing non-FG/GLFG/charged residues). Bright green shows the FG repeats, red the GLFG repeats, and white the charged groups.

Figure 12.10.25 shows the pore in the silicon nitride synthetic membrane (SiN) membrane before (left) and after derivitization with NupX.

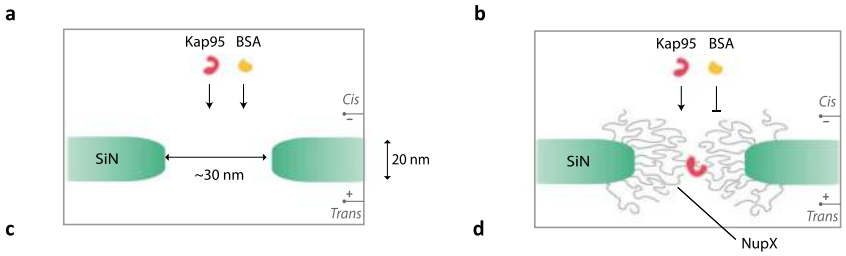


Figure 12.10.25: Pore in the SiN silicon nitride synthetic membrane before (left) and after derivatization with NupX

Figure 12.10.26 shows increasing size derivatized pores in the SiN membrane. In the smallest particle, the NupXs don't fit and are expelled from the pore. In the 30 nm particle, the NupX cohesive domains plug the hole. In the 45 nm pore, a hole appears.

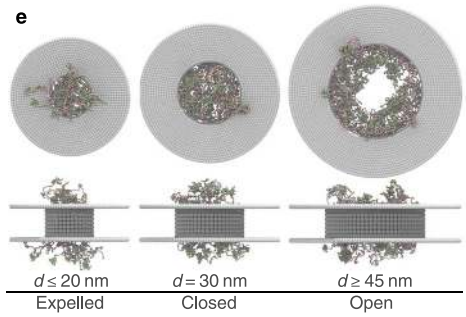


Figure 12.10.26: Derivatized pores in the silicon nitride (SiN) membrane of increasing size

Lastly, Figure 12.10.27 shows snapshots of 30 nm pores lined with NupXs. In panel c, an importin, Kap95 (spheres with orange binding spots), is shown forming a transient interaction with and translocating through the NupX-lined hole. BSA (sphere without binding sites) may interact weakly but does not translocate through the pore.

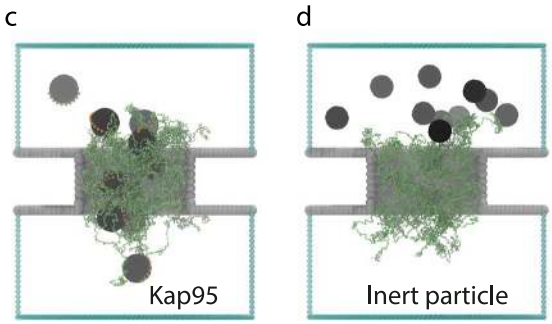


Figure 12.10.27: Snapshots of 30 nm pores lined with NupXs and interacting with an importin, Kap95 (spheres with orange binding spots, panel C) and BSA (sphere, panel D)

12.10.5: Summary

This chapter provides a comprehensive overview of membrane transport mechanisms with a focus on active transport processes essential for cell function. The material contrasts passive transport methods, such as facilitated diffusion and diffusion through channels or pores, with active transport, which moves solutes against their concentration gradients using various energy sources. Key topics include:

- **Energy-Driven Transport:**
  The chapter explains how cells harness different energy sources—ATP hydrolysis, energy from oxidation, light, and the collapse of ion gradients—to drive active transport. For example, ATP powers the Na^+/K^+ ATPase, while light energy is utilized by bacteriorhodopsin in photosynthetic bacteria.
- **Types and Classifications of Transporters:**
  Active transporters are classified into primary transporters, which directly use energy from ATP or photons, and secondary transporters, which rely on coupled gradients (cotransport systems). The text also distinguishes between uniporters, symporters, and antiporters, with specific examples such as GLUT1 for facilitated diffusion, the glucose symporter (SGLT1), and the sodium-calcium exchanger.
- **Major Facilitator Superfamily (MFS) Transporters:**
  The MFS is highlighted as the largest group of secondary transporters. Their characteristic structure—composed of 12 transmembrane helices arranged in two 6-helix bundles—supports a “clamp and switch” mechanism that allows solute binding on one side of the membrane, occlusion, and subsequent release on the other side. Detailed case studies, such as human monocarboxylate transporters and lactose permease, illustrate how key residues facilitate substrate binding and proton coupling.
- **Membrane ATPases and Ion Pumps:**
  The discussion extends to P-type ATPases, which are critical for establishing ion gradients. The chapter examines the Ca^{2+} pump in the sarcoplasmic/endoplasmic reticulum and the Na^+/K^+ ATPase, detailing their conformational transitions between E1 (high-affinity) and E2 (low-affinity) states, and how ATP hydrolysis drives these processes. Inhibitory effects of agents like vanadate are also considered.
- **ABC Transporters:**
  Another major transporter family covered is the ABC (ATP-binding cassette) transporters. These proteins, which feature distinct transmembrane and nucleotide-binding domains, are involved in both nutrient uptake and the export of toxic compounds, with clinical implications in multidrug resistance highlighted by the example of P-glycoprotein.
- **Nuclear Pore Complex and Selective Transport:**
  The chapter concludes by exploring the nuclear pore complex, a unique structure that regulates the transport of large macromolecules between the nucleus and cytoplasm. This section emphasizes the roles of FG-nucleoporins, karyopherins (importins and exportins), and Ran GTPase in ensuring selective and energy-dependent transport across the nuclear envelope.

Overall, the chapter integrates structural, kinetic, and thermodynamic principles to elucidate how a variety of transport proteins work in concert to maintain cellular homeostasis, adapt to environmental challenges, and facilitate communication between cellular compartments. This detailed analysis equips biochemistry majors with both theoretical frameworks and practical examples necessary for a deeper understanding of membrane transport processes.

This page titled 12.10: Active Transport is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.5: Active Transport** by Henry Jakubowski and Patricia Flatt has no license indicated.
- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

CHAPTER OVERVIEW

13: Biochemistry of Our Senses

13.1: Cell Signaling
13.2: Ligand-gated Ion Channel Receptors
13.3: Nuclear Hormone Receptors
13.4: G-protein Coupled Receptors (GPCRs)
13.5: Receptor Tyrosine Kinases (RTKs)
13.6: Gated Ion Channels - Neural Signaling
13.7: Signal Transduction - Vision and Olfaction
13.8: Signal Transduction - Taste (Gustation)
13.9: Signal Transduction - Temperature
13.10: Signal Transduction - Pressure

13: Biochemistry of Our Senses is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

13.1: Cell Signaling

How do cells receive signals from their environment and how do they communicate among themselves? It is intuitively obvious that even bacterial cells must be able to sense features of their environment, such as the presence of nutrients or toxins, if they are to survive. In addition to being able to receive information from the environment, multicellular organisms must find ways by which their cells can communicate among themselves. Since different cells take on specialized functions in a multicellular organism, they must be able to coordinate activities perfectly like the musicians in an orchestra performing a complicated piece of music. Cells grow, divide, or differentiate in response to specific signals. They may change shape or migrate to another location. At the physiological level, cells in a multicellular organism, must respond to everything from a meal just eaten to injury, threat or the availability of a mate. They must know when to repair damage to DNA, when to undergo apoptosis (programmed cell death) and even when to regenerate a lost limb. A variety of mechanisms have arisen to ensure that cell-cell communication is not only possible, but astonishingly swift, accurate and reliable.

How are signals sent between cells?
Like pretty much everything that happens in cells, signaling is dependent on molecular recognition. The basic principle of cell-cell signaling is simple. A particular kind of molecule, sent by a signaling cell, is recognized and bound by a receptor protein in (or on the surface of) the target cell. The signal molecules are chemically varied- they may be proteins, short peptides, lipids, nucleotides or catecholamines, to name a few. The chemical properties of the signal determine whether its receptors are on the cell surface or intracellular. If the signal is small and hydrophobic it can cross the cell membrane and bind a receptor inside the cell. If, on the other hand, the signal is charged, or very large, it would not be able to diffuse through the plasma membrane. Such signals need receptors on the cell surface, typically transmembrane proteins that have an extracellular portion that binds the signal and an intracellular part that passes on the message within the cell.

Receptors are specific for each type of signal, so each cell has many different kinds of receptors that can recognize and bind the many signals it receives. Because different cells have different sets of receptors, they respond to different signals or combinations of signals. The binding of a signal molecule to a receptor sets off a chain of events in the target cell. These events could cause change in various ways, including, but not limited to, alterations in metabolic pathways or gene expression in the target cell.

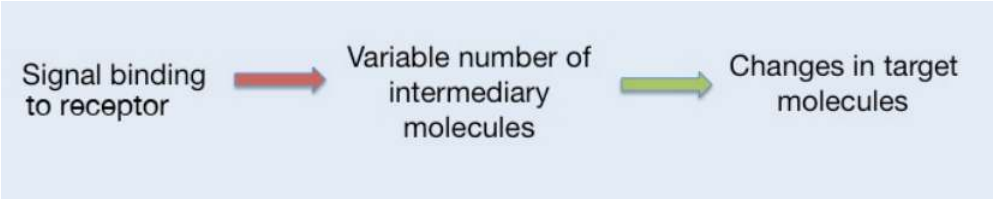


Figure 8.1.1: *Cellular Signaling*

How the binding of a signal to a receptor brings about change in cells is the topic of this section. Although the specific molecular components of the various signal transduction pathways differ, they all have some features in common:

- The binding of a signal to its receptor is usually, though not always, followed by the generation of a new signal(s) within the cell. The process by which the original signal is converted to a different form and passed on within the cell to bring about change is called signal transduction.
- Most signaling pathways have multiple signal transduction steps by which the signal is relayed through a series of molecular messengers that can amplify and distribute the message to various parts of the cell.
- The last of these messengers usually interacts with a target protein(s) and changes its activity, often by phosphorylation.

When a signal sets a particular pathway in motion, it is acting like an ON switch. This means that once the desired result has been obtained, the cell must have a mechanism that acts as an OFF switch.

Understanding this underlying similarity is helpful, because learning the details of the different pathways becomes merely a matter of identifying which molecular component performs a particular function in each individual case. We will consider several different signal transduction pathways, each mediated by a different kind of receptor. The first two examples we will examine are those with the fewest steps between the binding of the signal by a receptor and a cellular response.

This page titled 13.1: Cell Signaling is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **8.1: Cell Signaling** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

13.2: Ligand-gated Ion Channel Receptors

The simplest and fastest of signal pathways is seen in the case of signals whose receptors are gated ion channels. Gated ion channels are made up of multiple transmembrane proteins that create a pore, or channel, in the cell membrane. Depending upon its type, each ion channel is specific to the passage of a particular ionic species. The term "gated" refers to the fact that the ion channel is controlled by a "gate" which must be opened to allow the ions through. The gates are opened by the binding of an incoming signal (ligand) to the receptor, allowing the almost instantaneous passage of millions of ions from one side of the membrane to the other. Changes in the interior environment of the cell are thus brought about in microseconds and in a single step.

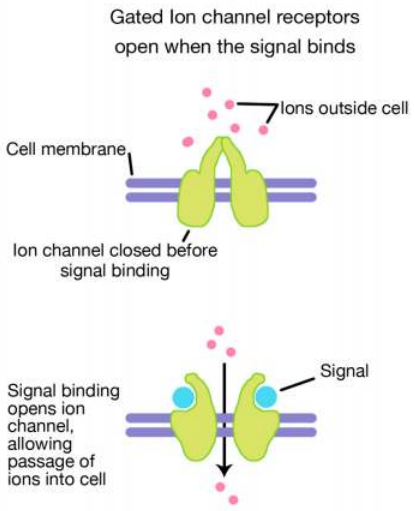


Figure 8.2.1: *Signaling Through Gated Ion Channels*

This type of swift response is seen, for example, in neuromuscular junctions, where muscle cells respond to a message from the neighboring nerve cell. The nerve cell releases a neurotransmitter signal into the synaptic cleft, which is the space between the nerve cell and the muscle cell it is "talking to". Examples of neurotransmitter signal molecules are acetylcholine and serotonin, shown in Figure 8.2.2.

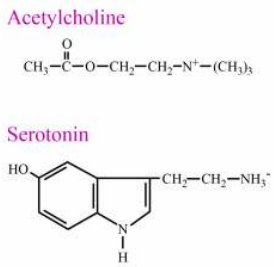


Figure 8.2.2: *Neurotransmitter*

When acetylcholine molecules are released into the synaptic cleft (the space between the pre- and post-synaptic cells) they diffuse rapidly till they reach their receptors on the membrane of the muscle cell. The binding of the acetylcholine to its receptor, an ion channel on the membrane of the muscle cell, causes the gate in the ion channel to open. The resulting ion flow through the channel can immediately change the membrane potential. This, in turn, can trigger other changes in the cell. The speed with which changes are brought about in neurotransmitter signaling is evident when you think about how quickly you remove your hand from a hot surface. Sensory neurons carry information to the brain from your hand on the hot surface and motor neurons signal to your muscles to move the hand, in less time than it took you to read this sentence!

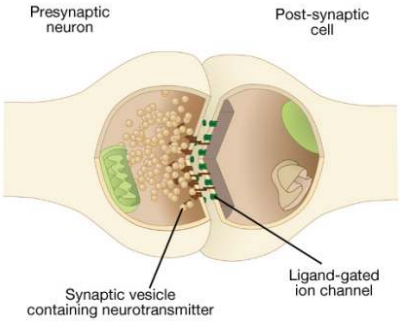


Figure 8.2.3: *Signaling across nerve cells*

This page titled 13.2: Ligand-gated Ion Channel Receptors is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **8.2: Ligand-gated Ion Channel Receptors** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

13.3: Nuclear Hormone Receptors

Another type of relatively simple, though much slower, signaling is seen in pathways in which the signals are steroid hormones, like estrogen or testosterone, pictured below. Steroid hormones, as you are aware, are related to cholesterol, and as hydrophobic molecules, they are able to cross the cell membrane by themselves. This is unusual, as most signals coming to cells are incapable of crossing the plasma membrane, and thus, must have cell surface receptors.

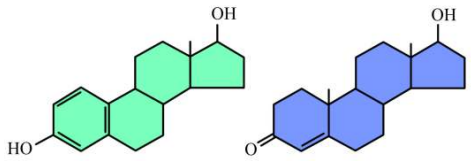


Figure 8.3.1: *Estradiol and Testosterone*

By contrast, steroid hormones have receptors inside the cell (intracellular receptors). Steroid hormone receptors are proteins that belong in a family known as the *nuclear receptors*. Nuclear hormone receptors are proteins with a double life: they are actually dormant transcription regulators. In the absence of signal, these receptors are in the cytoplasm, complexed with other proteins (HSP in Figure 8.3.2) and inactive. When a steroid hormone enters the cell, the nuclear hormone receptor binds the hormone and dissociates from the HSP. The receptors, then, with the hormone bound, translocate into the nucleus.

In the nucleus, Nuclear hormone receptors regulate the transcription of target genes by binding to their regulatory sequences (labeled HRE for hormone- response elements). The binding of the hormone-receptor complex to the regulatory elements of hormone-responsive genes modulates their expression. Because these responses involve gene expression, they are relatively slow. Most other signaling pathways, besides the two we have just discussed, involve multiple steps in which the original signal is passed on and amplified through a number of intermediate steps, before the cell responds to the signal.

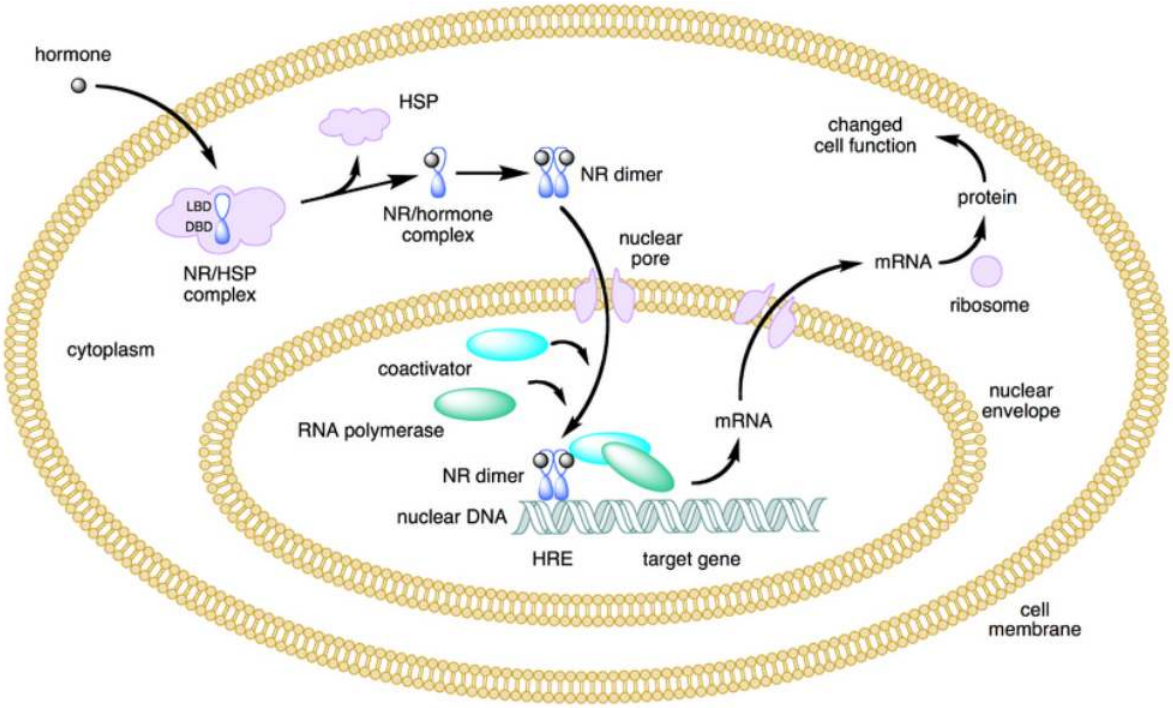


Figure 8.3.2: *Steroid hormones act by modulating expression of hormone-responsive genes*

We will now consider two signaling pathways, each mediated by a major class of cell surface receptor- the G-protein coupled receptors (GPCRs) and the receptor tyrosine kinases (RTKs). While the specific details of the signaling pathways that follow the binding of signals to each of these receptor types are different, it is easier to learn them when you can see what the pathways have

in common, namely, interaction of the signal with a receptor, followed by relaying the signal through a variable number of intermediate molecules, with the last of these molecules interacting with target protein(s) to modify their activity in the cell.

This page titled 13.3: Nuclear Hormone Receptors is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **8.3: Nuclear Hormone Receptors** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

13.4: G-protein Coupled Receptors (GPCRs)

G-protein coupled receptors are involved in responses of cells to many different kinds of signals, from epinephrine, to odors, to light. In fact, a variety of physiological phenomena including vision, taste, smell and the fight-or-flight response are mediated by GPCRs.

What are G-protein coupled receptors?

G-protein coupled receptors are cell surface receptors that pass on the signals that they receive with the help of guanine nucleotide binding proteins (a.k.a. G-proteins). Before thinking any further about the signaling pathways downstream of GPCRs, it is necessary to know a few important facts about these receptors and the G-proteins that assist them. Though there are hundreds of different G-protein coupled receptors, they all have the same basic structure: they all consist of a single polypeptide chain that threads back and forth seven times through the lipid bilayer of the plasma membrane. For this reason, they are sometimes called seven- pass transmembrane (7TM) receptors.

One end of the polypeptide forms the extracellular domain that binds the signal while the other end is in the cytosol of the cell.

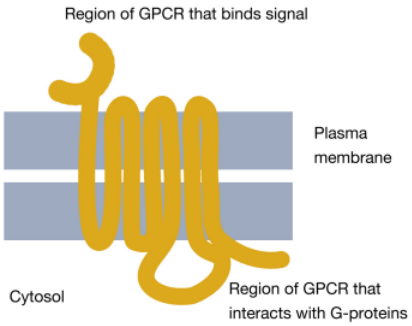


Figure 8.4.1: *G-protein coupled receptor*

When a ligand (signal) binds the extracellular domain of a GPCR, the receptor undergoes a conformational change that allows it to interact with a G-protein that will then pass the signal on to other intermediates in the signaling pathway.

What is a G-protein?

As noted above, a G-protein is a guanine nucleotide-binding protein that can interact with a G-protein linked receptor. G-proteins are associated with the cytosolic side of the plasma membrane, where they are ideally situated to interact with the cytosolic tail of the GPCR, when a signal binds to the GPCR.

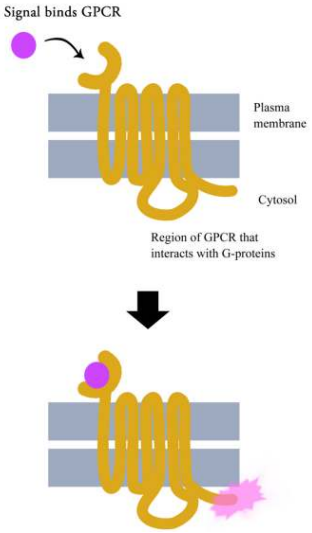


Figure 8.4.2: *G-protein coupled receptor signaling*

There are many different G-proteins, all of which share a characteristic structure- they are composed of three subunits called alpha, beta and gamma (aß.). Because of this, they are sometimes called heterotrimeric G proteins (hetero=different, trimeric= having three parts). The a subunit of such proteins can bind GDP or GTP and is capable of hydrolyzing a GTP molecule bound to it into GDP. In the unstimulated state of the cell, that is, in the absence of a signal bound to the GPCR, the G-proteins are found in the trimeric form (aß. bound together) and the a subunit has a GDP molecule bound to it.

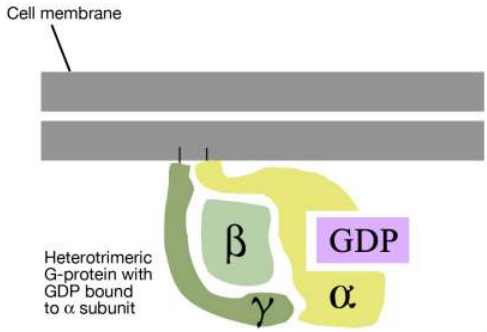


Figure 8.4.3: *G-protein with GDP Bound*

With this background on the structure and general properties of the GPCRs and the G-proteins, we can now look at what happens when a signal arrives at the cell surface and binds to a GPCR. The binding of a signal molecule by the extracellular part of the G-protein linked receptor causes the cytosolic tail of the receptor to interact with, and alter the conformation of, a G-protein. This has two consequences:

- First, the alpha subunit of the G- protein loses its GDP and binds a GTP instead.
- Second, the G-protein breaks up into the GTP-bound a part and the ß. part.

These two parts can diffuse freely along the cytosolic face of the plasma membrane and act upon their targets.

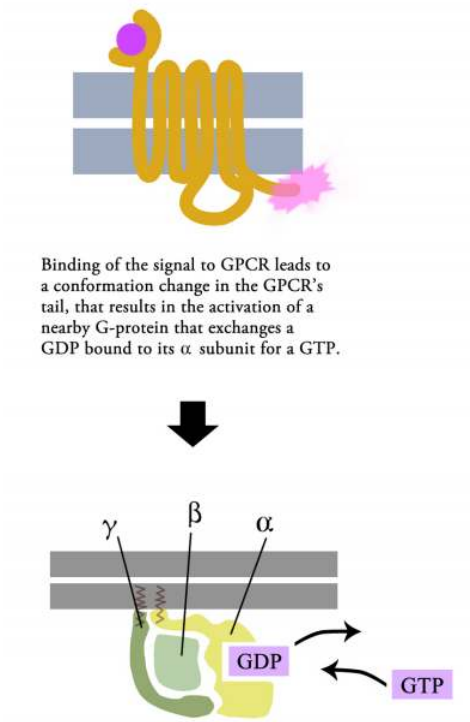


Figure 8.4.4: *G-protein Activation*

What happens when G-proteins interact with their target proteins? That depends on what the target is. G-proteins interact with different kinds of target proteins, of which we will examine two major categories:

Ion Channels

We have earlier seen that some gated ion channels can be opened or closed by the direct binding of neurotransmitters to a receptor that is an ion-channel protein. In other cases, ion channels are regulated by the binding of G-proteins. That is, instead of the signal directly binding to the ion channel, it binds to a GPCR, which activates a G-protein that then binds and opens the ion channel. The change in the distribution of ions across the plasma membrane causes a change in the membrane potential.

Specific Enzymes

The interaction of G-proteins with their target enzymes can regulate the activity of the enzyme, either increasing or decreasing its activity. Often the target enzyme will pass the signal on in another form to another part of the cell. As you might imagine, this kind of response takes a little longer than the kind where an ion channel is opened instantaneously. Two well-studied examples of enzymes whose activity is regulated by a G-protein are adenylate cyclase and phospholipase C. When adenylate cyclase is activated, the molecule cAMP is produced in large amounts.

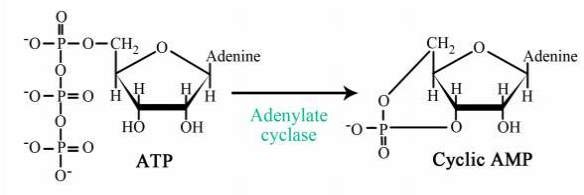


Figure 8.4.5: *Synthesis of cAMP*

When phospholipase C is activated, the molecules inositol trisphosphate (IP3) and diacylglycerol (DAG) are made. cAMP, IP3 and DAG are second messengers, small, diffusible molecules that can "spread the message" brought by the original signal, to other parts of the cell.

In these cases, the binding of a signal to the GPCR activated a G- protein, which in turn, activated an enzyme that makes a second messenger that can amplify the message in the cell. We will first trace the effects of activating adenylate cyclase and the resulting increase in cAMP.

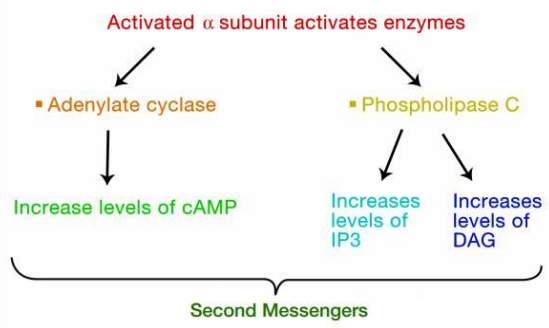


Figure 8.4.6: *Second Messengers*

What is the effect of elevated cAMP levels?

cAMP molecules bind to, and activate an enzyme, protein kinase A (PKA). PKA is composed of two catalytic and two regulatory subunits that are bound tightly together. Upon binding of cAMP the catalytic subunits are released from the regulatory subunits, allowing the enzyme to carry out its function, namely phosphorylating other proteins.

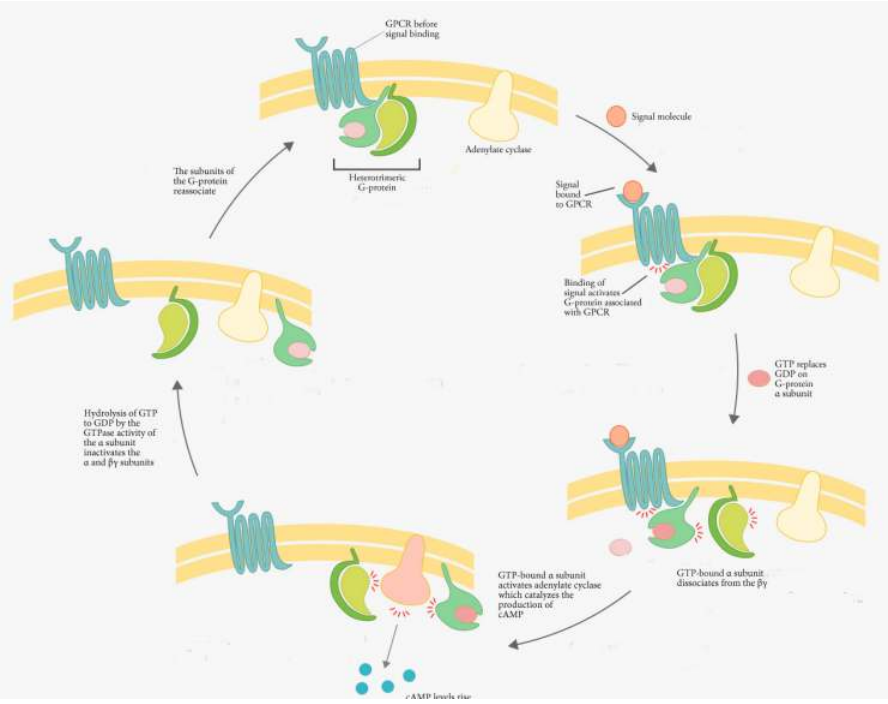


Figure 8.4.7: *G-protein Signaling Cycle*

Thus, cAMP can regulate the activity of PKA, which in turn, by phosphorylating other proteins can change their activity. The targets of PKA may be enzymes that are activated by phosphorylation, or they may be proteins that regulate transcription. The phosphorylation of a transcriptional activator, for example, may cause the activator to bind to a regulatory sequence on DNA and to increase the transcription of the gene it controls. The activation of previously inactive enzymes alters the state of the cell by changing the reactions that are occurring within the cell.

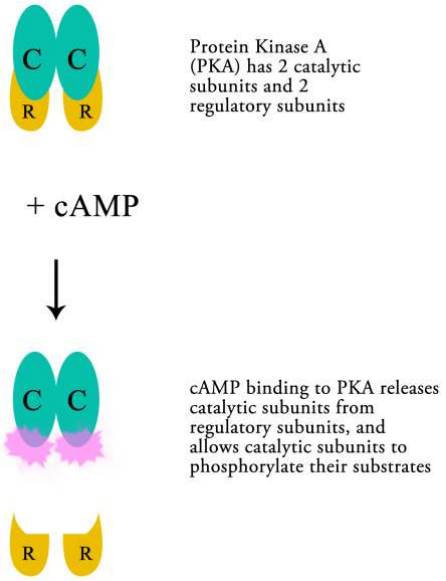


Figure 8.4.8: *Protein Kinase A Activation*

For example, the binding of epinephrine to its receptor on the cell surface, activates, through the action of G-proteins, and subsequent activation of PKA, the phosphorylation of glycogen phosphorylase. The resulting activation of glycogen phosphorylase leads to the breakdown of glycogen, releasing glucose (in the form of glucose-1-phosphate) for use by the cell. Changes in gene expression, likewise, lead to changes in the cell by altering the production of particular proteins in response to the signal.

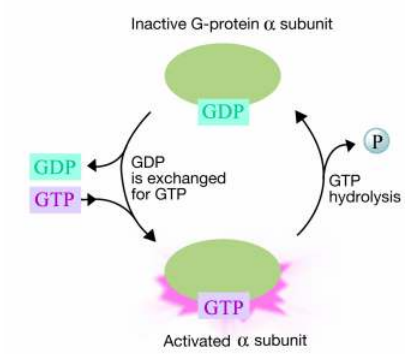


Figure 8.4.9: *G-protein nucleotide swapping*

Although the steps described above seem complicated, they follow the simple pattern outlined at the beginning of this section:

- Binding of signal to receptor
- Several steps where the signal is passed on through intermediate molecules (G-proteins, adenylate cyclase, cAMP, and finally, PKA)
- Phosphorylation of target proteins by the kinase, leading to changes in the cell.

Finally, if the signal binding to the receptor serves as a switch that sets these events in motion, there must be mechanisms to turn the pathway off. The first is at the level of the G-protein. Recall that the alpha subunit of the G-protein is in its free and activated state when it has GTP bound and that it associates with the beta- gamma subunits and has a GDP bound when it is inactive. We also know that the alpha subunit has an activity that enables it to hydrolyze GTP to GDP, as shown in the figure above left. This GTP-hydrolyzing activity makes it possible for the alpha subunit, once it has completed its task, to return to its GDP bound state, re-associate with the beta-gamma part and become inactive again.

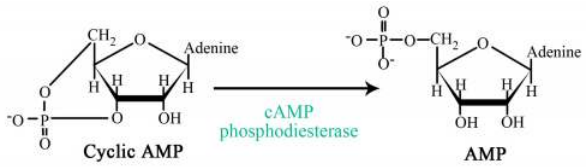


Figure 8.4.10: *cAMP Breakdown*

The second "off switch" is further down the signaling pathway, and controls the level of cAMP. We just noted that cAMP levels increase when adenylate cyclase is activated. When its job is done, cAMP is broken down by an enzyme called phosphodiesterase. When cAMP levels drop, PKA returns to its inactive state, putting a halt to the changes brought about by the activation of adenylate cyclase by an activated G-protein.

Let us now examine the events that follow the activation of Phospholipase C (PLC) by a G-protein. As we noted earlier, the activation of PLC results in the production of the second messengers IP3 and DAG. What do these molecules do?

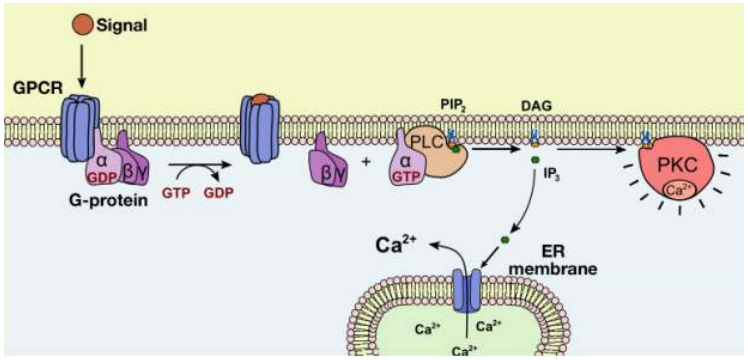


Figure 8.4.11: *Phospholipase C Signaling*

The IP3 and DAG produced by activated phospholipase C work together to activate a protein kinase. First, IP3 diffuses to the endoplasmic reticulum membrane where it binds to gated calcium ion channels. This causes calcium channels in the ER membrane to open and release large amounts of calcium into the cytoplasm from the ER lumen, as shown in the figure below.

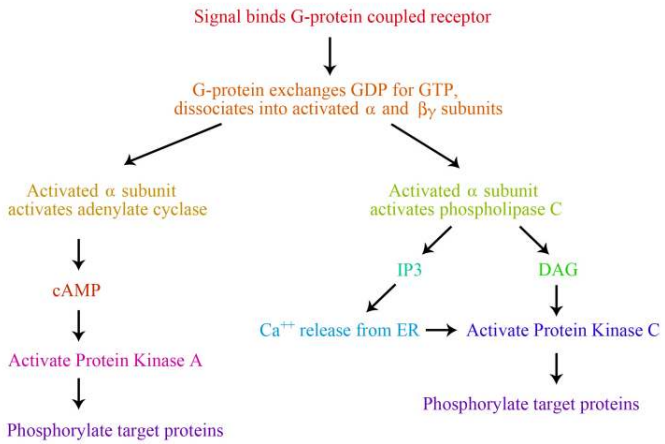


Figure 8.4.12: *Signaling Outcomes*

The increase in cytosolic calcium ion concentration has various effects, one of which is to activate a protein kinase called protein kinase C (C for calcium), together with the DAG made in the earlier step. Like PKA, Protein kinase C phosphorylates a variety of proteins in the cell, altering their activity and thus changing the state of the cell.

The pathways leading to PKC and PKA activation following the binding of a signal to a GPCR are summarized in Figure 8.4.12.

This page titled 13.4: G-protein Coupled Receptors (GPCRs) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **8.4: G-protein Coupled Receptors (GPCRs)** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

13.5: Receptor Tyrosine Kinases (RTKs)

Receptor tyrosine kinases mediate responses to a large number of signals, including peptide hormones like insulin and growth factors like epidermal growth factor. Like the GPCRs, receptor tyrosine kinases bind a signal, then pass the message on through a series of intracellular molecules, the last of which acts on target proteins to change the state of the cell.

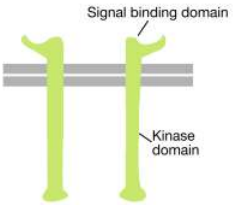


Figure 8.5.1: *Receptor Tyrosine Kinase before signal binding*

As the name suggests, a receptor tyrosine kinase is a cell surface receptor that also has a tyrosine kinase activity. The signal binding domain of the receptor tyrosine kinase is on the cell surface, while the tyrosine kinase enzymatic activity resides in the cytoplasmic part of the protein (see figure above). A transmembrane alpha helix connects these two regions of the receptor.

What happens when signal molecules bind to receptor tyrosine kinases?

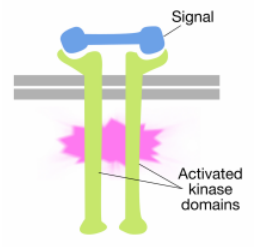


Figure 8.5.2: *Signal binding causes dimerization of receptor and activation of tyrosine kinase domains*

Binding of signal molecules to the extracellular domains of receptor tyrosine kinase molecules causes two receptor molecules to dimerize (come together and associate). This brings the cytoplasmic tails of the receptors close to each other and causes the tyrosine kinase activity of these tails to be turned on. The activated tails then phosphorylate each other on several tyrosine residues. This is called autophosphorylation.

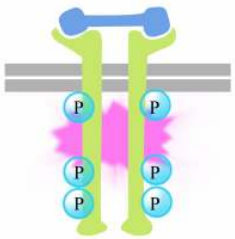


Figure 8.5.3: *Activated tyrosine kinase domains add phosphate onto each other*

The phosphorylation of tyrosines on the receptor tails triggers the assembly of an intracellular signaling complex on the tails. The newly phosphorylated tyrosines serve as binding sites for signaling proteins that then pass the message on to yet other proteins. An important protein that is subsequently activated by the signaling complexes on the receptor tyrosine kinases is called *Ras*.

The Ras protein is a monomeric guanine nucleotide binding protein that is associated with the cytosolic face of the plasma membrane (in fact, it is a lot like the alpha subunit of trimeric G-proteins). Just like the alpha subunit of a G- protein, Ras is active when GTP is bound to it and inactive when GDP is bound to it.Also, like the alpha subunit, Ras can hydrolyze the GTP to GDP.

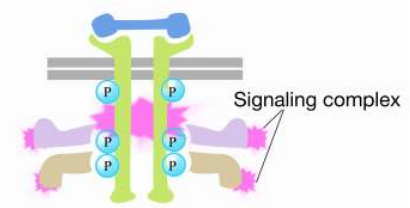


Figure 8.5.4: *Complex of signaling proteins assembles on phosphorylated RTK tails. This complex can activate Ras.*

When a signal arrives at the receptor tyrosine kinase, the receptor monomers come together and phosphorylate each others' tyrosines, triggering the assembly of a complex of proteins on the cytoplasmic tail of the receptor. One of the proteins in this complex interacts with Ras and stimulates the exchange of the GDP bound to the inactive Ras for a GTP. This activates the Ras.

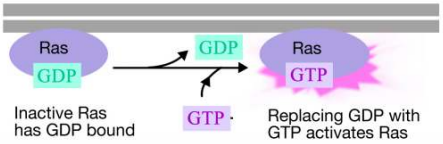


Figure 8.5.5: *Ras Activation*

Activated Ras triggers a phosphorylation cascade of three protein kinases, which relay and distribute the signal. These protein kinases are members of a group called the MAP kinases (Mitogen Activated Protein Kinases). The final kinase in this cascade phosphorylates various target proteins, including enzymes and transcriptional activators that regulate gene expression.

The phosphorylation of various enzymes can alter their activities, and set off new chemical reactions in the cell, while the phosphorylation of transcriptional activators can change which genes are expressed. The combined effect of changes in gene expression and protein activity alter the cell's physiological state.

Once again, in following the path of signal transduction mediated by RTKs, it is possible to discern the same basic pattern of events: a signal is bound by the extracellular domains of receptor tyrosine kinases, resulting in receptor dimerization and autophosphorylation of the cytosolic tails, thus conveying the message to the interior of the cell.

The message is passed on via a signalling complex to Ras which then stimulates a series of kinases. The terminal kinase in the cascade acts on target proteins and brings about in changes in protein activities and gene expression.

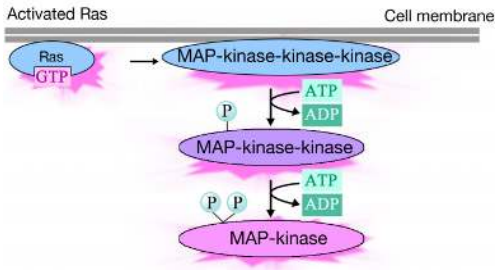


Figure 8.5.6: *Activated Ras Cascade*

The descriptions above provide a very simple sketch of some of the major classes of receptors and deal primarily with the mechanistic details of the steps by which signals received by various types of receptors bring about changes in cells. A major take-home lesson is the essential similarity of the different pathways.

Another point to keep in mind is that while we have looked at each individual pathway in isolation, a cell, at any given time receives multiple signals that set off a variety of different responses at once. The pathways described above show a considerable degree of "cross-talk" and the response to any given signal is affected by the other signals that the cell receives simultaneously. The multitude of different receptors, signals and the combinations thereof are the means by which cells are able to respond to an enormous variety of different circumstances.

This page titled 13.5: Receptor Tyrosine Kinases (RTKs) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal via source content that was edited to the style and standards of the LibreTexts platform.

- **8.5: Receptor Tyrosine Kinases (RTKs)** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

13.6: Gated Ion Channels - Neural Signaling

Learning Goals (ChatGPT o3-mini)

1. **Understand the Fundamental Role of Ion Channels in Neural Signaling:**
   - Explain how ion channels facilitate the movement of ions across the hydrophobic membrane.
   - Differentiate between constitutively open (leakage) channels and gated channels that open or close in response to stimuli such as ligands, changes in voltage, or mechanical forces.
2. **Describe the Structure and Functional Organization of Neurons:**
   - Identify the key structural components of neurons (cell body, dendrites, axon, synapses) and their roles in signal transmission.
   - Compare and contrast neuromuscular junctions with central nervous system synapses, emphasizing differences in neurotransmitter release and receptor activation.
3. **Analyze the Biophysical Basis of Membrane Potential Generation and Maintenance:**
   - Describe how ion gradients, created and maintained by ion pumps (e.g., Na^+/K^+-ATPase) and selective permeability through channels, contribute to the resting membrane potential.
   - Interpret how differential ion permeabilities (as demonstrated by synthetic vesicle studies versus biological membranes) influence the electrochemical potential across cell membranes.
4. **Explain the Mechanisms Underlying the Action Potential:**
   - Outline the sequential changes in ion permeability (Na^+, K^+, Cl^-) that produce the phases of depolarization, repolarization, and hyperpolarization in an action potential.
   - Evaluate the impact of voltage-gated ion channels in generating rapid electrical signals essential for neuronal communication.
5. **Differentiate Between Ionotropic and Metabotropic Neurotransmitter Receptors:**
   - Define ionotropic receptors and describe how they directly gate ion flow upon neurotransmitter binding, resulting in rapid synaptic transmission.
   - Define metabotropic receptors and discuss how they indirectly modulate ion channel activity via G-protein-coupled second messenger cascades, leading to slower, modulatory effects.
6. **Assess the Impact of Neurotoxins on Ion Channel Function:**
   - Identify key neurotoxins (e.g., tetrodotoxin and saxitoxin) and explain their modes of action in blocking specific ion channels.
   - Discuss how these toxins help reveal the structural features and selectivity mechanisms of ion channels.
7. **Evaluate How Neurotransmitter Binding Influences Neuronal Excitability:**
   - Explain how binding of excitatory neurotransmitters (e.g., acetylcholine, glutamate) and inhibitory neurotransmitters (e.g., GABA) alters membrane potential.
   - Describe how changes in membrane potential can trigger or prevent the generation of an action potential.
8. **Integrate the Roles of Ion Channels and Neurotransmitter Receptors in Synaptic Transmission:**
   - Discuss how the combined activities of ligand-gated and voltage-gated channels determine the overall synaptic response.
   - Explain the importance of synaptic integration in neural networks, including how single neurons can process multiple excitatory and inhibitory signals.
9. **Relate Neural Signaling to Higher-Level Brain Functions:**
   - Appreciate the complexity of the human brain by relating the molecular mechanisms of neural signaling to processes such as memory, cognition, and behavior.
   - Consider how alterations in ion channel function and neurotransmitter signaling can contribute to neurological disorders.
10. **Apply Quantitative and Structural Approaches to Neural Signaling:**
    - Interpret experimental data and computational models (e.g., electrophysiological recordings, ion permeability measurements) to explain how ion fluxes generate membrane potentials.
    - Use structural models of ion channels and receptors to understand their function and specificity in neural signaling.

These learning goals will help students build a comprehensive understanding of the molecular and biophysical underpinnings of neural communication and its significance in health and disease.

13.6.1: Introduction to Neural Signaling

Ion channels are membrane proteins that allow the flow of normally impermeant ions across the hydrophobic "sea" of cell or organelle membranes. The channel proteins can be constitutively open or "gated" open (or closed) by signals that affect the conformation of the protein. The signals can be ligands

(in the form of hormones or neurotransmitters), post-translational modifications (mostly phosphorylation, dephosphorylation), or more difficult-to-understand events (change in transmembrane potential, pressure, or temperature). Gated channels are perhaps best studied through their role in neural signaling.

Neurochemistry is one of the most explosive areas of biological research. Scientists are unraveling the molecular bases for memory, cognition, emotion, and behavior. The next decades will bring a greater understanding of brain chemistry, along with it the potential to alter human mood and memory, and to treat mental illnesses such as schizophrenia much more effectively. The human brain has about 86 billion neurons. Compare this with estimates for the number of stars in the Milky Way galaxy (about 200 billion, derived from luminosity and mass measurements, but some estimates suggest 400 billion). Now imagine that each neuron can form connections - synapses - with 1000 to 10,000 other neurons. Throw in another 86 billion nonneuronal cells, like glial cells, and you have one of the most complex structures in the universe. A recent high-resolution map of the cerebral cortex shows that a single cubic millimeter of the brain contains 57,000 cells, about 150 million synapses (connections between cells), and about 0.2 meters of blood vessels. This section will explore the biology and chemistry of neurons.

Recent studies suggest around 3300 different types of cells in the brain. We will discuss two kinds of neurons - those that interact with muscles at the neuromuscular junction and those that interact with other neurons in the central nervous system. Neurons consist of a single, nucleated cell body with multiple signal-receiving outgrowths (dendrites, for receiving electrical impulses, short) and multiple signal-sending outgrowths (axons, single projection for sending electrical impulses, long and myelinated for insulation) which end in a terminal button.

These interact through the synapse with dendrites on other neurons. These characteristics are shown in Figure 13.6.1.

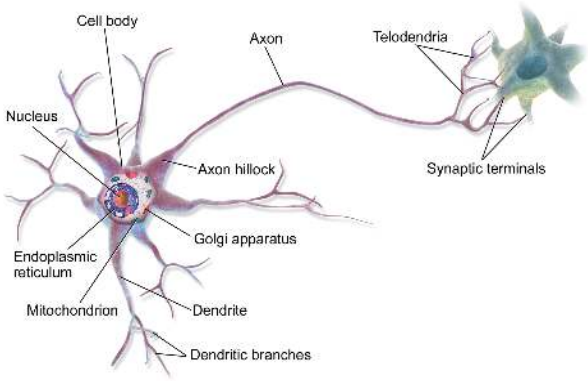


Figure 13.6.1: A neuron. Bruce Blaus - Own work, CC BY 3.0, https://commons.wikimedia.org/w/inde...curid=28761830

The high-resolution map of the human cortex shows that an axon most commonly (96%) forms one synapse with a target cell. 3% of the axons form 2 synapses with the target cell. However, neurons that had lots of input onto their dendrites (they were highly innervated with more than 3000 synapses to their dendrites) could form 7 or more synapses from their axon to a target cell.

A presynaptic neuron can stimulate an adjacent postsynaptic neuron by releasing a neurotransmitter into the synapse between the cells, which binds to a receptor in the membrane of the postsynaptic cell, stimulating the cell, as shown in Figure 13.6.2.

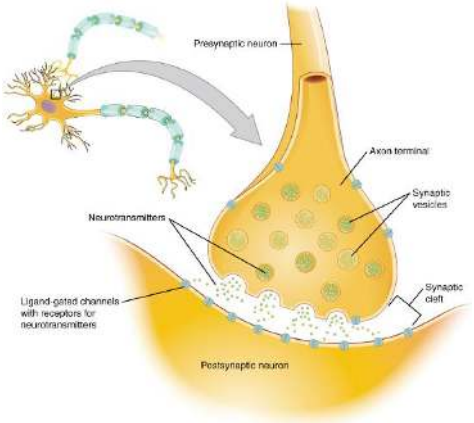


Figure 13.6.2: The synapse. https://cnx.org/contents/FPtK1zmh@8.25:fEI3C8Ot@10/Preface..Creative Commons Attribution 4.0 International license.

We will discuss the events that cause the post-synaptic cell to "fire," but we will not discuss the immediate events that lead to the release of neurotransmitters by the presynaptic neuron.

Neurons (like all cells) have a transmembrane potential across the membrane. Transient changes in the membrane potential are associated with neuron activation or inhibition. This arises partly due to the imbalance of ions across the membrane, which the Na^+-K^+-ATPase establishes in the membrane. This

electrogenic antiporter transfers 3 Na ions out of the cytoplasm for every 2 K ions it transports in. Likewise, Cl^- has a much higher level outside the cell. Membrane potentials are determined by the size of the ion gradients across the membrane and by the differential permeability of membranes to ions. As we saw previously, synthetic bilayer membranes are not very permeable to ions, as shown in Table **13.6.1** below.

Ion permeability of phosphatidyl serine vesicles

ION	PERMEABILITY (cm/s)
sodium	$<1.6 \times 10^{-13}$ (lowest)
potassium	$<9 \times 10^{-13}$
chloride	$<1.5 \times 10^{-11}$ (highest)

Table **13.6.1**: Ion permeability of phosphatidyl serine vesicles

Sodium would be expected to have a lower permeability than potassium since it has a higher charge density. It is also the largest in effective size due to the larger hydration sphere around the ion arising from its higher charge density. Chloride would be expected to have the highest permeability since it has the lowest charge density (due to the repulsion of the electron cloud in the negative ion). Intracellular charged proteins (which are mostly negative) are not permeable and help create the negative charge imbalance across the membrane.

Much work has been done on the giant axon of the squid, which has uniquely high intracellular potassium (400 mM compared to 20 mM outside) and high extracellular fluid sodium (440 mM vs 50 mM inside) and chloride (560 mM vs 52 mM inside). Mammalian cell concentrations are much lower, but the relative size of the gradient is about the same. Typical ion concentrations and permeabilities for mammalian membranes are shown in Table **13.6.2** below.

T

Ion	Cell (mM)	Blood (mM)	Permeability (cm/s)
potassium	140	5	5×10^{-7}
sodium	5-15	145	5×10^{-9}
chloride	4	110	1×10^{-8}
X^- (negative macromolecules)	138	9	0

Table **13.6.2**: Typical ion concentrations and permeabilities for mammalian membranes

How can we account for the markedly greater permeabilities of ions (1000x to 1,000,000x) in mammalian cell membranes compared to synthetic lipid vesicles? Previously, we showed that glucose has a greater permeability through red blood cell membranes than synthetic liposomes because of a membrane receptor allowing facilitated diffusion across the membrane and down a concentration gradient. The same thing is true of ion permeabilities in intact biological membranes. These membranes have different selective ion channels (non-gated - always open, and gated - open only after specific conformational changes). The non-gated channels dramatically increase the permeability of membranes to ions, as the glucose transport protein increases the permeability to glucose. It turns out that this differential permeability contributes to the transmembrane potential. Ion channels in nerves and muscles can move ions across the membrane at a rate of up to 10^9/s, comparable to k_{cat} for the best enzymes.

If we envision channels as pores, how can we account for the channel's selectivity to specific ions? A larger pore should admit any ion less than a maximal size for the pore, so it is hard to imagine the nature of the selectivity filter. Because of this, many people discounted the ideas of channels in favor of a transporter, which would bind the ion selectivity and then, through conformational changes, move the ion across (much like the Na/K ATPase we discussed in the previous guide). This model could not, however, account for the incredible rates of ion flux across the membrane. Selectivity can be accounted for by a channel that contains a narrow opening that acts as an ion sieve. The ion loses most of its hydration sphere and forms specific interactions with amino acid side chains in the pore region. Such an interaction would be transient and not too tight since the ion must pass through the membrane. As we will see later, these ion channels:

- pass ions down a concentration gradient in a thermodynamically favorable process
- are specific for certain ions (although a few are less selective and will pass Na, K, Ca, and Mg ions)
- allow ion flow through either ungated or gated channels.,
- saturate with increasing ion concentration (even though concentration increases, the ions have a greater thermodynamic drive to pass through the channel). This is consistent with the ion binding at a selectivity filter in the narrow part of the pore. The K_D for the interaction is usually in the mM range and indicates weak binding with large dissociation rate constants (k_{off}).

13.6.2: Transmembrane Potentials

Several questions arise about ions' distribution and the magnitude of the transmembrane potential.

1. How are the ion gradients established?
2. How does the transmembrane ion distribution contribute to the membrane potential?
3. How can the resting electrochemical potential and the ion distribution be maintained?

The answer to these questions will be illustrated using studies on two types of brain cells, glial cells (which function as protectors, scavengers, and feeders for brain neurons) and neurons. Both types of cells have transmembrane potentials.

Glial Cells

Different mechanisms can establish the transmembrane ion gradients for ions. One uses ion-specific ATPases (P-type ion transporters), as we discussed with the Na/K ATPase. This transporter ejects three sodium ions from inside the cell for every two potassium ions it transports in, all against a concentration gradient. Since it is an electrogenic antiporter, it helps generate the potential. Specific ion channels also contribute (as described below) to the transmembrane gradients and potentials.

The harder question is how the ion distribution contributes to the membrane potential. Two things must occur for a membrane potential to exist: First, there must be a concentration gradient of charged ions (for example, sodium, potassium, or chloride) across the membrane. Second, the membrane must be differentially permeable to different ions. If the membrane were entirely impermeable to ions, then no movement of ions across the membrane could occur, and no membrane potential would arise. If membranes are differentially permeable to the ions, an electrical potential across the membrane can arise. Remember, synthetic bilayers are quite impermeable to ions, given the hydrophobicity of the internal part of the bilayer. Likewise, it is quite impermeable to glucose. It turns out that glial cells appear to have only a non-gated potassium channel, which allows the outward flow of potassium ions down the concentration gradient. The inside will then have a net negative charge since impermeable anions remain.

The chemical potential gradient causes this outward flow of potassium ions. As more ions leave, the inside gets more negative, and a transmembrane potential develops, which resists further efflux of potassium. Eventually, they balance, and the net efflux of potassium stops. The resting transmembrane potential reaches around -75 mV, the value obtained from the equations we will derive below. Since glial cells appear only to express a non-gated (or leakage) potassium channel, their resting potential equals the potassium equilibrium potential. Figure **13.6.3** shows hypothetical transmembrane potentials (Ψ), and the electrical and chemical potentials (ΔG) for K^+ loaded vesicles with and without a non-gated K^+ channel

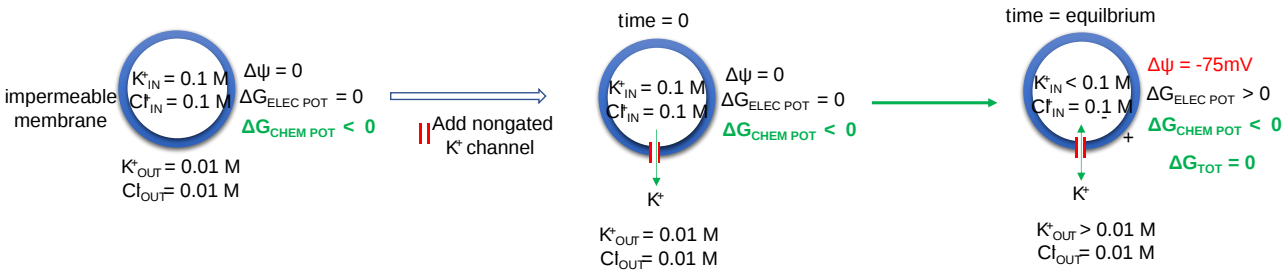


Figure **13.6.3**: Visualizing the transmembrane potential in K^+ loaded vesicles with and without a non-gated K^+ channel

As shown previously, very little K^+ efflux is required to develop a transmembrane potential. At equilibrium, the K^+_{IN} is simply shown as < 0.1 M while K^+_{OUT} is shown as > 0.1 M

A PhET simulation shows the differences between non-gated (leakage, always open) and gated channels (opened in response to stimuli like ligands, transmembrane potential, mechanical forces, etc).

13.6.3: Neurotransmitter Activation of Neurons

What happens when a neurotransmitter binds to a receptor on the postsynaptic cell? We will study two examples. The first is the simplest: binding the neurotransmitter acetylcholine, released by a motor neuron, to its receptor on the muscle. This region is called the neuromuscular junction. Acetylcholine binding will lead to a transient depolarization of the muscle cell. Next, we will discuss the interaction of a neurotransmitter with a postsynaptic neuron in the central nervous system. This is a much more complex system. Their differences are described below:

In neurons interacting with muscles:

- Most muscle fibers are innervated by only one neuron, a motor neuron
- Neurotransmitter release at the neuromuscular junction leads only to muscle excitation, not inhibition.
- All fibers are excited by the same neurotransmitter, acetylcholine.

In the central nervous system, life is more complicated:

- Stimuli are received from hundreds to thousands of different neurons.
- Nerves receive both excitatory and inhibitory stimuli from neurotransmitters.
- Different kinds of receptors are present to receive stimuli, which control the activity of different channels.
- Neuronal ion channels are gated by various mechanisms in addition to changes in membrane potential, including gating by heat, cold, stretch, or covalent modification.
- Most nerve cells have a resting potential of about -65 mV compared to -90 mV for a muscle cell.

What happens when a neurotransmitter binds to the receptor on the postsynaptic cell? A depolarization occurs (mediated by conformational changes in the transmitter-receptor complex), raising the membrane potential from the resting equilibrium level. What happens next depends on the identity of the post-synaptic cell. In the muscle cell, the rising potential caused by acetylcholine binding ultimately leads to muscle contraction by opening the intracellular organelle membrane calcium channels. In a neuron, the rising potential triggers an action potential by opening voltage-gated sodium channels. The potential rises to about + 35 mV but does not reach the Na^+ ion equilibrium potential, because the high positive potential opens a voltage-gated potassium

channel. The potential then falls until it reaches the K^+ ion equilibrium potential when the cells are hyperpolarized. It slowly then relaxes back to the resting potential of -60 mV. This wave of changes in potential sweeps down the post-synaptic cell membrane and is the basis for the "firing" of the neuron. A plot of transmembrane voltage changes vs time for a typical action potential is shown in Figure 13.6.4.

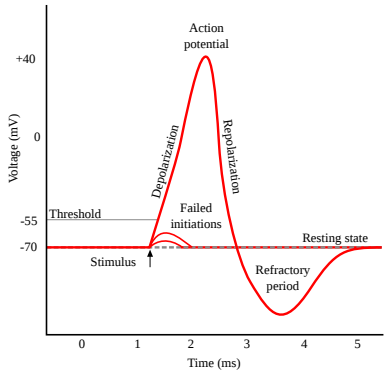
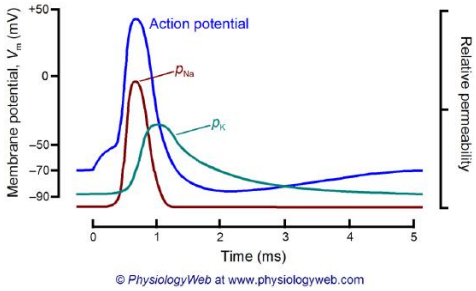


Figure 13.6.4: Action potential of a neuron. Original by en:User:Chris 73, updated by en:User:Diberri, **GNU Free Documentation License**,

Figure 13.6.5 shows the actual changes in ion permeability in various phases of the action potential (replace figure, seek permission)



mine

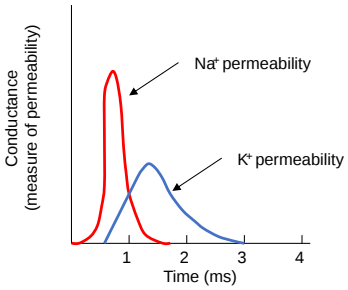
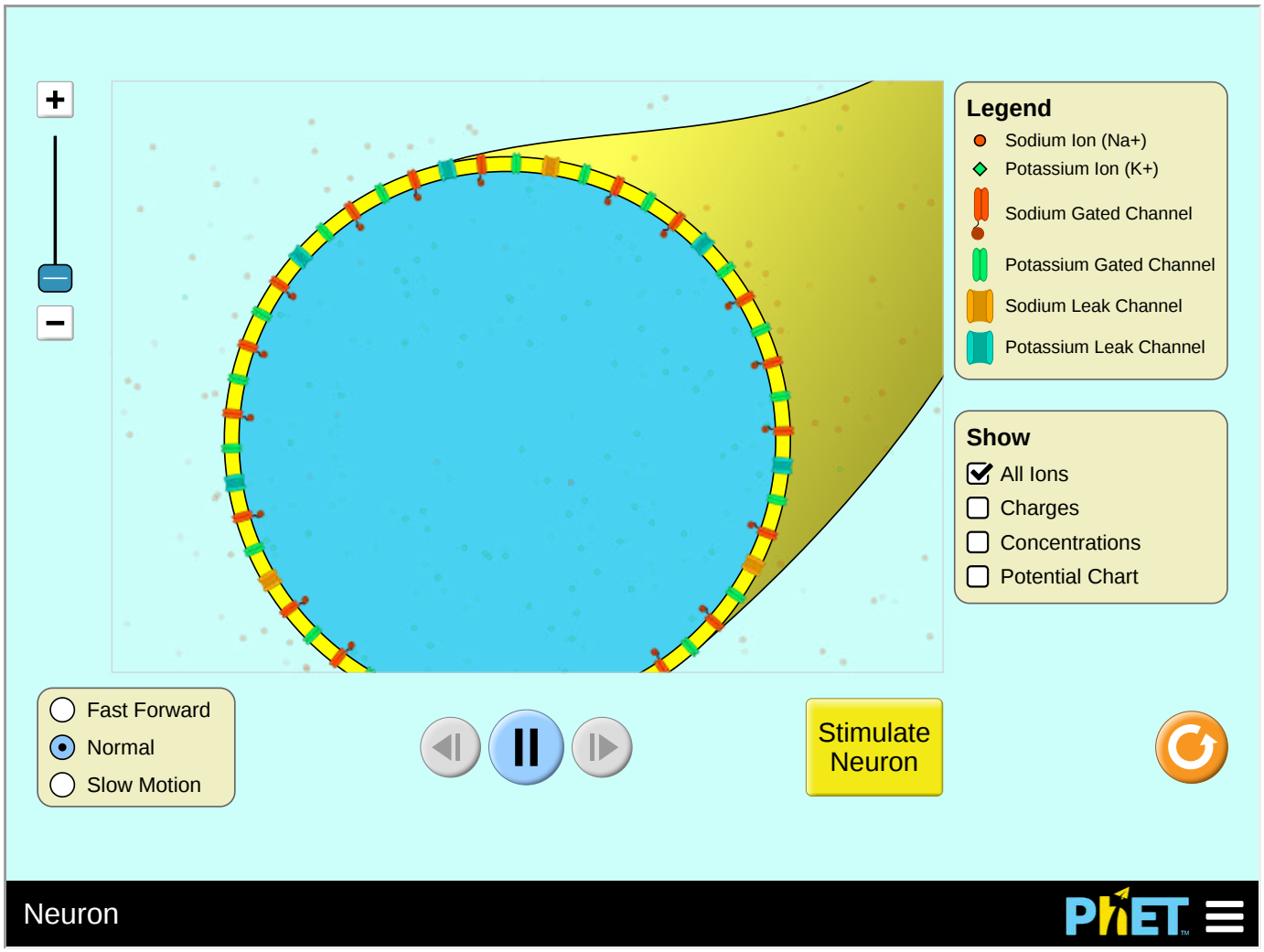


Figure 13.6.5: Na^+ and K^+ permeability during the action potential

Figure 13.6.6 shows an animation of a neuron firing, showing all the key players. (produced by PhET, University of Colorado, Boulder).



13.6.4: Proteins of the Neural Synapse

We must now account for the rise and fall of the membrane potential to a variety of neurotransmitters, including the cholinergic transmitters (ex., acetylcholine), catecholamines (dopamine, epinephrine, norepinephrine), amino acid derivatives (ex., Glu, Asp, N-methyl-D-Asp, Gly, gamma-amino-butyric acid -GABA), and peptides (endorphins, enkephalins). We will consider five membrane proteins as shown below in Figure **13.6.7**.

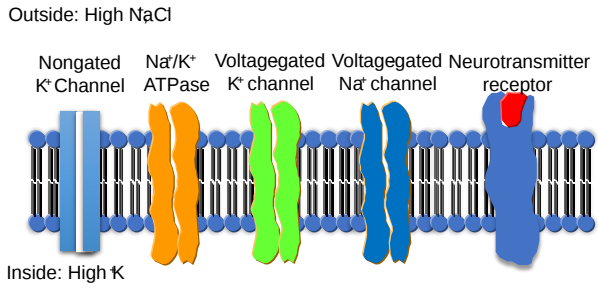


Figure **13.6.7**: 5 key membrane proteins involved in the action potential

We have explored many of these in detail in Chapter 11.3: Diffusion Across a Membrane—Channels, so in the following section, we will focus on just one new one, the chloride channel.

Na^+-K^+-ATPase: It transports sodium and potassium ions against a concentration gradient using ATP as an energy source. The protein is a sodium-dependent ATPase. The membrane potential could not be maintained without this protein since the sodium and potassium gradient would collapse. It also contributes to the potential since it is an electrogenic antiporter. (We have also seen that ungated potassium and sodium channels are present.)

Neurotransmitter receptor: The receptors we will consider here are typically ligand-gated ion channels. Once the ligand binds, a conformational change occurs in the protein, allowing a flow of ions down a concentration gradient. Depending on the nature of the ion, the channel either initiates depolarization

(when Na^+ enters from the outside and raises ΔΨ) or inhibits depolarization (when Cl^- enters from the outside and lowers ΔΨ). When chloride channels open, they hyperpolarize the transmembrane potential. Stimulatory neurotransmitters (like glutamate) lead to depolarization of the membrane, while inhibitory neurotransmitters (like gamma-aminobutyric acid) lead to hyperpolarization of the membrane (making the potential more negative).

Na^+ channel (voltage-gated): When the ligand-gated channel depolarizes the membrane to some threshold value, sodium channels undergo a conformational change and open, allowing Na^+ ions to flood into the cell, raising the potential to a positive value. 33 mV (a value lower than the equilibrium sodium potential). This membrane protein is a voltage-gated channel, not a ligand-gated one.

Two potent neurotoxins, tetrodotoxin (from Pufferfish) and saxitoxin, bind to the channel and act as antagonists (inhibit the activity of the receptor by blocking sodium influx). Their structures are shown in Figure **13.6.8**.

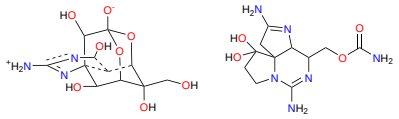


Figure **13.6.8**: Structures of tetrodotoxin and saxitoxin

The guanidino group of the tetrodotoxin appears to bind with high affinity to the channel's entrance that interacts with the hydrated sodium ion. Affinity chromatography using tetrodotoxin beads has been used to purify the protein. Figure **13.6.9** shows the relative sizes of ions used to probe size requirements for the channel. The hydrated K^+ ion can not pass through.

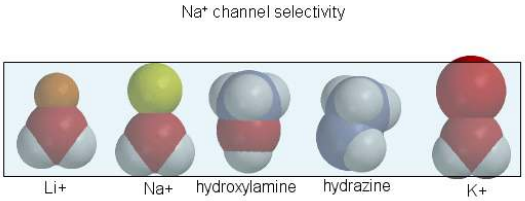


Figure **13.6.9**: Na^+ channel ion selectivity

Depolarization of the membrane potential may result in an outward movement and rotation of the positively charged helices containing Lys and Arg side chains, presumably forming salt bridges (ion-ion interactions) with negatively charged side chains within the protein. Depolarization of the membrane breaks a few salt bridges and effectively leads to the movement of 1 or 2 charges on the helix through the membrane. Work occurs when charges are moved through an electric field. Work is also related to the ΔG for the system, which is also dependent on the ratio of the open and closed forms of the channel. Other voltage-gated ion channels (for potassium and chloride) have a similar membrane topology and an S4 voltage-sensor helix.

K^+ channel *(voltage-gated)*: When the membrane potential reaches around +25 mV or so, the K+ channel, a voltage-gated membrane protein, alters its conformation, allowing K^+ efflux from the cells. This lowers the potential until it reaches the potassium equilibrium potential. It then slowly relaxes back to the cell resting potential of about -60 mV.

Cl^- channel: If these channels (typically ligand-gated) are open, they will hyperpolarize the cell and make it more difficult to fire.

The selectivity filter comprises many stacked rings of oxygens that can interact with a dehydrated K ion but not with a dehydrated Na^+ ion, which can not approach close enough to form significant interactions. Surrounding the filter are twelve aromatic amino acids that constrain the size of the pore opening. The interactions of the filter O's with the K ion make up for the energetically disfavored dehydration of the ion. The filter contains two K ions, which repel each other, assisting in the vectorial discharge of the ions through the membrane. These ions must form weak interactions with the selectivity filter. The actual pore is hydrophobic primarily, which facilitates the flow-through of the ions. The central cavity of the pore can hold some water molecules in addition to the K ions, which helps stabilize the ion in the middle of the pore.

13.6.5: Inhibitory Neurotransmitters - The GABA receptor

The main inhibitory neurotransmitters are GABA (gamma-aminobutyric acid), made from glutamic acid through decarboxylation of the α-C group, and glycine. They bind to transmitter-gated chloride channels, which, when open, hyperpolarize (make more negative) the transmembrane potential. Benzodiazepines (like Valium and Librium - anti-anxiety and muscle-relaxing agents) and barbiturates (like phenobarbital hypnotics) bind at allosteric sites on the GABA receptor and potentiate the binding of each other and GABA. This receptor is also affected by alcohol and anesthetics. Let's focus on the GABAA channels (also called the GABA(A) receptor or GABA(A)R) since they bind so many interesting pharmaceutical drugs.

Figure **13.6.10** shows an interactive iCn3D model of the human full-length heteromeric α1-β3-γ2L GABA(A)R in complex with picrotoxin, GABA, and megabody Mb38 (6huj) derived from cryo-EM (loads slowly).

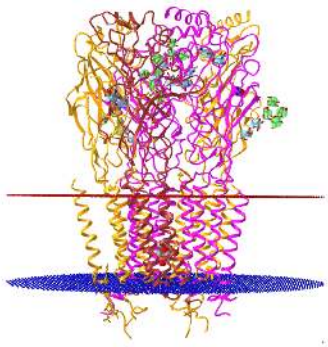


Figure 13.6.10: Human full-length heteromeric alpha1beta3gamma2L GABA(A)R in complex with picrotoxin, GABA and megabody Mb38 (6huj) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...VdExv4QhFPsVX8 (loads slowly)

This structure is a Type-A γ-aminobutyric ($GABA_A$) receptor. It consists of two alpha chains (orange), two beta chains (magenta), and one gamma chain (brown). Two GABAs (spacefill, CPK colors) are bound between the orange and magenta subunits (between the α and β chains). Picrotoxin is shown in spacefill CPK colors in the central pore (bottom), near the cytoplasmic end (blue bilayer dummy atoms). The extracellular domain (above the red) has glycans shown in SNFG cartoon form and shown in Figure 13.6.11.

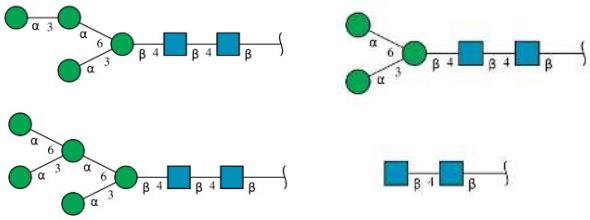
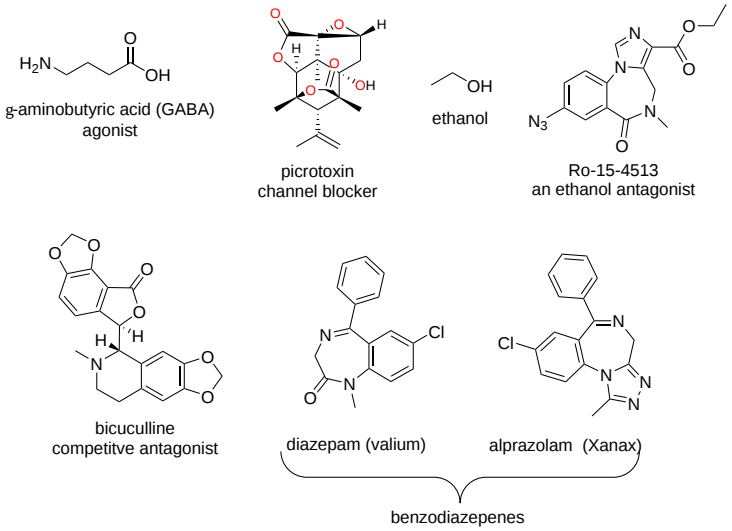


Figure 13.6.11: Glycan structure for the Type-A γ-aminobutyric ($GABA_A$) receptor

The structure of the receptor bound to many pharmacological agents has been solved. The structure of GABA (the agonist), bicuculline (a competitive antagonist), the benzodiazepines alprazolam (Valium) and diazepam (Xanax), a channel blocker (picrotoxin), ethanol, and Ro-15-4513 (an ethanol antagonist) are shown in Figure 13.6.12.



(\PageIndex{12}\): Structures of GABA-R binding molecules -GABA (the agonist), bicuculline (a competitive antagonist), the benzodiazepines alprazolam (Valium) and diazepam (Xanax), a channel blocker (picrotoxin), ethanol, and Ro-15-4513 (an ethanol antagonist)

The benzodiazepine diazepam (Xanax) binds at an allosteric site and promotes GABA binding, so it is considered an "**indirect agonist**". It is also an anxiolytic (a drug used to reduce anxiety) and an anticonvulsant. Ethanol activates the inhibitory GABA channel. Too much ethanol leads to hyperpolarized cells, so neural responses are greatly inhibited, possibly leading to death. Ethanol acts synergistically with benzodiazepines, which makes this combination so lethal.

Figure **13.6.13** shows an interactive iCn3D model of the human full-length heteromeric α1-β3-γ2L GABA(A)R in complex with diazepam (Valium), GABA (6HUP) (loads slowly). https://structure.ncbi.nlm.nih.gov/i...fTZMqePnpdwTy8

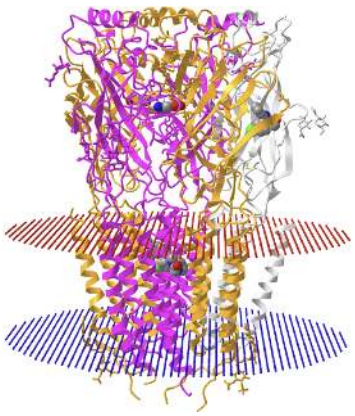


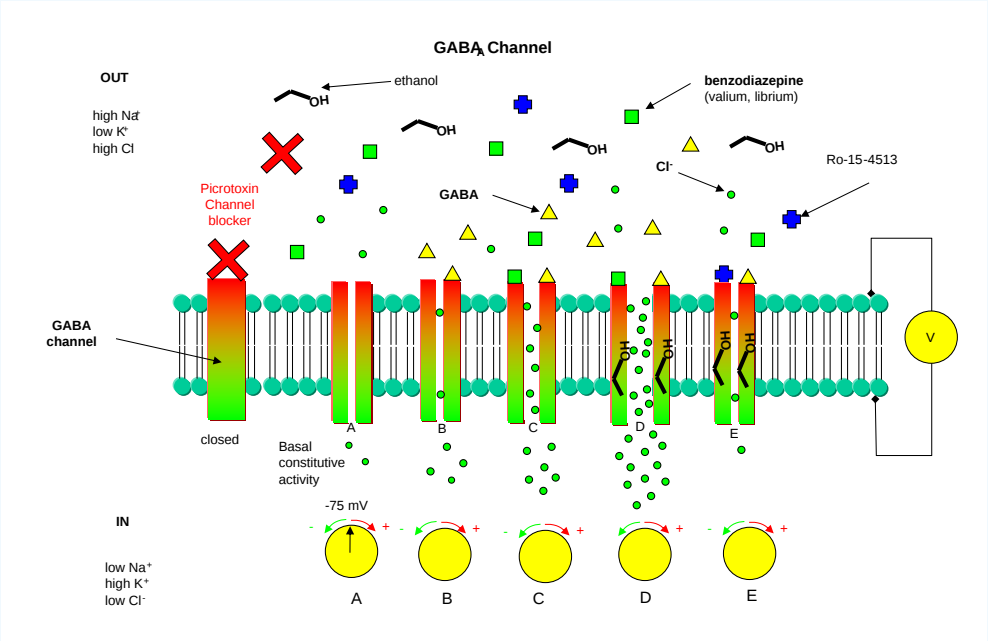
Figure **13.6.13**: Human full-length heteromeric α1-β3-γ2L GABA(A)R in complex with diazepam (Valium), GABA (6HUP). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...fTZMqePnpdwTy8

The structure shows two alpha chains (orange), two beta chains (magenta), and one gamma chain (gray). Two GABAs (spacefill, CPK colors) are bound between the orange and magenta subunits (between the α and β chains). Three Valiums (spacefill, CPK) are shown in the structure. One is bound to an allosteric site between the orange (alpha) and gray (gamma) chains. Two others are bound within the bilayer, this time between the alpha (orange) and beta chains (magenta). Different benzodiazepines appear to bind in slightly different sites, and binding elicits their specific effects.

A drug, Ro-15-4513, was developed in the 1980s that antagonizes the effect of ethanol. It has a complicated pharmacology. It is considered a benzodiazepine "**partial inverse agonist**". It has no effect by itself. It reverses the anticonvulsant effects of benzodiazepines and blocks the Cl^- effects of ethanol, so it is an inverse agonist of GABA-mediated Cl^- flux. Hence, it antagonizes the effect of benzodiazepines and alcohol. If given to intoxicated mice, they act normally! Such drugs have not reached the market as their use poses significant ethical issues.

The effect of pharmacological agents on the GABA receptor

In the "voltmeters" below, draw an arrow indicating if the transmembrane potential becomes more negative or more positive for the conditions given.



Answer

Add texts here. Do not delete this text first.

13.6.6: Metabotropic Neural Receptors

Some signaling molecules, whose effects are regulated by kinases (β-adrenergic and some olfactory signals by PKA, and acetylcholine by PKC, for example), are neurotransmitters. In all the previous examples, neurotransmitters directly gate inactive ion channels. These types of membrane receptors are classified as **ionotropic**. Typical examples of neurotransmitter-gated ion channels are the acetylcholine receptor in neuromuscular junctions and the central nervous system's Glu, Gly, and GABA receptors. These receptors are multimeric proteins. Receptors with direct gating of ion flow are fast, with activities that last milliseconds, and are used in eliciting behavioral responses.

However, ion channels can also be gated indirectly when the neurotransmitter binds to its receptor and leads to downstream events, subsequently opening an ion channel distinct from the receptor. In this case, the occupied receptor indirectly communicates to an ion channel through a G protein, for example. Examples of this indirect gating of ion channels include the serotonin, adrenergic, and dopamine receptors in the brain. These receptors are classic single-protein serpentine GPR receptors with seven transmembrane helices and intracellular domains that can interact with G proteins as described above to increase second messenger levels (cAMP, DAG) in the neuron. The receptors are classified as **metabotropic** since they must activate a series of metabolic steps before ion channels are opened. The second messengers can either activate kinases in the cell, which phosphorylate ion channels to open or close them, or bind directly to the channel and modulate its activity through an allosteric conformational change. In some cases, the G protein directly acts on the ion channel. These different ways are described in Figure 13.6.14.

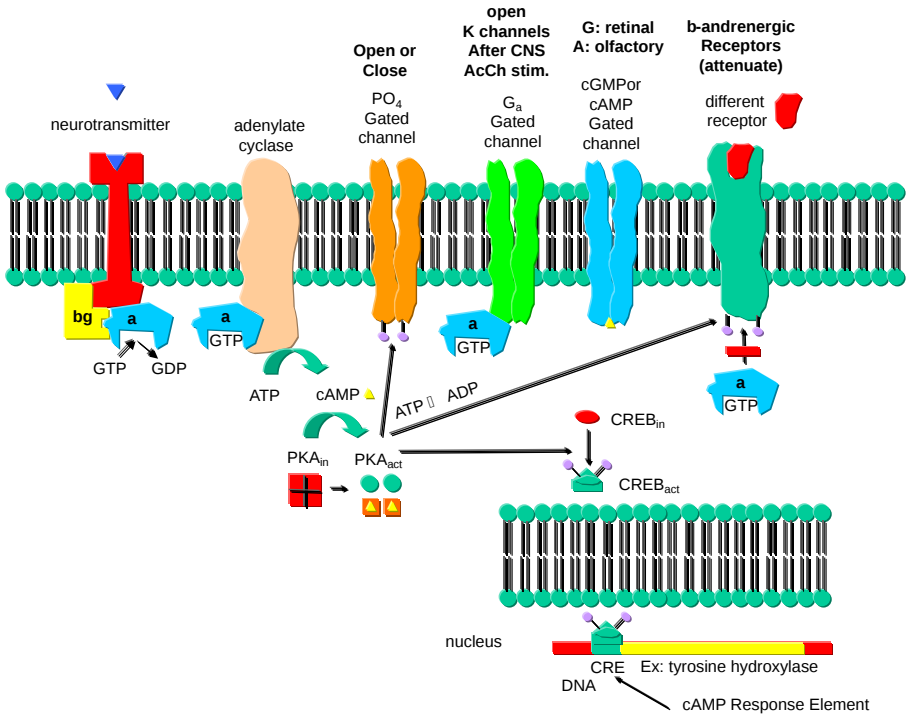


Figure 13.6.14: Neurotransmission: Gating through G Linked Receptors

In contrast to direct gating, receptors that indirectly gate ion channels have slow activities that last seconds to minutes. These receptors are usually involved in modulating behavior by changing the excitability of neurons and the strength of neural connections, hence modulating learning and memory. These changes can occur in many ways, summarized below:

Phosphorylating ion channels: Receptors that act through a second messenger system can change ion channel activity by activating kinases that phosphorylate the channels. This may:

- open the channel normally closed at the resting potential, producing an effect like gating.
- Close a channel that usually opens at the resting potential (such as non-gated K channels, which, when closed, would depolarize the cell and make it more excitable).

Gα interaction with ion channels:

- the Ga subunit of the G protein interacts with K channels after stimulation of the CNS Acetylcholine receptor, opening the channel and hyperpolarizing the cell

Second messenger interaction with ion channels:

- cGMP opens cation channels in retinal cells after activation of the photoreceptor by photons
- cAMP opens cation channels in olfactory cells after activation of the olfactory receptor by odorants.

Second messenger effects on proteins other than ion channels (usually different receptors):

- the β-adrenergic receptors are phosphorylated by PKA and PKC (activated by stimulation of a different neurotransmitter receptor linked through a G protein to produce increased levels of second messengers cAMP and diacylglycerol). When phosphorylated, the β-adrenergic receptor, itself activated through G protein) can't bind Gs. This attenuates the response of the β-adrenergic receptor to its neurotransmitter, which leads to desensitization to that signal.

Second messengers regulate gene expression:

- cAMP-activated PKA can phosphorylate an inactive transcription factor in the cell, which then can bind to a section of DNA called the cAMP Response Element (CRE), which is upstream of specific genes, leading to the transcription of the genes. The transcription factor is called CREB for cAMP Response Element Binding protein. Example: Tyrosine hydroxylase (a monooxygenase) is a key enzyme in synthesizing epinephrine and norepinephrine. The activity of this protein is increased when PKA phosphorylates it. Hence, this modification of the already present protein can increase its activity quickly. If an animal is subjected to severe or long-term stress (cold or immobilization), presynaptic cells with norepinephrine will be stimulated to release the neurotransmitter. This requires continual synthesis of the neurotransmitter by the presynaptic cell. The increase in the synthesis of this neurotransmitter is caused by the presynaptic cell being stimulated by another neuron, which leads to increased levels of cAMP, and ultimately activation of CREB, which increases transcription of the hydroxylase gene.

Caffeine

Caffeine produces a state of arousal in the central nervous system. High levels block the binding of an inhibitory neurotransmitter, adenosine, to the A2A adenosine receptor. Without caffeine, adenosine levels rise during the day, promoting interaction with its receptor, leading to increased sleepiness and lack of concentration. When adenosine binds normally to its receptors, it activates the adenylate cyclase cascade, which activates PKA, leading to changes in the phosphorylation state of many proteins inside the cell, including protein phosphatase 2A. These changes inhibit neural firing. Caffeine blocks these changes.

Hallucinogenic drugs

Illicit drugs like LSD, psilocybin, and ecstasy can produce hallucinations as they have profound effects on consciousness and perception of self and reality. Recent clinical studies have shown that under tightly controlled conditions and doses, these drugs might have significant therapeutic effects in the treatment of mental health issues such as depression and post-traumatic stress disorder. Hence, their mechanisms of action have been the source of many studies.

All these drugs bind to the human serotonin 2A receptor (5-HT_{2A}R), a metabotropic GPCR receptor of serotonin (5-hydroxytryptamine)**.** When serotonin binds, it leads to GPCR signaling, partly through beta-arrestins. These **adapter proteins** form complexes with ligand-bound and activated GPCRs and GPCR protein kinases.

Structures show that LSD binds to the orthosteric binding site (i.e., the active site, not an allosteric site) for serotonin. The study also found that serotonin and psilocin binding extends into an adjacent site called the extended binding pocket (EBP). These differences suggest that it should be possible to design agonists that have therapeutic, but not hallucinogenic, properties. Serotonin binding does not cause hallucinations. A drug, IHCH-7086, in mouse studies does not seem to provoke hallucinations but appears to have antidepressant effects (based on observations of mouse behaviors like twitching and freezing). The structures of serotonin (5HT) and psychotropic drugs that bind the $5HT_{2A}$R are shown in Figure 13.6.15.

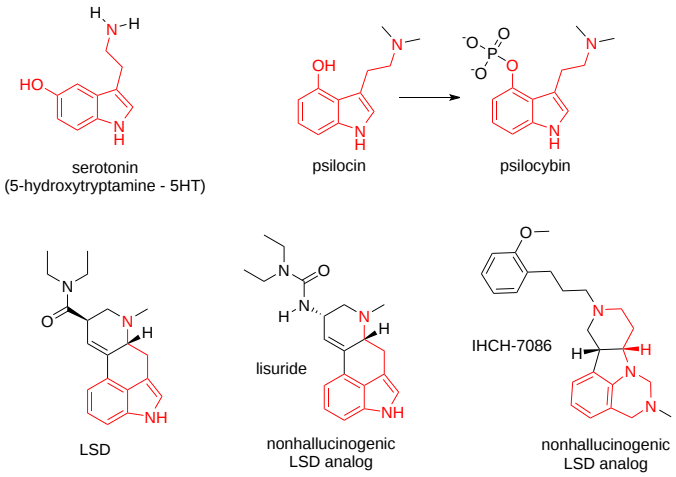


Figure 13.6.15: Structures of serotonin (5HT) and psychotropic drugs that bind the $5HT_{2A}$R

Figure 13.6.16 shows a model of the serotonin receptor (5HT-2A) with each of the bound drugs. Note the different occupancy of the orthosteric and extended binding pockets.

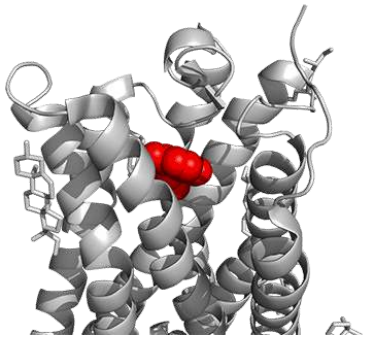


Figure 13.6.16: The serotonin receptor (5HT-2A) with bound serotonin and orthosteric binding analogs.

The GPCRs have been aligned for each of the poses. The ligands are represented in spacefill in a single color, as listed below:

- red: serotonin (5HT), the physiological agonist, 7WC4
- yellow: psilocin, 7WC5 LSD, 7WC6
- magenta: LSD, 7WC6
- cyan: IHCH-7086

An alternative explanation for why serotonin does not cause hallucinations and why the other agonists do is that, in addition to binding plasma membrane GPCRs, the agonist might have additional effects that are associated with hallucinations by binding intracellular membrane GPCRs.

The cerebral cortex plays a key role in what it means to be human. It is key to higher-level processing involving thought, learning, emotion, personality, language, and memory. As such, its dysregulation is involved in mental health conditions such as depression, et al., in which synaptic plasticity is decreased as evidenced by a reduction in dendritic structure required for synaptic connections. The serotonin (5HT) reuptake inhibitors (SSRIs) used to treat depression appear to increase neural plasticity and connections over time (so they don't give immediate responses). Hallucinogens that target the 5-HT2A receptor (5-HT2AR) also promote neuroplasticity. However, neuron growth in cortical cultures is not affected by serotonin. What is the mechanism that allows various combinations of neuron plasticity and hallucinations, all from the same receptor?

One factor determining these differential effects is that serotonin (5-HT) is more polar and cannot readily enter cells. At the same time, the HT2AR agonists that are hallucinogenic are less polar (more lipophilic) and could potentially enter cells where they might elicit hallucinogenic effects. Intracellular hallucinogens would bind internal 5-HT2AR receptors, which are found in cortical neurons on the Golgi and other organelles that are more acidic than the cytoplasm or extracellular environment. (In fact, in cortical neurons, the main location of the 5-HT2ARs is intracellular.) This might lead to more prolonged retention of intracellular hallucinogens and longer signaling (LSD has a profound hallucinogenic effect that lasts for 10 hours or more), leading to neuronal growth as well.

If serotonin is given to wild-type mice, no hallucinations occur, as evidenced by the lack of a head twitch response (HTR). However, if given to mutant mice who expressed serotonin transporter (SERT) on cortical neurons, the HTR response was observed. The import of serotonin led to neuroplasticity. It could be that serotonin is not the physiological ligand for intracellular HT2ARs. It might also imply that, as with the case of endocannabinoids, we have endogenous psychedelics. The subtle difference in the binding of serotonin and the hallucinogens in Figure 16 might have little to do with their tendency to produce hallucinations.

Recent Updates: LSD binding to the BDNF receptor TrkB 06/09/23

Much is still not known about how antidepressants work and what causes delays in their therapeutic effects. Increased synaptic connections through expanded neuroplasticity appear to be required for their therapeutic effects. The major action of drugs like Prozac, a serotonin reuptake inhibitor, is blocking the serotonin transporter (SERT), increasing extracellular serotonin levels in the neural synapse. Tricyclic anti-depressants as well as monoamine oxidase inMAOIstors (MAOI), also increase monoamine neurotransmitters in the synapse with delayed therapeutic effects (the monoamines act quickly, however). Perhaps these agents bind other receptors!

They do. The binding of both typical and fast antidepressants also occurs in the transmembrane domain of tyrosine kinase receptor 2 (TRKB), the brain-derived neurotrophic factor (BDNF) receptor, which is linked to increases in neuroplasticity. The receptor also binds cholesterol, which modulates its activity. The antidepressant binding site is formed on dimerization of their transmembrane domains. Mutations in the transmembrane region block the efficacy of the antidepressants.

LSD and other psychedelics also produce fast and long-lasting antidepressant effects promoted by increases in neuroplasticity. Studies have shown that LSD and psilocin bind to slightly overlapping sites in the transmembrane domain of the BDNF receptor as antidepressants, but with a 1000-fold higher affinity. If the LSD binding site on the serotonin 2A receptor (5-HT_{2A}) is blocked, LSD still has antidepressant and increased neuroplasticity effects.

Figure 13.6.17 shows the interaction of LSD and the metabolite of psilocybin, psilocin (PSI), with the TRKB receptor.

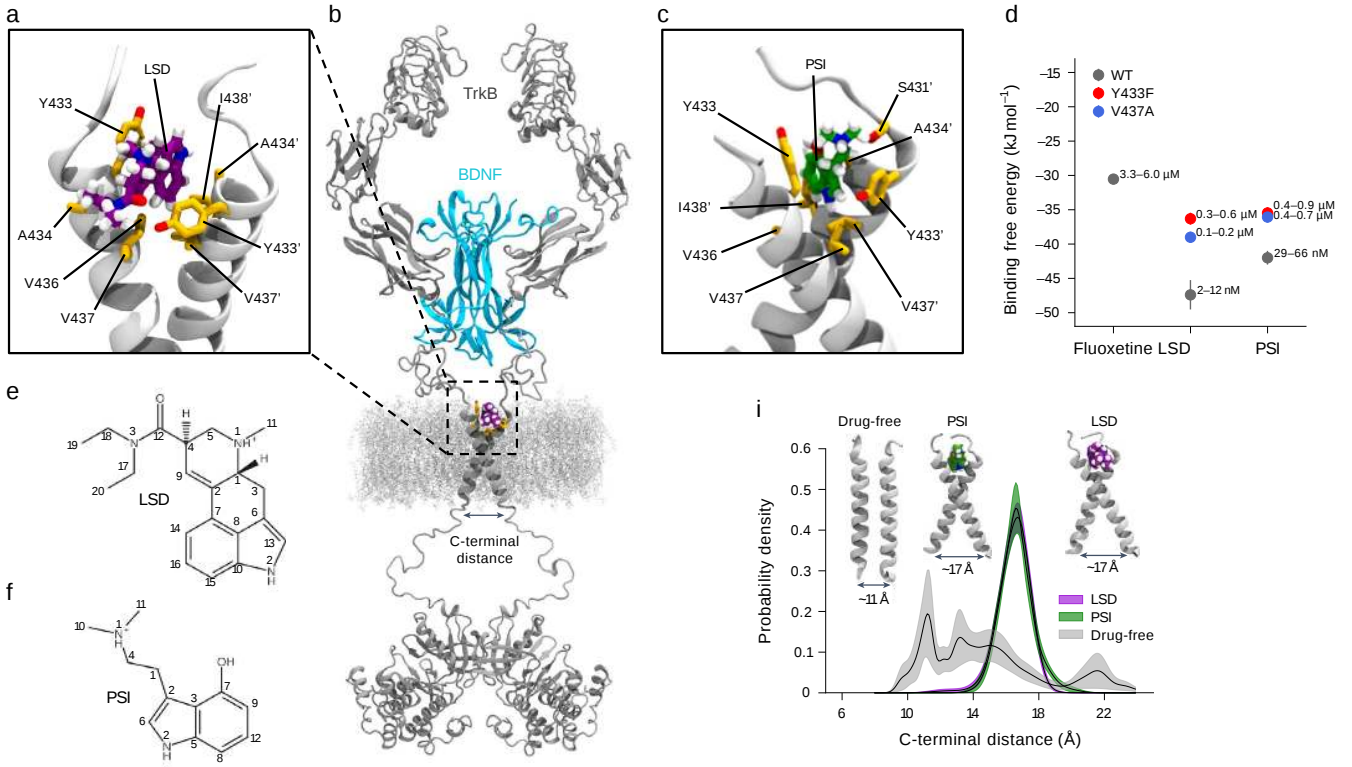


Figure **13.6.17**: **Characterization of the psychedelics binding site in the TrkB TMD.** Moliner, R., Girych, M., Brunello, C.A. *et al.* Psychedelics promote plasticity by directly binding to the BDNF receptor TrkB. *Nat Neurosci* **26**, 1032–1041 (2023). https://doi.org/10.1038/s41593-023-01316-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel a–**c** show representative MD snapshots showing the binding pocket for LSD (purple) (**a**) and PSI (green) (**c**) in the extracellular-facing crevice of the TrkB TMD dimer (gray). Side chains (yellow) of relevant binding site residues are displayed. A structural model of full-length TrkB dimer (gray) embedded in a lipid membrane is shown with bound BDNF (blue) and LSD (purple) (**b**).

Panel d shows in silico binding free energy estimations for fluoxetine, LSD, and PSI. Each free energy estimate (ΔG, circles) and its statistical error (SE, bars) were estimated from a separate set of FEP simulations ($n = 1$). Dissociation constants are given as a range with upper and lower bounds converted from ΔG − SE and ΔG + SE, respectively.

Panels e,**f** show chemical structures of LSD (**e**) and PSI (**f**) with atom numbers annotated.

Panel i shows the distributions of TMD dimer C-terminal distance and that LSD and PSI stabilize the cross-shaped conformation of TrkB favorable for receptor activation in a 40 mol% CHOL membrane. Lines represent the mean distribution, and bands represent the standard errors ($n = 10$ independent simulations). TMD conformations corresponding to indicated C-terminal distances and drug-bound states are shown in the inset.

Figure **13.6.18** shows the different binding modes of LSD and Fluoxetine (Prozac), a selective serotonin reuptake inhibitor (SSRI) for TrkB. Fluoxetine is used to treat depression, obsessive-compulsive disorder (OCD), bulimia nervosa, and panic disorder.

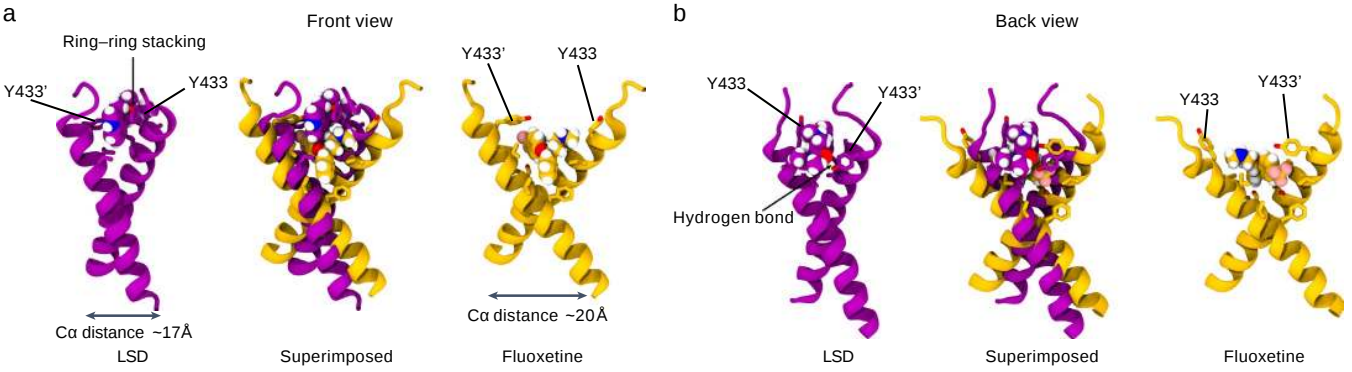


Figure **13.6.18**: **Different TrkB binding modes of LSD and fluoxetine.** Moliner et al., ibid.

Panels a,**b**, show representative snapshots of atomistic MD simulations showing the front (**a**) and back (**b**) views of the binding pockets for LSD (purple) and fluoxetine (yellow) in the extracellular-facing crevice of TrkB TMD dimers. Side chains of relevant binding site residues are displayed. Superimposed structures of TrkB optimally bound to LSD or fluoxetine reveal that, while some residues involved in binding are shared (Y433 and V437), the binding modes are different. Fluoxetine binds at a site deeper within the dimer, locking the TMD dimers in a more open cross-shaped conformation (distance between the center of mass L451–L453 Cα atoms of each monomer ~20 Å). In contrast, LSD binds closer to the N-terminus of the TrkB TMD and establishes more stable interactions with the dimer: a hydrogen bond between the oxygen atom of the diethylamide group of LSD and the Y433 residue of one monomer, and pi-stacking of the aromatic backbone of the drug with the Y433 residue of the second monomer, locking the TMD dimer in a tighter cross-shaped conformation (L451–L453 Cα distance ~17 Å) compared with fluoxetine. Drugs are shown in van der Waals representation. Oxygen, nitrogen, and hydrogen atoms are shown in red, blue, and white, respectively.

13.6.7: Summary

In this chapter, we explore the molecular and biophysical foundations of neural signaling, focusing on how ions traverse cellular membranes via specialized ion channels and how neurotransmitters modulate these processes to initiate and regulate electrical activity in neurons. Key topics include:

- **Ion Channels and Membrane Permeability:**
  Ion channels are integral membrane proteins that provide pathways for the rapid, selective movement of ions (e.g., Na^+, K^+, Cl^-) across the otherwise impermeable lipid bilayer. The chapter contrasts synthetic lipid bilayers—which have extremely low ion permeability—with biological membranes that express various gated channels and transporters. These channels are responsible for the establishment and dynamic modulation of the transmembrane potential.
- **Generation and Maintenance of Membrane Potentials:**
  The resting membrane potential arises from a combination of ion gradients—maintained by pumps such as the Na^+/K^+-ATPase—and differential membrane permeability. Glial cells, for example, maintain a resting potential largely defined by a non-gated potassium leakage channel. This potential is crucial for neuronal excitability and signal propagation.
- **Action Potential Dynamics:**
  Neuronal signaling involves the generation of action potentials, rapid fluctuations in membrane potential caused by the sequential opening and closing of voltage-gated sodium and potassium channels. The chapter discusses the phases of depolarization and repolarization, and how ion fluxes contribute to the rapid firing of neurons.
- **Synaptic Transmission and Neurotransmitter Receptors:**
  At synapses, presynaptic neurons release neurotransmitters that bind to receptors on the postsynaptic cell. Ionotropic receptors, such as the acetylcholine receptor at the neuromuscular junction, directly gate ion channels to produce fast, transient responses. In contrast, metabotropic receptors, which are typically G protein-coupled, trigger intracellular signaling cascades that modulate ion channel activity and gene expression, producing slower and longer-lasting effects.
- **Ion Channel Selectivity and Neurotoxins:**
  The selectivity of ion channels is achieved through specialized narrow pores that strip ions of part of their hydration shells and coordinate them via specific interactions. Neurotoxins like tetrodotoxin and saxitoxin, which block sodium channels, provide insights into channel structure and function and underscore the importance of precise ion flow in neural activity.
- **Integration of Neural Signaling:**
  Neural communication is not only determined by the rapid opening and closing of channels but also by the coordinated activity of multiple receptor types and signaling proteins. This integration underpins complex processes such as synaptic plasticity, learning, memory, and behavior. The chapter highlights how disruptions in these mechanisms can contribute to neurological disorders.

Overall, the chapter presents a detailed view of how ion channel function, neurotransmitter receptor dynamics, and synaptic architecture converge to produce the electrical signals that underpin neural activity. By linking structural features to functional outcomes, it sets the stage for understanding the complex interplay between molecular events and higher-order brain functions in both health and disease.

This page titled 13.6: Gated Ion Channels - Neural Signaling is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

13.7: Signal Transduction - Vision and Olfaction

Learning Goals (ChatGPT o3-mini)

1. **Describe Specialized Cellular Structures:**
   - Explain the unique architecture of photoreceptors (rods and cones) and olfactory sensory neurons (OSNs), including the role of specialized cilia and membrane structures in signal detection.
   - Compare the structural differences between rod outer segments (stacked disk membranes) and cone outer segments (plasma membrane invaginations), and relate these differences to their functional roles in light detection.
2. **Understand Signal Detection Mechanisms:**
   - Outline how visual pigments in photoreceptors, with covalently bound chromophores, convert light into a molecular signal.
   - Describe how odorant receptors in OSNs detect volatile molecules through noncovalent, reversible interactions, and how these receptors initiate the olfactory transduction cascade.
3. **Analyze Transduction Cascades:**
   - Detail the molecular steps of the phototransduction cascade, including the activation of rhodopsin, the role of transducin, and the subsequent changes in cGMP levels that lead to hyperpolarization.
   - Compare and contrast the olfactory transduction cascade, highlighting the roles of odorant receptors, Golf, adenylyl cyclase 3 (AC3), cAMP production, and the opening of cyclic nucleotide-gated channels leading to depolarization.
4. **Evaluate Signal Amplification Strategies:**
   - Explain the mechanisms that amplify the initial sensory signal in photoreceptors, such as the multiple activation of transducin by a single photoactivated rhodopsin and the cooperative gating of CNG channels.
   - Describe how OSNs achieve signal amplification via a secondary Ca^{2+}-activated Cl^- current following cAMP-mediated activation, and discuss the advantages of using Cl^- over Na^+ in the olfactory cilia.
5. **Examine Inactivation and Adaptation Processes:**
   - Illustrate how photoreceptors inactivate their response through phosphorylation, arrestin binding, and GTP hydrolysis, ensuring rapid reset for subsequent stimuli.
   - Discuss the feedback mechanisms mediated by Ca^{2+} in both photoreceptors and OSNs that lead to adaptation, including the roles of GCAPs, recoverin, and Na^+/Ca^{2+} exchangers in maintaining sensitivity over a wide dynamic range.
6. **Integrate Electrophysiological Techniques:**
   - Describe how loose-patch (suction pipette) recording techniques are used to measure stimulus-induced electrical responses in photoreceptors and OSNs.
   - Interpret electrophysiological data showing graded potentials in photoreceptors versus action potentials in OSNs, and relate these differences to their respective sensory functions.
7. **Relate Molecular Mechanisms to Sensory Function and Disease:**
   - Connect the structural and functional properties of phototransduction and olfactory transduction to their roles in high-sensitivity signal detection and discrimination.
   - Analyze how mutations or disruptions in transduction components (e.g., visual pigment misfolding, CNG channel dysfunction, or OR mutations) can lead to sensory disorders, such as retinitis pigmentosa or anosmia, and discuss potential therapeutic strategies.

By achieving these learning goals, students will develop a comprehensive understanding of how specialized sensory neurons convert external stimuli into electrical signals, how these signals are amplified and terminated, and how these processes are critical for proper sensory function and adaptation.

This chapter section is taken in its entirety from Genovese et al Front. Cell. Neurosci., 08 October 2021 | https://doi.org/10.3389/fncel.2021.761416. Creative Commons Attribution License (CC BY)

13.7.1: Sensory Transduction in Photoreceptors and Olfactory Sensory Neurons

Photoreceptors and olfactory sensory neurons (OSNs) have highly specialized structures that enable them to capture their respective stimuli of light and odorant ligands. Both photoreceptors and OSNs have evolved highly specific abilities to detect and discriminate light wavelengths or odors. They use intricate transduction mechanisms to convert sensory stimuli into electrical signals. Their transduction cascades can greatly amplify the signal and enhance the signal-to-noise, enabling these cells to detect and distinguish minute stimuli within very noisy background conditions. Such transduction mechanisms provide for modulation at multiple steps to adapt the sensory neurons to different background stimulation and optimize the capture of useful information about the surrounding world.

In this review, we summarize some of the key structural and functional features of vertebrate rod and cone photoreceptors and OSNs, as well as the molecular mechanisms underlying their function. While describing the features of both cell types, we emphasize the similarities and differences between photoreceptors and OSNs and the unique features of each cell type that make them perfectly suited to perform their function.

13.7.2: Signal Detection in Photoreceptors and Olfactory Sensory Neurons' Specialized Cilia

Vertebrate rod, cone photoreceptors, and OSNs are ciliary neurons, as shown in Figure 13.7.1. They have specialized cilia, where the sensory stimulus's initial detection occurs to activate a sensory transduction cascade. Rods and cones have a single cilium that has evolved to accommodate a stack of ~1,000 membrane disks where the visual pigment is expressed at a very high 3–5 mM concentration. In the case of rods, the disks are enveloped by the plasma membrane, whereas in cones the disks are formed by invaginations of the plasma membrane. As light enters the eye and reaches the retina, it travels along the length of the rod and cone outer segments. The orientation of the elongated outer segments along the light path and the high density of visual pigment in their disks result in ~50% probability that a visual pigment molecule absorbs an incident photon. In the case of OSNs (right panel), odorant ligands are detected in the ~20 cilia protruding from each dendritic knob which are immersed in the mucus layer covering the olfactory epithelium. The olfactory cilia, which are motile in amphibians but not in rodents, are only about 0.1–0.2 μm thin but can reach up to 100 μm in length depending on the species. While this greatly increases the surface membrane area available to incorporate olfactory receptor (OR) proteins to detect odorants, it also greatly reduces the ciliary volume with potentially detrimental effects (see below).

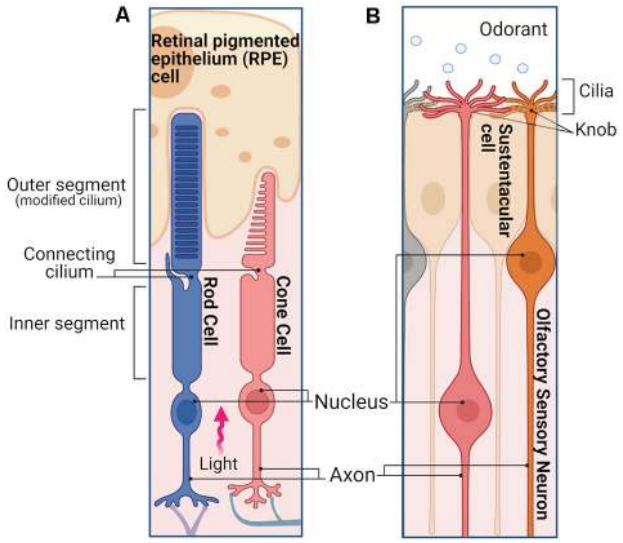


Figure 13.7.1: Photoreceptors and olfactory sensory neurons (OSNs).

(A) Simplified schematic representation of a rod and a cone in the retina. Photoreceptors are polarized neurons with a specialized morphology optimized to detect light stimuli. The outer segments of both rods and cones are modified sensory cilia, containing

membrane disks organized in a stack. In the case of rods, the outer segment has a slim rod-like structure in which the plasma membrane encloses the disks. The outer segment of the cones has a stocky conical-shaped structure, in which invaginations of the plasma membrane constitute the disks. The outer segment does not contain any proteins of the cell translation machinery, which are mostly localized in the inner segment, including the endoplasmic reticulum, Golgi, and mitochondria. The connecting cilium connects the outer and inner segments. In contrast, distal to the inner segment is the cell body containing the nucleus, followed by the axon and synaptic termini that extend into the outer plexiform layer, where they synapse with the second-order neurons. When the light enters the eye, after reaching the retina, it travels along the length of the rod and cone inner segment until finally reaching the outer segments.

(B) Simplified schematic of an OSN in the olfactory epithelium. OSNs are ciliated bipolar neurons; their apical dendrites extend to the surface of the epithelium, terminating with a spherical structure called a dendritic knob, from which the sensory cilia enter the mucus layer. The ciliary membrane contains the olfactory receptors (ORs) necessary to detect different odorants. Distal from the knob is the cell body of the OSN with its nucleus, followed by a long axon that projects to the olfactory bulb, where it synapses with the second-order neurons. Images created with BioRender.com.

13.7.3: Electrophysiological Approaches to Record Light- and Odorant-Induced Responses

The similar morphological structure of rods, cones, and OSNs, with a ciliary part able to detect the respective stimuli and an adjacent cell body, allows similar electrophysiological approaches to recording stimulus-induced responses in these cell types. The cell body of a photoreceptor or an OSN can be sucked into the tip of a recording pipette by using a loose-patch (or suction pipette) recording configuration. This leaves the outer segment of photoreceptors or the olfactory cilia exposed and accessible to bath solution changes, e.g., the application of pharmacological agents or odorants, in the case of OSNs. Suction pipette recordings can be performed from isolated sensory neurons, as shown in Figure **13.7.2** (A, B, and D show a salamander rod, salamander cone, and salamander OSN) but also from dissected retina tissue, as in the case of the outer segment of a mouse rod drawn in the recording electrode from a piece of the retina (C). This recording configuration measures the transduction current entering the photoreceptor outer segment or olfactory cilia, and leaving *via* the cell body.

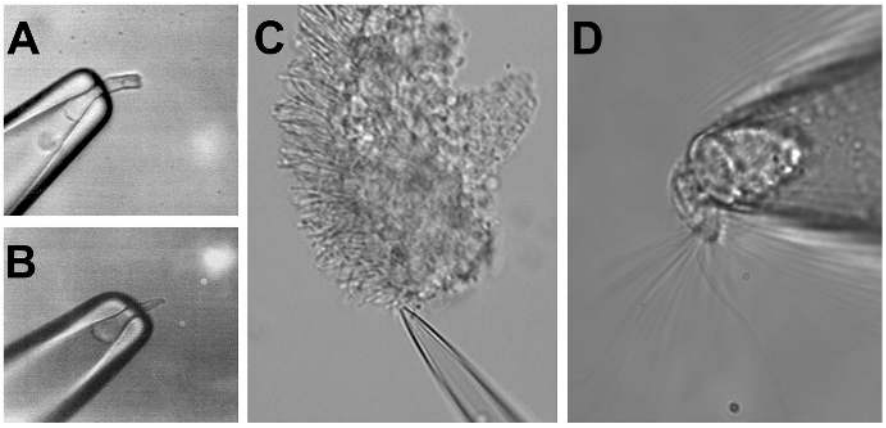


Figure \(\
PageIndex
{2}\):
Figures modified from
Reisert and Matthews (2001)
,
Kefalov et al. (2010)
, and
Kefalov (2012)
with permission.

A fundamental difference between photoreceptors' and OSNs' responses to stimuli lies in their polarity. In the absence of light, rods and cones are kept depolarized by a standing inward current of approximately 20–40 pA for amphibian cones and rods, and 7–15

pA for mouse photoreceptors. This depolarizing current is gradually suppressed upon light stimulation until, for sufficiently high light intensities, it is reduced to zero, as shown in Figure **13.7.3** (A, B, mouse rod, and cone responses, respectively), leading to photoreceptor hyperpolarization. Similar to rods, but unlike cones, the OSNs show comparatively little spontaneous activity without stimuli. Different OSNs show varying levels of spontaneous basal activity determined by the constitutive activity of their ORs.

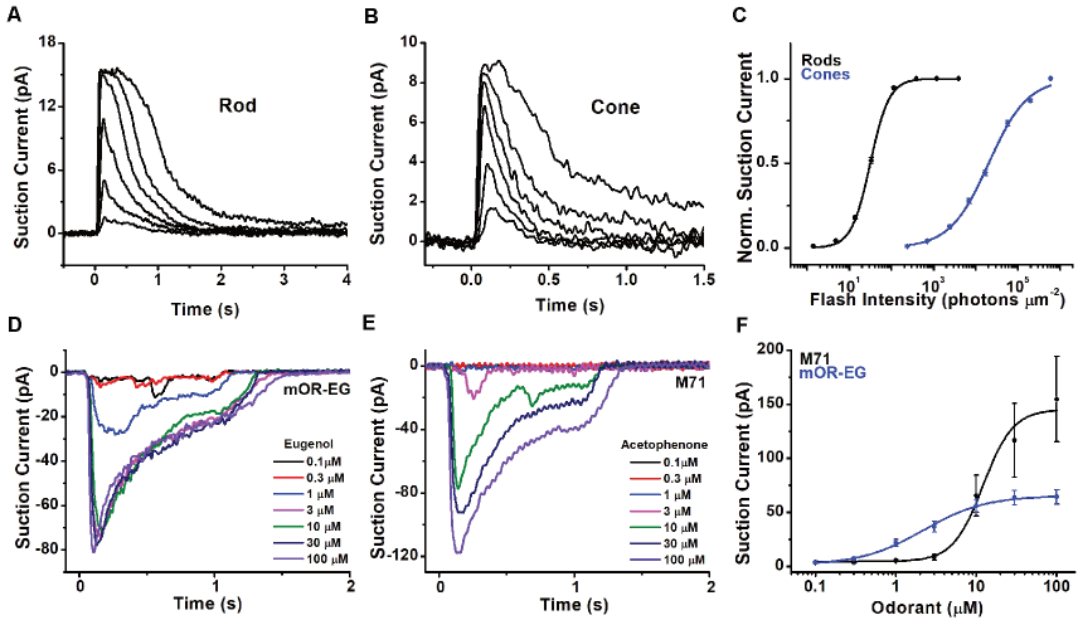


Figure **13.7.3**: Modified from Kefalov et al. (2010), Sakurai et al. (2011), Deng et al. (2013), and Dibattista and Reisert (2016) with permission.

In the presence of odorants, OSNs generate an inward receptor current that leads to depolarization and the generation of action potentials. This receptor current is odorant concentration-dependent and increases progressively with increasing stimulation until it eventually saturates at high odorant concentrations. Responses recorded from OSNs expressing different olfactory receptors can generate fairly different response amplitudes when stimulated with their respective agonists (Panel D, E): responses recorded from mouse OSNs that express the mOR-EG or the M71 olfactory receptor, which are activated by the ligands eugenol and acetophenone, respectively.

The hyperpolarization and signals carried by graded potentials in photoreceptors vs. depolarization and signals carried by action potentials in OSNs represent another fundamental difference between these two types of sensory neurons. These topics and the differences in synaptic structure and transmission between photoreceptors and OSNs go beyond the focus of this review.

13.7.4: Sensitivity of Photoreceptors and Olfactory Sensory Neurons

Due to their unique structure, photoreceptors and, to a lesser extent, OSNs have achieved exquisite sensitivity that optimizes the detection of stimuli within the respective sensory organs. In addition, both sensory receptors use a transduction cascade to amplify the signal (see below). As a result, rod photoreceptors can reliably detect single photons, enabling humans to perceive light with as few as six photons detected by adjacent rods. This renders rods perfectly suited for dim light vision, with a dynamic range spanning lights from a dark, cloudy night to sunrise. Cones are ~100-fold less sensitive than rods, making them suited for daytime light conditions. Figure **13.7.3***C* compares the intensity-response function of mouse rods and cones, demonstrating the much lower sensitivity of cones compared to rods.

Most OSNs respond to odor concentrations in the low micromolar range, but they can also reach exquisite sensitivity and are capable of detecting odors at the nanomolar concentration range. Picomolar sensitivity is reached by a subset of OSNs that express receptors specialized in detecting amines, the trace-amine-associated receptors. Compared to rods, OSNs do not reach such high sensitivity. A single odorant molecule cannot activate them; instead, it requires around 30 odorant binding events to begin firing action potentials reliably. The detection of odorants in the olfactory epithelium can be further enhanced by expressing a wider number of different OR genes, more than 350 in humans and 1,000 in mice, with overlapping response profiles to odorants. A

larger number of OSNs, particularly in species relying heavily on their sense of smell, may further enhance the detection of odorants. For instance, the human olfactory epithelium covers ~3–4 cm^2 and contains approximately 5–6 million OSNs, while in the case of dogs, the area of the olfactory epithelium is 18–150 cm^2 and contains 150–300 million OSNs.

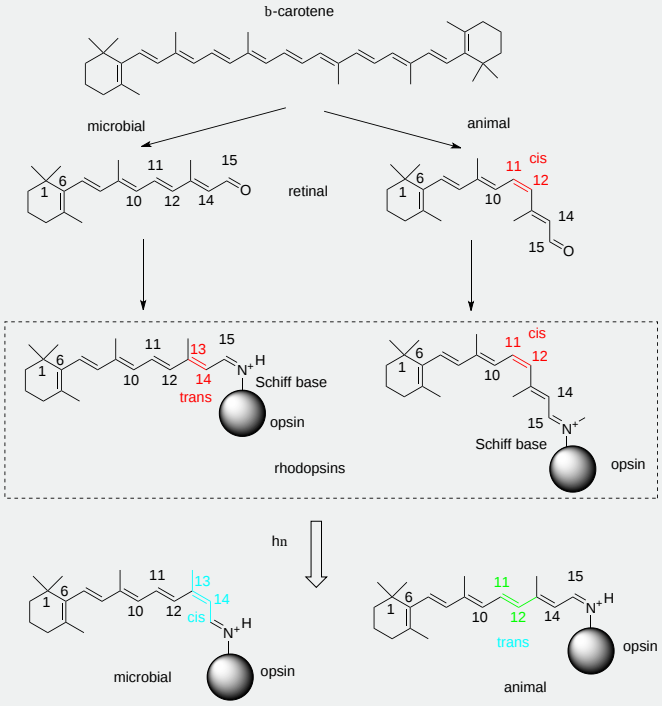
13.7.5: Detection of Stimuli

In both photoreceptors and OSNs, the detection of stimuli is mediated by G protein-coupled receptors. In photoreceptors, this function is achieved by rod and cone visual pigments, which consist of a protein, opsin, covalently attached to the visual chromophore, typically 11-cis-retinal. The chromophore is a reverse agonist, keeping the receptor molecule in the inactive ground state. Absorption of a photon by 11-cis-retinal triggers its conformational change to all-trans-retinal, which, in turn, results in the rearrangement of the opsin transmembrane helices and switch of the visual pigment molecule into its active state.

We discussed bacterial rhodopsin in an earlier section. We present it here again to highlight its similarity to animal light-sensing proteins.

Review: Retinal, Opsin and Light Transduction in Bacteria

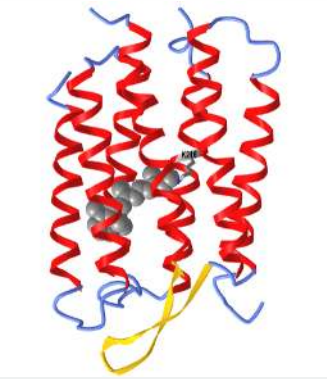
Microbial **rhodopsins** are phototransducing proteins with a conjugated chromophore retinal, covalently attached to the protein opsin through a Schiff base (imine) linkage. The holoprotein (opsin with the attached retinal) is called rhodopsin. Retinal is derived from beta-carotene. The structures of animal and microbial retinals are shown below.



Structures of animal and microbial retinals

When light of the correct wavelength is absorbed, an electron in a pi molecular orbital in retinal is promoted to a pi antibonding molecular orbital, breaking a 2-electron pi bond in the structure at a certain site in the isoprenoid chain, allowing rotation around the now single bond. After the electrons return to the ground state, the final result is photoisomerization of the trans 13-14 and cis 11-12 bonds in microbial and animal retinal, respectively, to their respective cis 13-14 and trans 11-12 configuration. This conformational change in the bound retinal induces a conformational change in the protein opsin, leading to signaling.

The figure below shows an interactive iCn3D model of one monomer of bacterial rhodopsin (1C3W) containing retinal attached through a Schiff base to Lys 216.



NCBI iCn3D Bacterial rhodopsin (1C3W). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...9HzvfdFvb9kcc6

The covalently attached retinal is shown in gray spacefill. Opsin is a membrane protein that spans the membrane with seven helices. Hence, it is very similar to a GPCR.

The activated visual pigment then binds to a G protein, transducin, activating it. The transducin activation triggers the transduction cascade that ultimately generates the cellular response. Eventually, the all-trans-retinal chromophore is released from opsin after the covalent Schiff base between them is hydrolyzed, leaving behind chromophore-free opsin. Notably, without a chromophore, opsin has residual activity and, in sufficient quantities, can produce steady activation of the photoreceptors, similar to steady background light, thus modulating the sensitivity of photoreceptors. This process is known as bleaching adaptation, indicating the production of free opsin after the photoactivation of the visual pigment and dissociation of the visual chromophore.

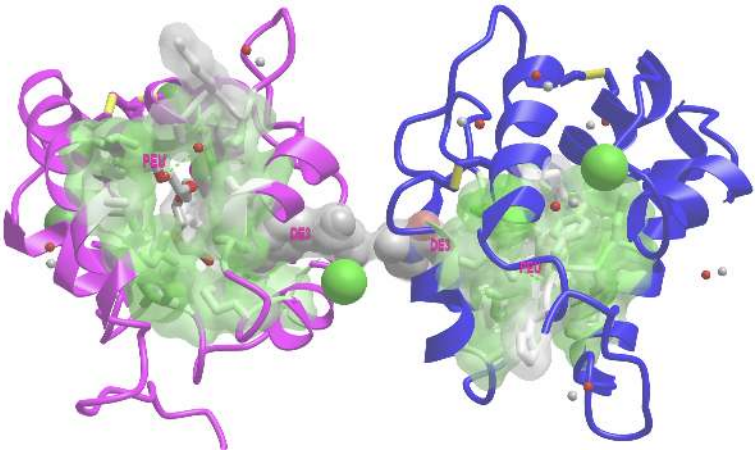
Unlike in photoreceptors, where the ligand, a light-sensitive reverse agonist, is covalently attached to opsin, in olfaction, the ligands are dissolved in the mucus covering the surface of the olfactory epithelium and come into direct contact with the OR proteins expressed in the OSN ciliary membrane. This results in the activation of the receptor protein that, in turn, is transduced downstream to a G protein to trigger a transduction cascade resulting in the cellular response. The binding of the ligand to the receptor protein is noncovalent and rapidly reversible. ORs, like other G protein-coupled receptors, display antagonism, inverse agonism, and partial agonism, leading to suppressed responses to their agonists, a reduction in basal activity without stimulation, or suppression of the maximal response.

Recent Updates: 7/25/24

13.7.6: Detection of Odorants in Insects - The Effects of DEET

DEET (N,N-diethyl-meta-toluamide) is the active ingredient in many mosquito and insect repellents. It's been widely used for a long time, but its biochemistry is still unclear. It inhibits olfactory signaling pathways in insects. These pathways involved olfactory receptors (OR), odorant-binding proteins (OBPs), and other mostly downstream proteins. Studies have progressed furthest on the small soluble OBPs found in mosquito lymph, which consists of three classes: pheromone-binding proteins, general OBPs, and antennal binding proteins X. The OBPs likely bind and transport generally low-solubility odorants as part of the initial steps in signal transmission. Likely, the odorant:OBP complex binds to an odorant receptor. An individual OBP also likely binds to different odorants with similar structural topology and chemical properties, as the number of recognized odorants greatly exceeds the number of OBPs.

The classic OBP contains six alpha-helices and three disulfide bonds. The figure below shows an interactive iCn3D model of the Odorant Binding Protein 1 from Anopheles gambiae (AgamOBP1) with bound DEET (N,N-Diethyl-meta-toluamide) and PEG 216 (3N7H).



NCBI iCn3D Odorant Binding Protein 1 from Anopheles gambiae (AgamOBP1) with DEET (N,N-Diethyl-meta-toluamide) and PEG 216 (3N7H). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wWGmkkJeoQvC38

The crystal structure is a dimer (colored magenta and blue), which is likely the biological unit since the protein concentration is high in insect lymph (which could drive dimerization). OBP1 has two ligands bound in each subunit, polyethylene glycol (PEU, a crystallization agent) and DEET (DE3). The iCn3D structure above shows the amino acids in both dimers that are 5 angstroms (Å) from the bound PEGs. The PEG binds in a hydrophobic, L-shaped tunnel (green surface coloring indicates a hydrophobic pocket). The interactions are predominantly nonpolar with one hydrogen bond to a water molecule. DEET binds with a reported K_d of 31 μM near the entrance in the short segment of the L-shaped tunnel. Since DEET and PEG bind to different parts of the tunnel, the tunnel can likely accommodate other odorants of similar topology and hydrophobicity. Little conformational change occurs on the binding of DEET compared to the unliganded protein. The protein has the Uniprot ID Q8WRX3, but nothing about its interaction partners is described.

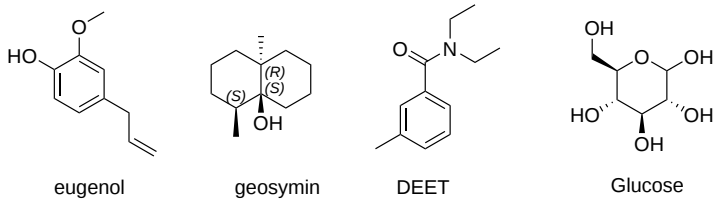
In contrast, insect odorant receptors (OR) are tetrameric, ligand-gated ion channels. On binding of the odorant, the channel is gated open. The OR often binds coreceptors that facilitate channel restructure and opening. It is unclear from the literature if the odorant binding protein 1 is a coreceptor for an odorant receptor or delivers odorants (such as DEET) to the receptor. The animations below show the closed and open form of the *Ap* OR5-Orco heterocomplex from the pea aphid *Acyrthosiphon pisum.* A conformational change occurs when the odorant geranyl acetate is added (not shown). The active structure contains the OR5 receptor and the Orco coreceptor.

Figure: *Ap* OR5-Orco heterocomplex from the pea aphid *Acyrthosiphon pisum* morphing from the closed to the open channel on binding of the odorant geranyl acetate (8Z9Z to 8Z9A). Image made using Pymol.

Insect odorant receptors have an intracellular N-terminus, an extracellular C-terminus, and seven transmembrane helices. Hence, they are very similar to GPCRs but with a flipped orientation, so they can't couple to intracellular G-proteins. Rather, they are

ligand-gated ion channels. They bind proteins named Orco, which act as co-receptors.

The olfactory receptor *Mh*OR5 from the jumping bristletail *Machilis hrabei* is a homotetrameric odorant-gated ion channel that binds a broad range of odorants. Several are shown below, along with the non-odorant molecule glucose.



The activity of several odorants, including eugenol (an oily substance found in clove, nutmeg, cinnamon, basil, and bay leaf), geosmin (a fungal metabolite that confers the smell in the air after fresh rain), DEET, and a negative control, glucose, is shown in the figure below.

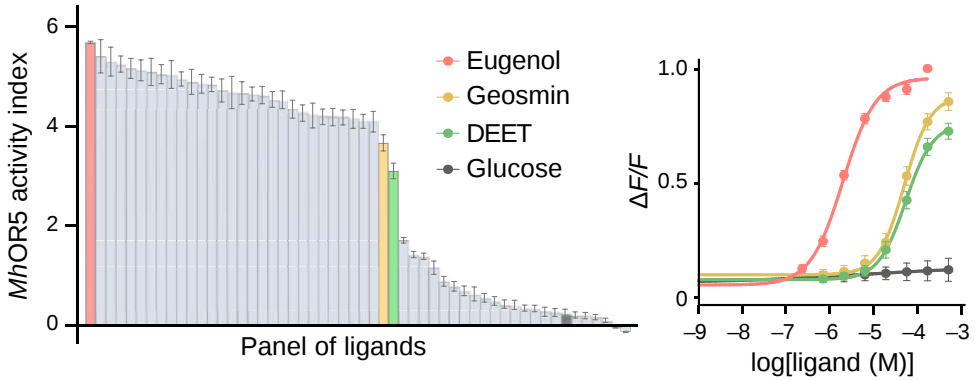


Figure: Left, activity of *Mh*OR5 evoked by a panel of 54 small molecule ligands. Right, dose–response curves of *Mh*OR5 to selected ligands. del Mármol, J., Yedlin, M.A. & Ruta, V. The structural basis of odorant recognition in insect olfactory receptors. *Nature* **597**, 126–131 (2021). https://doi.org/10.1038/s41586-021-03794-8. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Notice that Eugenol binds with an apparent Kd of around 2 μM (antilog of -5.7, the log [Ligand] at half-maximal binding (as measured by changes in fluorescence). At the same time, geosmin and DEET have a Kd closer to 100 μM.

(back to main article)

13.7.7: Discrimination Between Stimuli

The spectral sensitivity of individual rod and cone photoreceptors is dictated by the absorption properties of their visual pigments. Typically, each photoreceptor type expresses only one type of opsin; in the case of the human retina, rods express rod opsin, whereas cones express long wavelength (LW, red), middle wavelength (MW, green), or short wavelength (SW, blue) opsin. When bound to the chromophore, the amino acid structure of each opsin determines the optical properties of the resulting visual pigment and the spectral sensitivity of the photoreceptors expressing it. As a result, species existing in environments with characteristic light distribution, such as deep-sea fish, have visual pigments that have evolved to optimize their spectral sensitivity. A second factor controlling the optical properties of the visual pigment is the structure of the visual chromophore. Most species, including mice and humans, use 11-cis-retinal, a derivative of Vitamin A, also known as A1. However, some amphibians and fish also use 3,4-dehydro-11-cis retinal, also known as A2. This chromophore has an extra conjugated double bond in its structure, which shifts the absorption spectrum of A2 visual pigments to longer wavelengths than the A1 visual pigment embedded in the same opsin molecule. Some aquatic and amphibian species use the A1/A2 chromophores to shift their spectral sensitivity from murky waters dominated by longer wavelengths of light to seawater and air, dominated by shorter wavelengths of light. One notable example includes the toad, where the retina is populated by A1 visual pigment in its ventral section, receiving light from above the surface of the water, and by A2 visual pigment in its dorsal section, receiving light from below the surface of the water. A shift in the chromophore can also occur during the lifetime of the animal as its environment changes, such as the A2 to A1 shift in salamanders

as they metamorphose from the larval (aquatic) to the adult (terrestrial) stage, or the A2 to A1 shift in Atlantic salmon during migration from sea to freshwater.

Similar to photoreceptors, the ligand specificity of the OSNs is also dictated by the expression of OR genes in their cilia. As photoreceptors, each OSN generally expresses only one receptor gene, so the OR's structure determines its ligand specificity expressed in that particular cell. However, photoreceptors typically use no more than five opsin genes to cover the visible spectrum. In contrast, OSNs can use hundreds, in the case of humans, to thousands or more, for rodents and dogs, OR genes to cover the odor space. The same OR can be activated by multiple odorants with different sensitivities. A given odorant can activate different ORs with different half-maximal concentrations. This generates a complex mosaic of ORs and odorant response pairs. Figure 13.7.3*F* compares the dose responses of OSNs expressing either the mOR-EG or the M71 OR to eugenol and acetophenone, respectively. In this case, mOR-EG OSNs display higher sensitivity to their agonist than M71 OSNs. However, this does not preclude the possibility that the M71 OR is more sensitive to another ligand, resulting in a more left-shifted dose-response relation than the one seen with acetophenone. Conversely, the dose-response relation of M71 OSNs to benzaldehyde is approximately 10-fold right-shifted compared to acetophenone.

Determining the ligand specificity of ORs is an ongoing endeavor. Due to the large number and diversity of OR genes and the nearly endless number of odorant molecules, understanding the overall mechanisms that control their ligand binding affinity and specificity remains challenging. Receptor modeling approaches to understanding and predicting OR–odorant molecule interactions can provide valuable insights. Still, they are somewhat hampered by the lack of a crystal structure of any vertebrate OR. The rhodopsin structure is often used as a guide and homology model to predict the structure of ORs.

13.7.8: Sensory Transduction Activation

In both photoreceptors and OSNs, the detection of stimuli by their respective G protein-coupled receptors is converted into electrical signals *via* the activation of a G protein-coupled to a second messenger transduction cascade. The two pathways, though distinct, share an amazing level of similarity, as shown in Figure 13.7.4. Thus, the second messenger in both cases is a cyclic nucleotide, cGMP in photoreceptors and cAMP in OSNs. As a result, the activation of both transduction cascades results in a rapid shift in the equilibrium between synthesis and hydrolysis of the respective cyclic nucleotide, which is then sensed by the cyclic nucleotide-gated (CNG) transduction channels in the plasma membrane of the photoreceptor outer segment or olfactory cilium.

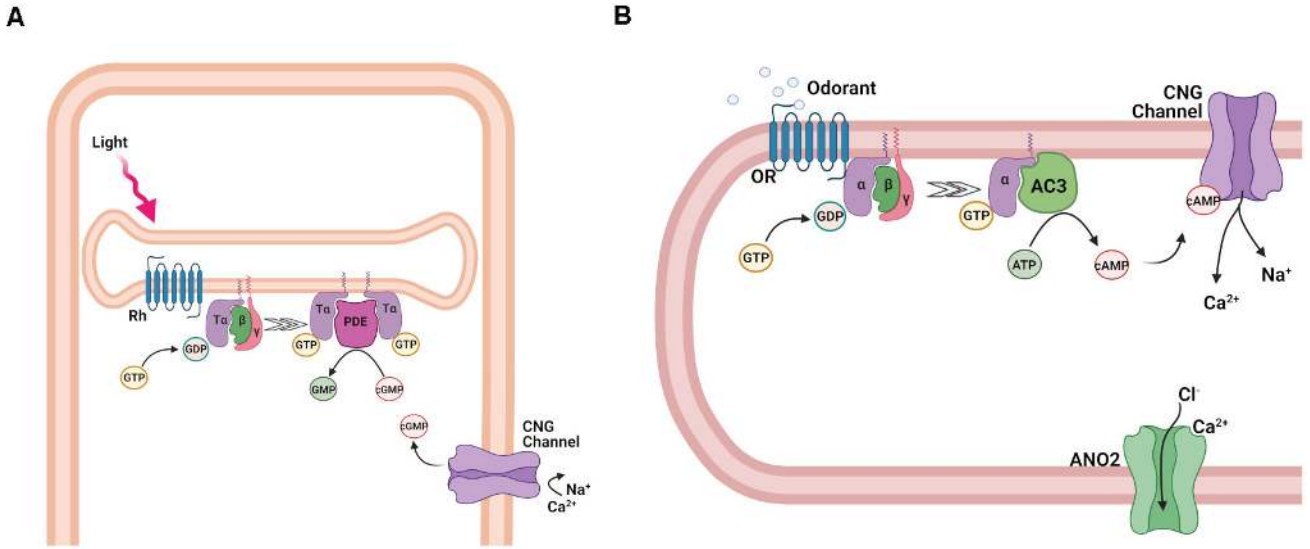


Figure 13.7.4: Activation of the transduction cascade in rod photoreceptors and OSNs.

(A) Schematic representation of the phototransduction cascade in rods. Abbreviations: rhodopsin (Rh), Tα, β, and γ subunits of the retinal G protein, transducin (T), guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), phosphodiesterase (PDE), guanosine monophosphate (GMP), and cyclic guanosine monophosphate (cGMP), and cyclic nucleotide-gated (CNG) channel.

(B) Schematic representation of the olfactory transduction cascade in OSNs. Abbreviations: Olfactory receptor (OR), guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), $G\alpha_{olf}$, β, and γ, subunits of the olfactory G protein; adenylyl cyclase 3 (AC3), adenosine-5′-triphosphate (ATP), cyclic adenosine monophosphate (cAMP), cyclic nucleotide-gated (CNG) channel; Ca^{2+}-activated Cl^- channel anoctamin 2 (ANO2). Images created with BioRender.com.

In the case of photoreceptors, the photoactivated visual pigment binds to and activates the trimeric G protein transducin (T) (Panel A), causing the exchange of GDP for GTP on its α-subunit, which is part of the $G\alpha_t$ protein family. Following the subsequent dissociation of the α-subunit (Tα) from its β/γ complex (Tβγ), Tα then binds to the cGMP phosphodiesterase (PDE) complex, relieving the inhibition of its catalytic α- and β-subunits by its inhibitory γ-subunits. All these transduction proteins are embedded in or tethered to the disc membranes inside rods or contained in cones' cell membranes. As a result of their activation, the hydrolysis of free cGMP in the outer segment is upregulated, causing its rapid decline, and partial or complete closure of the cGMP-gated channels expressed in the rod and cone cell membranes. The closure of the CNG channels leads to the reduction of the inward transduction current, followed by the hyperpolarization of the cells, and a reduction of neurotransmitter release to second-order neurons within the retina. Inversely, in the absence of light, the opening of CNG channels and the resulting inward transduction current is sustained by the continuous cGMP production by guanylyl cyclase (GC).

Similarly, in OSNs (Panel B), the ligand-activated OR proteins bind to the G protein G_{olf}, causing its dissociation into active $G\alpha_{olf}$ and olfactory β- and γ-subunits, $G\beta\gamma_{olf}$. In contrast to transducin, however, $G\alpha_{olf}$ is part of the $G\alpha_s$ protein family and binds to adenylyl cyclase 3 (AC3), activating it. As a result, the synthesis of cAMP in the olfactory cilia is upregulated, causing its rapid increase and the opening of cAMP-gated channels.

While photoreceptors and OSNs use CNG transduction channels, their respective channels have different subunit compositions. Rods and cones express heterotetramers consisting of the main A1 and A3 and the modulatory B1a and B3 subunits in 3:1 and 2:2 stoichiometric ratios, respectively. The olfactory CNG channel is a heterotetramer consisting of two units of the main A2 subunits and one each of the modulatory A4 and B1b subunits. Interestingly, the rod and the olfactory CNG channels express different splice variants of the same B1 subunit. In OSNs, the initial inward Na^+ and Ca^{2+} current generated by the opening of the CNG channel raises ciliary Ca^{2+} and opens a secondary ion channel, the Ca^{2+}-activated Cl^- channel Anoctamin 2. A high intraciliary Cl− maintained by the $Na^+/K^+/2Cl^-$ cotransporter 1 ensures a Cl^- efflux that further depolarizes the OSNs. This depolarization triggers the generation of action potentials, which further propagate along the axons, inducing glutamate release at synapses with the second-order neurons in the olfactory bulb. In photoreceptors, the transduction cascade upon stimulation does not ultimately generate action potentials in the receptor cell, but only a graded receptor potential that directly causes a change in neurotransmitter release.

13.7.9: Amplification

As for any other sensory modality, proper amplification of the signal is required to detect small stimuli, and the resulting high sensory sensitivity is critical for the survival and propagation of the species. Nature has reached the highest physically possible sensitivity in the case of rod photoreceptors that can produce a detectable electrical response to the absorption of a single photon. This impressive feat is achieved by employing a transduction cascade that allows tremendous signal amplification. During the ~50 ms active lifetime, a single photoactivated rhodopsin molecule activates ~20 transducins, producing an immediate 20-fold amplification. The following activation of PDE by transducin does not directly produce amplification, as each transducin has to bind to a PDE molecule to activate it. However, once activated, each PDE enzyme can hydrolyze thousands of cGMP molecules. Lastly, as the binding of cGMP to the CNG transduction channels is cooperative, a slight change in cGMP levels can reduce the number of cGMP molecules bound to the channel from 3 to 2. This results in channel closure and a sharp reduction in the transduction current, further enhancing the detection of photostimulation. Despite the similarities in the transduction cascades of rods and cones, the amplification in cone photoreceptors is substantially lower due to fine-tuning at several of the phototransduction steps. Interestingly, even though rod and cone visual pigments activate transducin with similar efficiencies, the lower thermal stability of the cone visual pigment results in higher intrinsic activity in cones compared to rods in darkness, effectively desensitizing the cones and shifting their function towards brighter daytime light conditions as shown in Figure $13.7.3C$.

Curiously, the activation of G_{olf} by the OR molecule does not result in amplification. Indeed, the dwell time of the odorant ligand on the OR appears to be very short and on a millisecond timescale. As a consequence, on average, this results in the activation of less than one G protein per activated receptor. As such, in contrast to phototransduction, where the lifetime of the activated

rhodopsin greatly influences the response size and kinetics, in OSNs the response depends more prominently on the coupling efficacy of downstream transduction components. At the same time, the odorant presence keeps the OR activated. OSNs employ a secondary amplification step on top of the cAMP transduction cascade to compensate for the lack of initial amplification at the G protein level. The activation of AC3 by Golf results in the synthesis of most likely hundreds of cAMP molecules, the opening of the CNG channels, which is followed by a unique secondary amplification based on excitatory Ca^{2+}-activated Cl^{-} channels in the cilia, as shown in Figure **13.7.4*B***. The Cl^{-} current carries up to 80% of the overall transduction current. Physiological experiments with pharmacological and genetic modulation of the Cl^{-} conductance indicate that the Cl^{-} channels set the length of the action potential train generated in response to odorant stimulation and to promote recognition of novel odorants.

A puzzling aspect of the secondary amplification step is why Cl^{-} is the charge-carrying ion, not Na^{+}. This could be achieved easily by increasing the expression level and/or the ion permeation and conductance of the olfactory CNG channel. Recent theoretical work hinted at two main advantages of Cl^{-}, instead of Na^{+}, as the charge carrier. As the external environment of cilia is the nasal mucus, currents will depend on the ion concentration in the mucus, which can be unstable. A current that depends on the intracellular ion concentration, as for Cl^{-} but not for Na^{+}, is much less dependent on the mucosal ion concentration. For instance, this could become an issue in the case of a cold with a runny nose or during swimming, when the mucus becomes diluted. The second advantage results from the “compromise” to increase the ciliary surface area, which has the expense of having a very small ciliary volume, in the femtoliter range. Even small ionic currents can lead to large changes in ion concentration and osmotic pressures in such small volumes. If the main charge carrier were Na^{+}, this would lead to a large increase (tens of mM). This would cause a large increase in osmotic pressure and prevent Ca^{2+} clearance *via* the olfactory Na^{+}/Ca^{2+}, K^{+} exchanger (see below) with greatly deleterious effects.

In contrast, high intracellular Cl^{-} is maintained throughout the OSN so that diffusion from the cell soma rapidly reverses its local depletion in the cilia upon ligand activation. Neither of these issues exist for photoreceptors as they are embedded in the interstitial fluid of the eye. Photoreceptors are sufficiently large, and their transduction currents are sufficiently small that ion concentration changes due to changes in transduction currents are relatively small. Nevertheless, rod photoreceptors undergo osmotically driven length changes upon light activation, an effect that is mitigated by the translocation of G protein subunits into the cytosol.

13.7.10: Receptor and G Protein Inactivation

Timely and effective transduction inactivation is critical for allowing sensory neurons to continue to detect stimuli with high temporal resolution. Equally important is to extract behaviorally relevant information from the presented stimuli. All active transduction components must be turned off in both photoreceptors and OSNs. The cyclic nucleotide level within the cells must be restored to the resting level before the sensory cell can be reset to the inactive state and become ready for subsequent activation, as shown in Figure **13.7.5**. In the case of photoreceptors, the identity of the step determining the overall kinetics of the photoresponse inactivation was the subject of intense research and debate over several decades. As the visual chromophore ligand is covalently attached to opsin, inactivation of the visual pigment could be extremely slow. Indeed, if left on its own, the active state of rhodopsin decays with a time constant of ~50 s. Its inactivation in photoreceptors is a two-step process, involving phosphorylation of the rhodopsin C-terminus by rhodopsin kinase (GRK1), which partially quenches its activity, followed by the binding of arrestin1, which completely inactivates the visual pigment (Panel A). Though the decay of the active state of cone pigment is significantly faster at ~2 s, this is still clearly too slow to enable the timely termination and reset of phototransduction. Thus, in both rods and cones, the visual pigments are inactivated by phosphorylation by rhodopsin kinase and the subsequent arrestin binding long before they decay spontaneously. The effective time constant of rod visual pigment inactivation is ~50 ms. The slowest step in the inactivation of rod phototransduction turned out to be the hydrolysis of GTP, which shuts off Tα, a reaction driven by the transducin GTPase activity and enhanced by a GTPase-activating protein (a GAP) complex consisting of Gβ5 and the membrane-anchoring protein R9AP. Inactivation of transducin results in its release from PDE, allowing the two PDE γ inhibitory subunits to resume their inhibition on this enzyme's two catalytic subunits (α and β). The kinetics of this reaction determine the overall kinetics of response inactivation in rod photoreceptors. In contrast, work from amphibian cones suggests that in cones, the photoresponse duration is Ca^{2+}-dependent and involves the quenching of the cone visual pigment.

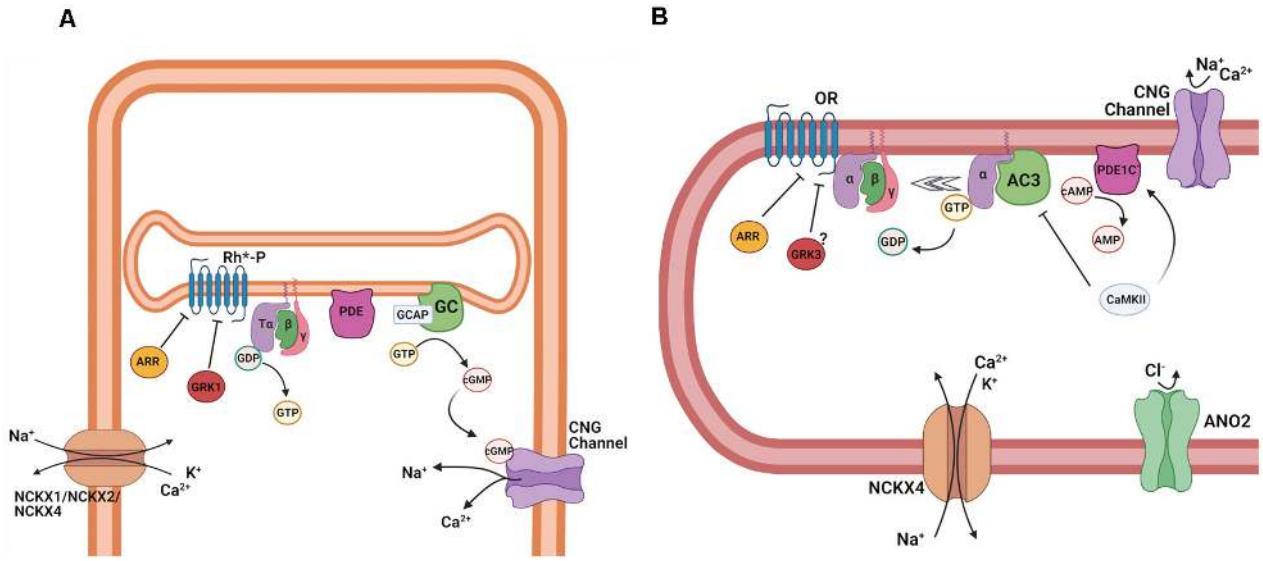


Figure 13.7.5: Termination of transduction cascade in rod photoreceptors and OSNs.

(A) Schematic representation of the termination of phototransduction in rods. Abbreviations: Phosphorylated light-activated rhodopsin (Rh*-P), arrestin (ARR), G protein-coupled receptor kinase 1 (GRK1), Tα, β, and γ subunits of the retinal G protein, transducin (T), guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), phosphodiesterase (PDE), guanosine monophosphate (GMP) and cyclic guanosine monophosphate (cGMP), and cyclic nucleotide-gated (CNG) channel, guanylate cyclase (GC), guanylate cyclase-activating protein (GCAP), cyclic nucleotide-gated (CNG) channels, K^+ dependent Na^+/Ca^{2+} exchanger 1, 2 and 4 (NCKX1, NCKX2, NCKX4).

(B) Schematic representation of the termination of the olfactory transduction cascade. Abbreviations: Olfactory receptor (OR), arrestin (ARR), G protein-coupled receptor kinase 3 (GRK3) $G\alpha_{olf}$, *β*, and *γ*, subunits of the olfactory G protein, guanosine-5′-triphosphate (GTP), guanosine-5′-diphosphate (GDP), adenylyl cyclase 3 (AC3), activated phosphodiesterase 1C (PDE1C), cyclic adenosine monophosphate (cAMP), adenosine monophosphate (AMP), Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII), cyclic nucleotide-gated channel (CNG); Ca^{2+}-activated Cl^-channel anoctamin 2 (ANO2), K^+-dependent Na^+/Ca^{2+} exchanger 4 (NCKX4). Images created with BioRender.com.

In OSNs, the inactivation by phosphorylation and arrestin is potentially not needed for the timely shutoff of the olfactory transduction cascade, due to the extremely short lifetime of the active ligand-bound receptor molecule. Early biochemical experiments suggested that OR phosphorylation does control cAMP kinetics. Still, it seems to play little, if any, role in controlling odorant-response kinetics for one particular OR, mOR-EG. It remains to be established whether this applies to all ORs or whether a subset of ORs is subject to phosphorylation and inactivation. β-arrestin interacts with ORs, mediating internalization during prolonged stimulation and altering adaptation to repetitive odor stimuli. Experiments on isolated human and rat OSNs suggested a role for protein kinases A (PKA) and C (PKC) in the termination of the olfactory response. Ca^{2+} imaging showed that the inhibition of PKA and PKC increases intracellular Ca^{2+} responses in the presence of odorant mixtures and blocks their termination after odorant stimulation ceases. While the inhibition of both PKA and PKC modulated the odor-induced intracellular Ca^{2+} increase in the human OSNs, only PKC and not PKA affected the Ca^{2+} response to odorants in rat OSNs, suggesting differences among species in the termination of the olfactory response.

The control of the lifetime of the olfactory G protein seems to be more complex and less well-understood compared to phototransduction. Ric-8B (resistant to inhibitors of cholinesterase-8B) has been identified as a GTP exchange factor (GEF) expressed in OSNs, which facilitates the exchange of GDP for GTP on $G\alpha_{olf}$ and its activation. Unusually, Ric-8B not only interacts with the G protein α-subunit but also with γ13, the olfactory γ-subunit. In a heterologous system, Ric-8B co-expression with olfactory transduction components can greatly increase cAMP production, suggesting it could modulate olfactory transduction. A knockout mouse for Ric-8B displays impaired olfactory behavior and, surprisingly, greatly reduced odorant responses. Ric-8B is localized primarily in the cell body and the dendritic knob of OSNs. Ric-8B knockout OSNs are devoid of $G\alpha_{olf}$, suggesting that this gene is needed for the stable expression of $G\alpha_{olf}$, and excludes addressing its potential role as a GEF in the odorant response. The Ric-8B knockout mice also display higher OSN cell death. Regulators of G protein signaling (RGS) are

GAPs that modulate the lifetime of an activated G protein as described above. RGS2, instead of functioning as a GAP, directly inhibits AC3 to control the size of the odorant response. However, inconsistent and contradictory data on RGS2 and RGS3 expression and their roles in OSNs suggest that more research is needed.

13.7.11: Adaptation

Adaptation plays a critical role in the capacity of our sensory neurons to remain able to detect stimuli above the background in a complex and rapidly changing environment. For instance, in constant light conditions, the dynamic range for rods and cones is only 100-fold, spanning a range from threshold stimulation to saturation as shown in Figure **13.7.3*C***. However, due to light adaptation, photoreceptors can shift their functional range over a wide range of light conditions, ranging from cloudy night to sunrise for rods, and starry night to bright sunny day for cones. Thus, using the adaptation of individual photoreceptors, the visual system can remain responsive to stimuli over a wide range of light conditions. In contrast, the ability of OSNs to adapt is somewhat limited even at modest levels of background odorant. Nevertheless, increasing concentrations of the same odorant can recruit less sensitive ORs, and therefore less sensitive OSNs, preserving its perception at higher concentrations and ensuring that it reports the presence of that odorant to the brain.

In both types of sensory neurons, adaptation is mediated by a change in Ca^{2+} upon stimulation. This change is sensed by several Ca^{2+}-binding proteins that trigger a negative feedback on the vision and olfaction transduction cascades by modulating several of their steps. In the outer segments of rods and cones and olfactory cilia, Ca^{2+} levels are controlled by the balance of influx *via* the CNG channels, whose current is carried in part by Ca^{2+}, and efflux *via* Na^+/Ca^{2+}, K^+ exchangers (NCKXs) that use the electrochemical gradient for Na^+ and K^+ to extrude Ca^{2+} as shown in Figure **13.7.5**. In rods (Panel A), this task is accomplished by rod-specific NCKX1, whereas cones employ two separate exchangers, NCKX2 and NCKX4. At rest, both in darkness and in steady-state light, the influx of Ca^{2+} is matched with its extrusion and, as a result, the level of free Ca^{2+} in the outer segments is maintained constant. Upon photostimulation, the transduction cascade is activated, resulting in the depletion of cGMP, closure of CNG channels, and reduction in the influx of Ca^{2+} into the outer segments. However, Ca^{2+} extrusion by the Na^+/Ca^{2+} and the K^+ exchangers continues for a while and, as a result, the level of Ca^{2+} in the outer segments declines. Direct Ca^{2+} measurements in amphibian photoreceptors indicate a dynamic range from 670 to 30 nM in rods and 400–5 nM in cones, in darkness and bright light, respectively.

The light-driven decline in Ca^{2+} causes its release from several Ca^{2+}-binding proteins. The dominant Ca^{2+}-dependent feedback mechanism in both rods and cones controls the synthesis of cGMP by membrane-bound GC *via* a pair of GC-activating proteins (GCAPs)—GCAP1 and GCAP2. When Ca^{2+} in the outer segments is high, Ca^{2+}-bound GCAPs bind to and partially inhibit the activity of GC. Upon photoactivation and the decline in Ca^{2+}, GCAPs become Ca^{2+}-free and release from GC, resulting in the upregulation of cGMP synthesis, which restores the dark current after photostimulation and modulates the activation of the transduction cascade in the presence of background light. Another mechanism by which Ca^{2+} modulates phototransduction involves the Ca^{2+}-binding protein recoverin. As a GCAP, recoverin is a member of the EF-hand protein family. When bound to Ca^{2+} in darkness, it inhibits rhodopsin kinase, thus slowing down the inactivation of the visual pigment. When the photoreceptors are activated and Ca^{2+} declines, it is released from the recovery. This then dissociates from rhodopsin kinase and relieves its inhibition. This enhances the phosphorylation of visual pigments and accelerates their inactivation, effectively reducing the background light's activation of the transduction cascade. Finally, direct modulation of the CNG channels has also been suggested. However, in the case of rods, such modulation appears to play a marginal, at best, role and is not mediated by the Ca^{2+}-binding protein calmodulin. In zebrafish cones, the modulation of the CNG channels appears to play a more substantial role and is mediated by the Ca^{2+}-binding protein CNG-modulin. It is still unclear whether the mammalian homolog of CNG modulin, EML1, plays a similar role in mammalian cones.

Adaptation in OSNs is less well understood than phototransduction. Early data, mostly of biochemical nature or obtained from heterologously expressed proteins of interest, suggested three main molecular targets for adaptation. All three of them are mediated by the Ca^{2+} influx during the odorant response: Ca^{2+}/calmodulin-mediated desensitization of the olfactory CNG channel to close the channel even in the presence of high cAMP; phosphorylation *via* CaM-kinase 2 of AC3 to reduce the rate of cAMP production; and Ca^{2+}-mediated upregulation of phosphodiesterase 1C, which is expressed in olfactory cilia, and is assumed to degrade cAMP to AMP to terminate the response. Follow-up experiments using recordings from OSNs all indicate that none of these mechanisms play as prominently or as originally thought in transduction. A mouse with a mutation in the CNGB1b channel subunit that entirely prevents desensitization by Ca^{2+}, surprisingly displays normal olfactory adaptation but instead shows a delayed response

termination, suggesting that Ca^{2+}/calmodulin-mediated desensitization of the CNG channel speeds up response termination. A mouse model with a mutation in AC3 that prevents phosphorylation does not show a discernible phenotype of the olfactory response. However, it might be possible that other, unknown phosphorylation sites in AC3 might be important. Finally, a knockout mouse for PCE1C has no deficits in response termination but shows much-reduced response amplitudes for unclear reasons. This begs the obvious question of what the role of PDE1C might be and what might occur to cAMP generated during the odorant response. For the latter, an interesting option is that cAMP diffuses out of the cilia into the cell body to reduce ciliary cAMP, allowing OSNs to recover from stimulation. One reasonably understood aspect is NCKX4, the Ca^{2+} exchanger in OSNs that is required to lower intraciliary Ca^{2+} during and after odorant stimulation, allowing the transduction cascade to recover from adaptation.

13.7.12: Diseases

Disorders affecting photoreceptors are among the leading causes of blindness in the human population. One of the prevalent visual disorders, called retinitis pigmentosa, is a complex disease caused by a wide range of mutations in photoreceptors. Many of these mutations affect the rod's visual pigment's expression, structure, and function. Because of the very high expression of opsin in the outer segments of rods, this protein plays not only a functional role. Still, it is also critical for properly forming the outer segment itself. As a result, mutations affecting the expression, folding, or targeting of opsin to the rod outer segments cause gradual degeneration of the rods. Other genes implicated in rod dysfunction and degeneration include those for phosphodiesterase (e.g., rd1, rd10; the CNG channels A and B subunits (channelopathies;), GC, and GCAPs. Another diverse set of visual disorders is caused by abnormal chromophore production or supply to photoreceptors, limiting the ability to detect light and leading to degeneration. Notably, the visual system's efficiency in producing chromophore seems to decline with age, which may result in poor rod function in dim light even in normally aging adults. It is also an early indicator of age-related macular degeneration. This devastating blinding disorder affects the function of cones in the central retina, which are responsible for acute vision and color discrimination. Interestingly, rods and cones seem to coexist synergistically in the retina, and diseases caused by rod-specific mutations that result in rod degeneration eventually lead to the loss of cones and central vision. Thus, considerable efforts are currently focused on developing methods for preserving rods even when they are not functional to protect daytime cone-driven vision. Because the eye is a relatively accessible organ, novel therapeutic approaches for vision protection and restoration have led the field, with successful gene and stem cell therapy examples in experimental and clinical trial phases.

Compared to vision, in olfactory transduction, very few mutations in transduction components are known that lead to deleterious effects. Several aspects might account for this. Mutations causing a partial reduction of olfaction might go unnoticed in the human population, as very little systematic olfactory testing is done. OSNs regenerate throughout life and only have a lifespan of a few weeks. Hence, as seen in photoreceptors, any slow degeneration might not manifest in that time window. In a screen of families with congenital anosmia, no potentially causative mutations were found in three main transduction proteins ($G_{\alpha olf}$, CNGA2, and AC3), with these genes also being under purifying selection. An interesting exception is patients suffering from retinitis pigmentosa, which is caused by mutations in the gene encoding the CNGB1 subunit expressed in rods and OSNs. Those patients, identified because of their visual function decline, were found to be hyposmic or anosmic when tested for their olfactory ability. If congenital anosmia is considered a relatively rare and little-understood condition, specific anosmias are better known and frequently detected, manifesting in the inability to detect certain odorants. This is the olfactory equivalent of color blindness caused by known OR mutations.

Arguably, the most common causes of smell loss are events that lead to the destruction of the olfactory epithelium and/or the olfactory nerves connecting it to the central nervous system (CNS). These events include head or face trauma, inhalation of toxic chemicals, or viral infection (such as SARS-CoV-2), and neurodegenerative diseases such as Alzheimer’s and Parkinson’s disease. In the former, the origin of the smell disorder can be traced to the periphery, the olfactory epithelium. In the case of neurodegenerative diseases, it has been thought that olfactory dysfunction originates centrally in the CNS. Still, it is becoming clearer now that peripheral olfaction can be affected in these cases as well, although the respective mechanisms have not been fully elucidated.

13.7.13: Summary

This chapter explores the molecular and cellular mechanisms by which specialized sensory neurons convert external stimuli into electrical signals. Focusing on vertebrate photoreceptors (rods and cones) and olfactory sensory neurons (OSNs), the chapter

highlights how distinct structures and signaling cascades are optimized for the detection of light and odorants, respectively, while also discussing common themes such as signal amplification, inactivation, and adaptation.

Specialized Cellular Architecture

- **Photoreceptors:**
  Rods and cones are highly polarized, ciliary neurons that use their modified outer segments for light detection. In rods, the outer segment contains a stack of approximately 1,000 membrane disks densely packed with visual pigments, where each disk is enclosed by the plasma membrane. Cones, on the other hand, feature disk structures formed by invaginations of the plasma membrane. This organization, coupled with the precise alignment of the outer segments to incoming light, maximizes photon absorption.

- **Olfactory Sensory Neurons:**
  OSNs extend multiple (approximately 20) slender, elongated cilia from their dendritic knobs into the mucus layer covering the olfactory epithelium. These cilia provide a large surface area for the expression of olfactory receptors (ORs), which are responsible for detecting a vast array of odorant molecules.

Signal Detection and Transduction Mechanisms

- **Detection by G Protein-Coupled Receptors (GPCRs):**
  Both photoreceptors and OSNs rely on GPCRs to detect stimuli. In photoreceptors, visual pigments consist of an opsin protein covalently bound to a chromophore (typically 11-cis-retinal). Absorption of light triggers isomerization of the chromophore to all-trans-retinal, inducing conformational changes in opsin that activate the transduction cascade.
  In OSNs, odorant molecules bind noncovalently to olfactory receptors located on the ciliary membrane. Despite similarities to other GPCRs, these receptors are part of a complex network that discriminates among a wide variety of odorants.

- **Transduction Cascades:**
  In photoreceptors, activated visual pigment (e.g., photoactivated rhodopsin) engages a specific heterotrimeric G protein (transducin). This leads to the exchange of GDP for GTP on the Gα subunit, which then dissociates and activates cGMP phosphodiesterase (PDE). The resultant decrease in cGMP levels causes the closure of cyclic nucleotide-gated (CNG) channels, reducing the inward current and hyperpolarizing the cell.
  In contrast, odorant binding in OSNs activates the olfactory G protein (Golf). The activated Gαolf stimulates adenylyl cyclase 3 (AC3), leading to an increase in cAMP levels. Elevated cAMP opens CNG channels, allowing an influx of cations (primarily Na^+ and Ca^{2+}) that depolarize the neuron and trigger action potentials. Notably, OSNs employ an additional amplification step via Ca^{2+}-activated Cl^- channels, which enhance the depolarization.

Signal Amplification and Sensitivity

- **Photoreceptor Amplification:**
  Phototransduction in rods is extraordinarily sensitive—capable of detecting single photons. This sensitivity is achieved by a cascade in which one activated rhodopsin molecule activates multiple transducin molecules, and each activated PDE hydrolyzes thousands of cGMP molecules. Furthermore, the cooperative nature of cGMP binding to CNG channels ensures that even small changes in cGMP levels produce significant changes in channel conductance and membrane potential.

- **Olfactory Sensitivity:**
  Although OSNs do not achieve the single-molecule sensitivity of rods, they are finely tuned to detect odorants even at nanomolar concentrations. The initial G protein activation by odorant-bound receptors may be modest; however, the robust secondary amplification via cAMP synthesis and Ca^{2+}-activated Cl^- channel activation ensures that even minimal odorant binding leads to a significant electrical response. In addition, the large number of OR genes and the diverse array of OSNs contribute to the broad detection and discrimination of odorant molecules.

Inactivation and Adaptation Mechanisms

- **Termination of the Signal:**
  For both photoreceptors and OSNs, timely inactivation of the signaling cascade is crucial to prevent continuous activation and

to allow rapid recovery for subsequent stimuli. In photoreceptors, active visual pigments are rapidly inactivated through phosphorylation by rhodopsin kinase followed by arrestin binding. GTP hydrolysis on transducin, accelerated by a GTPase-activating protein complex, further ensures rapid signal termination.
In OSNs, although receptor phosphorylation and arrestin-mediated internalization may play a role, the transient binding of odorants typically results in a short-lived active state. Additionally, proteins such as Ric-8B and regulators of G protein signaling (RGS) modulate the duration of the odorant response, although the precise mechanisms remain less well defined than in photoreceptors.

- **Adaptation Processes:**
  Adaptation allows sensory neurons to adjust their sensitivity based on the prevailing background levels of stimulation. In photoreceptors, adaptation is largely mediated by changes in intracellular Ca^{2+} levels. When light reduces cGMP and closes CNG channels, the subsequent decline in Ca^{2+} triggers feedback mechanisms via GCAPs and recoverin that restore cGMP synthesis and accelerate pigment inactivation. OSNs also utilize Ca^{2+}-dependent feedback mechanisms, such as Ca^{2+}/calmodulin-mediated modulation of CNG channel sensitivity and regulation of cAMP levels, to fine-tune their responses and maintain dynamic sensitivity in fluctuating odor environments.

Implications for Sensory Function and Disease

- **Functional Relevance:**
  The precise control of signal amplification, inactivation, and adaptation in photoreceptors enables the visual system to operate over an enormous range of light intensities—from starlight to bright daylight. Similarly, the complex transduction mechanisms in OSNs allow for the discrimination of an immense variety of odorants despite the inherently low stimulus concentrations.
- **Pathological Considerations:**
  Mutations or dysfunctions in key components of these signaling pathways can lead to sensory disorders. For example, mutations in visual pigment or CNG channel components are associated with retinitis pigmentosa and other forms of degenerative blindness. In the olfactory system, while genetic defects are less commonly identified, disruptions due to environmental damage or neurodegenerative diseases can result in anosmia (loss of smell) or hyposmia (reduced sense of smell).

In summary, this chapter integrates the structural, biochemical, and electrophysiological aspects of sensory transduction in photoreceptors and OSNs. It highlights how specialized cellular architectures and finely tuned signaling cascades enable these neurons to convert light and odorant stimuli into robust electrical signals, and it discusses the mechanisms that ensure both high sensitivity and rapid adaptation. Understanding these processes not only provides insights into the fundamental principles of sensory biology but also informs the development of therapeutic approaches for sensory disorders.

This page titled 13.7: Signal Transduction - Vision and Olfaction is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

13.8: Signal Transduction - Taste (Gustation)

Learning Goals (ChatGPT o3-mini)

- **Understand Taste GPCR Classification and Structure:**
  - Describe the two main families of taste GPCRs (Type I and Type II), including the heterodimeric TAS1Rs for sweet and umami tastes and the monomeric receptors for bitter and kokumi sensations.
  - Identify key structural domains (e.g., Venus flytrap domain, cysteine-rich domain, transmembrane domain) and explain how these contribute to receptor function and ligand specificity.
- **Explain Taste Signal Transduction Pathways:**
  - Outline the intracellular signaling cascades activated by taste GPCRs, emphasizing the roles of heterotrimeric G proteins (e.g., Gα-gustducin, Gβγ), PLCβ2, IP3-mediated Ca^{2+} release, and subsequent events leading to ATP release and neurotransmission.
  - Compare the mechanisms by which sweet, umami, bitter, and kokumi receptors transduce signals from tastants to the gustatory cortex.
- **Analyze Receptor–Ligand Interactions:**
  - Discuss how orthosteric agonists bind to specific receptor domains (e.g., VFT in TAS1Rs) to induce conformational changes necessary for receptor activation.
  - Evaluate the impact of allosteric modulators—including positive allosteric modulators (PAMs) and negative allosteric modulators (NAMs)—on receptor conformation and signal transduction, using examples such as lactisole, NHDC, and various synthetic peptides.
- **Explore Molecular Modeling and Mutagenesis Approaches:**
  - Explain how heterologous expression, site-directed mutagenesis, and computational docking (e.g., using AlphaFold and AutoDock Vina) contribute to identifying ligand binding sites and deciphering the structural basis of receptor activation and modulation.
  - Interpret data from molecular models (e.g., interactive iCn3D structures) to determine how key residues in receptor domains influence ligand binding and receptor function.
- **Examine Receptor Trafficking and Cell Surface Expression:**
  - Describe the importance of proper receptor folding, dimerization, and trafficking in ensuring effective cell surface expression of taste GPCRs, with an emphasis on the roles of specific extracellular and transmembrane regions.
- **Understand Bitter Taste Receptor Diversity and Specificity:**
  - Compare the ligand-binding properties of broadly tuned versus narrowly tuned bitter taste receptors, and discuss how receptor sequence variability influences agonist specificity and affinity.
  - Evaluate the pharmacological strategies used to identify bitter receptor inhibitors and the implications for reducing off-target bitterness in foods and drugs.
- **Investigate the Role of the Calcium-Sensing Receptor (CaSR) in Kokumi Sensation:**
  - Explain how CaSR, a class C GPCR, mediates kokumi taste through interactions with various ligands (e.g., calcium ions, amino acids, glutathione) and discuss its downstream signaling pathways.
  - Analyze how allosteric modulation of CaSR by endogenous and synthetic compounds contributes to flavor enhancement and the potential extraoral roles of kokumi peptides.
- **Integrate Physiological and Pharmacological Perspectives:**
  - Relate the molecular mechanisms of taste GPCR activation to physiological taste perception and ingestion decisions.

- Discuss the current challenges in taste receptor research, including difficulties in crystallizing taste GPCRs, and explore how emerging structural biology techniques (e.g., cryo-EM and serial femtosecond crystallography) may overcome these obstacles.

By achieving these goals, students will gain a comprehensive understanding of the structure, function, and pharmacology of taste GPCRs, enabling them to appreciate how these receptors contribute to taste perception and how they can be targeted to modulate taste responses in food science and medicine.

This chapter section is taken in large part from Ahmad and Dalziel. Front. Pharmacol., 30 November 2020 | https://doi.org/10.3389/fphar.2020.587664. Creative Commons Attribution License (CC BY)

13.8.1: G Protein-Coupled Receptors in Taste Physiology and Pharmacology

Heterotrimeric G protein-coupled receptors (GPCRs) comprise the largest receptor family in mammals and regulate most physiological functions. Besides mediating the sensory modalities of olfaction and vision, GPCRs also transduce signals for three basic taste qualities: sweet, umami (savory taste), and bitter, as well as the flavor sensation kokumi. Taste GPCRs reside in specialized taste receptor cells (TRCs) within taste buds. Type I taste GPCRs (TAS1R) form heterodimeric complexes that function as sweet (TAS1R2/TAS1R3) or umami (TAS1R1/TAS1R3) taste receptors, whereas Type II are monomeric bitter taste receptors or kokumi/calcium-sensing receptors. Sweet, umami, and kokumi receptors share structural similarities in containing multiple agonist binding sites with pronounced selectivity. In contrast, most bitter receptors contain a single binding site broadly tuned to a diverse array of bitter ligands in a non-selective manner. Tastant binding to the receptor activates downstream secondary messenger pathways, leading to depolarization and increased intracellular calcium in TRCs, innervating the brain's gustatory cortex. Despite recent advances in understanding the relationship between agonist binding and the conformational changes required for receptor activation, several major challenges and questions remain in taste GPCR biology, discussed in the present review. In recent years, intensive integrative approaches combining heterologous expression, mutagenesis, and homology modeling have provided insight regarding agonist binding site locations and orthosteric and allosteric modulation molecular mechanisms. In addition, studies based on transgenic mice, utilizing either global or conditional knock-out strategies, have provided insights into taste receptor signal transduction mechanisms and their roles in physiology. However, the need for more functional studies in a physiological context is apparent. A crystal structure of taste receptors would enhance it for a more complete picture of their pharmacological mechanisms.

13.8.2: Introduction

G protein-coupled receptors (GPCRs) are eukaryotes' largest and most diverse group of membrane receptors. A wide variety of ligands activate them in the form of light energy, lipids, sugars, peptides, and proteins, which convey information from the outside environment into the cell to mediate their corresponding functional responses. The conformational changes of GPCRs upon ligand binding initiate a series of biochemical reactions within the cell. These intracellular reactions regulate sensory functions of smell, taste, and vision, as well as a wide variety of physiological processes such as secretion, neurotransmission, metabolism, cellular differentiation, inflammation, and immune responses. Taste is one of the most important sensations in human life, enabling us to perceive different tastes from the diverse range of food available in nature, and is a major determinant of our ingestion decisions.

The anatomical units of taste detection are taste receptor cells (TRCs) assembled into taste buds distributed across different tongue and palate epithelium papillae. Taste processing is first achieved at the level of TRCs that specific tastants activate. They transmit information *via* sensory afferent fibers to the gustatory cortex in the brain for taste perception, as shown in Figure 13.8.1. Three different morphologic subtypes of TRCs in taste buds sense the different tastes we perceive. Type I glial-like cells detect salty taste, while type II cells expressing GPCRs detect sweet, umami, and bitter tastes. Type III cells sense sour stimuli.

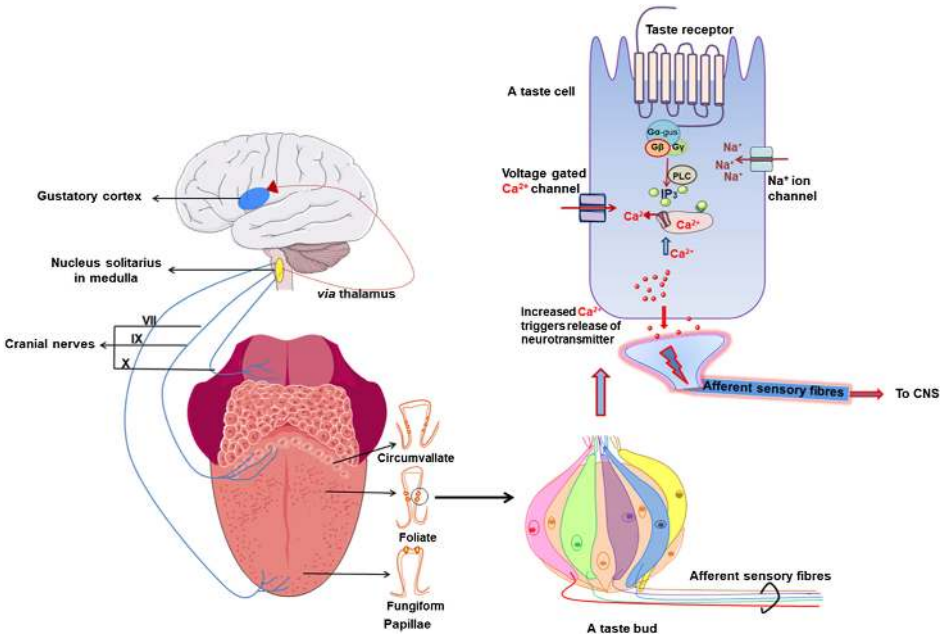


Figure 13.8.1: A schematic diagram shows taste signal transmission between the tongue and brain. Taste buds in different places on the tongue and palate papillae contain taste receptor cells (TRC), which contain taste G protein-coupled receptors (GPCRs). The left side shows how afferent nerves transmit a signal to the gustatory cortex in the brain via cranial/glossopharyngeal nerves. The right side shows taste buds with taste TRCs and a simplified signal transduction pathway of taste receptors. Taste GPCRs are activated by a tastant that recruits a specific G protein that further induces intracellular calcium release (created with BioRender.com).

Type 1 taste GPCRs transduce sweet and umami stimuli, while Type 2 taste GPCRs sense bitter taste. See Figure 13.8.2 and Table 13.8.1 for details. The more recently described kokumi sensation is mediated by another GPCR, the calcium-sensing receptor (CaSR). Taste GPCRs are activated by specific taste ligands present in foods and recruit G proteins to activate downstream signaling effectors, as shown in Figure 13.8.3 and Table13.8.1.

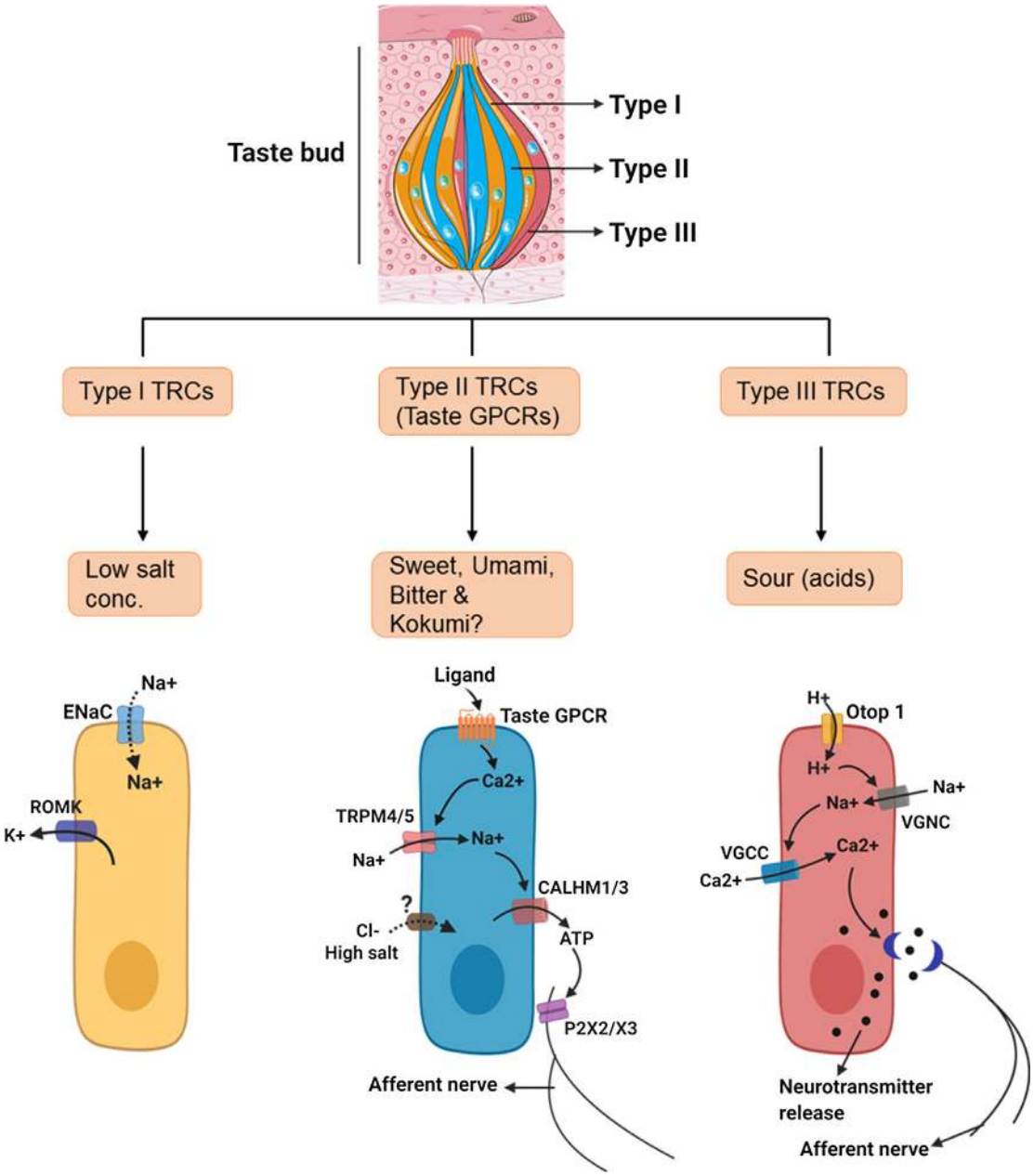


Figure 13.8.2: (created with BioRender.com).

TABLE 1 | Taste GPCRs classification and their downstream signaling regulators.

Taste/flavor sensation	Sweet	Umami	Bitter	Kokumi
GPCRs	TAS1R2 + TAS1R3	TAS1R1 + TAS1R3	TAS2Rs	Calcium sensing receptor
Class of GPCR	Class C	Class C	Class A	Class C
Type	Type 1 taste GPCRs	Type 1 taste GPCRs	Type 2 taste GPCRs	ND
Oligomerization	Heterodimer	Heterodimer	Monomer	Homodimer
GαSubunit	Gα-gustducin	Gα-gustducin	Gα-gustducin	Gq/11, Gi/o, G12/13, Gs
Signaling pathway	Gβγ-PLCβ2-TRPM5 pathway-semaphorin 7A	Gβγ-PLCβ2-TRPM5 pathway	Gβγ-PLCβ2-TRPM5 pathway-semaphorin 3A	Gβγ-PLCβ2-TRPM5 pathway
Ligand binding sites	VFT; TMD	VFT, TMD	ND	VFT, TMD
Potent ligands	Sucrose, aspartame, sucralose, cyclamate	L-amino acids, Gultamate	Peptides, alkaloids, flavonoids	Ca^{2+}, divalent and trivalent cations, Glutathione, *γ*-glutamyl peptides, polyamines, amino acids
Antagonists/NAMs	Lactisole, gymnemic acid	Lactisole, clofibric acid	GIV3727, probenecid	NPS2143
Food source	Sugar/sweets, dairy products	Savoury food, fermented dairy/meat/fish/chicken broth, mushroom, tomato	Wine, tea, coffee, cheese, broccoli	Garlic/onion meat/fish/broth/mushroom, soy sauce

TAS2R, bitter taste receptors; GPCR, G protein-coupled receptors; VFT, extracellular Venus flytrap domain; TMD, seven transmembrane domain; ND, not defined.

Table 13.8.1: Taste GPCRs classification and their downstream signaling regulators.

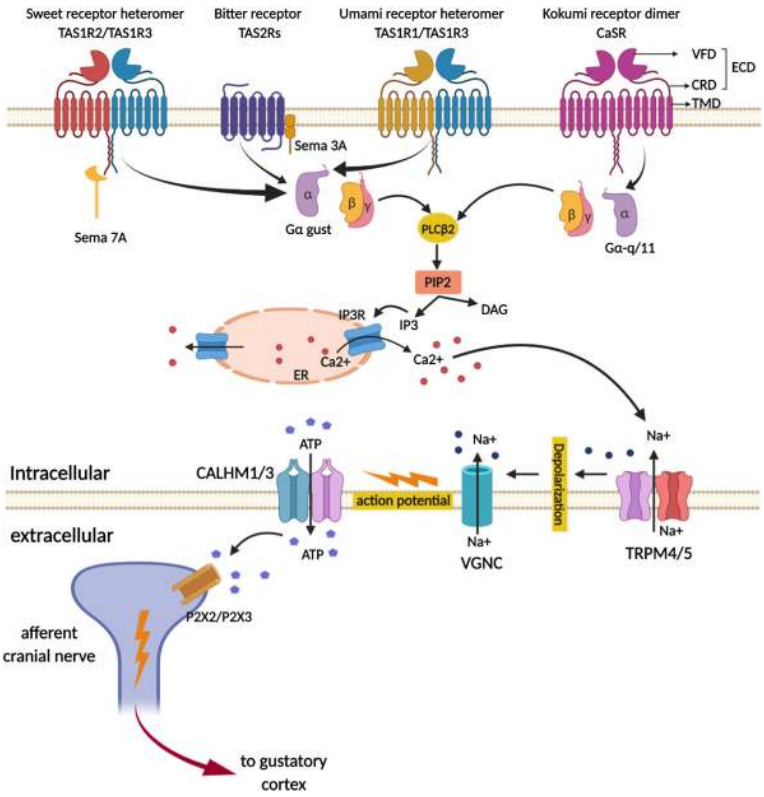


Figure 13.8.3: Schematic representation of signal transduction pathway of sweet, umami, bitter, and **kokumi-calcium sensing receptors (CaSR)** in taste receptor cells on the tongue.

Ligand-induced stimulation of the sweet (TAS1R2/TAS1R3), umami (TAS1R1/TAS1R3), bitter receptors (TAS2Rs), and kokumi sensation expressed in type II taste cells within taste buds activates a trimeric G protein composed of α-gustducin (Gα-gust) in sweet, umami, bitter, and Gα-q/11 in kokumi-receptor and a complex consisting of Gβγ proteins. The released Gβγ-complex activates phospholipase C isoform β2 (PLCβ2), which then induces the production of inositol 1,4,5-trisphosphate (IP3) and diacylglycerol (DAG); the second messenger IP3, in turn, activates the IP3 receptor (IP3R), an intracellular ion channel that allows Ca^{2+} release from the intracellular endoplasmic reticulum (ER store). An increase in intracellular Ca^{2+} then activates the complex

of transient receptor potential cation channel subfamily M member 4 and 5 (TRPM4/5), which are plasma membrane-localized sodium-selective channels, leading to depolarization and subsequent activation of voltage-gated sodium channels (VGSC). The combined action of increased Ca^{2+} and membrane depolarization activates the complex of calcium homeostasis modulator 1 and 3(CALHM1/3) and pannexin1 channels, thus releasing the neurotransmitter ATP. Increased ATP, in turn, activates P2X ionotropic purinergic receptors 2 and 3 (P2X2/P2X3) on the afferent cranial nerve, generating an action potential that subsequently signals to the gustatory cortex for sensory perception. Besides well-known taste GPCR pathways, connecting proteins semaphorin 7A (Sem 7A) and 3A (Sem 3A) are depicted in close contact with sweet and bitter receptors as they provide instructive signals that fine-tune to sweet or bitter ganglion neurons, respectively. VFT, Venus flytrap domain; CRD, cystine-rich domain; ECD, extracellular domain. (created with BioRender.com).

In this review, we will first explore the basic architecture of the gustatory sensory system and its peripheral signal transmission. Then we will discuss taste GPCR signal transduction mechanisms for the different taste modalities, their molecular structure, and the conformational changes following orthosteric/allosteric binding of endogenous and food-derived ligands.

13.8.3: Taste Buds and Neural Transmission

In mammals, taste buds on the tongue comprise 50–100 elongated epithelial cells and a small number of proliferative basal cells. Ultrastructural studies and gene expression patterns with cell function reveal three distinct anatomical types of TRCs within each taste bud: Type I, Type II, and Type III. See Figure 13.8.2: and Figure 13.8.2 for details.

TABLE 2 | Summary of taste receptor cell characteristics.

Taste receptor cells	Type I	Type II	Type III
Cell type	50% of the total population	20–30%	5–10%
Taste responses	Low salt taste amiloride sensitive response	Sweet, biiter, umami and probably kokumi flavor response (?) via GPCRs	Sour response via otopterin 1
Morphology	Spindle shaped with long brush like microvilli (1–2 mm), no synapse	Long and slender short microvilli, no synapse	Single large microvillus, have synaptic contact with afferent nerve
Other functions	Support function, ion redistribution, neurotransmitter clearance	High salt taste amiloride insensitive response?	Probable salt response?
Marker proteins	GLAST, K^+ channel (ROMK)	CALHM1, CALHM3, pannexin, connexins	Kir2.1, PDK2L1, SNAP25, synapsinII, NCAM

GLAST, glutamate aspartate transporter; ROMK, renal outer medullary potassium channel; PDK2L1, polycystin 2 like 1, transient receptor potential cation channel; CALHM, calcium homeostasis modulator; Kir2.1, inward rectifier K(+) channel; NCAM, neural cell adhesion molecule; SNAP25, synaptosome associated protein 25 kDa.

Table 13.8.2 Summary of taste receptor cell characteristics.

Type II TRCs express either sweet, umami, or bitter taste receptors at their cell surface. These receptors share some commonality with their signal transduction mechanisms intrinsic to TRCs. Taste GPCRs (sweet, umami, and bitter) couple to heterotrimeric G proteins that include Gα-gustducin, Gβ3, and Gγ13 and initiate a series of signal transduction cascades involving activation of phospholipase C-β2 (PLCB2), production of inositol-1,4,5-trisphosphate (IP3), and IP3-dependent Ca^{2+} release from the endoplasmic reticulum (ER) via the IP3 receptor (IP3R). The increased intracellular $[Ca^{2+}]_i$ then activates the transient receptor potential cation channel subfamily M member 4 and 5 (TRPM4/5) in the basolateral plasma membrane, leading to membrane depolarization that triggers Na^+ action potential firing, and depolarization-induced release of ATP. In turn, ATP acts as the primary neurotransmitter, stimulating purinergic receptors 2 and 3 (P2X2 and P2X3) on afferent cranial nerves, whose activation triggers an action potential that subsequently activates the gustatory cortex in the brain. α-gustducin is a distinct G protein selectively expressed in ∼30% of type II TRCs and shares 80% identity with retinal protein α-transducin and is a key contributor to signal transduction for sweet and bitter taste receptors.

An important aspect of taste transduction is how ATP signaling is conducted. Recent studies have discovered that calcium homeostasis modulators 1 and 3 (CALHM1/3) are enriched in type II TRCs, where they interact and form a functional complex. Their genetic deletion abolishes responses to sweet, bitter, and umami tastes, supporting the requirement of the CALHM1/3 complex as an ATP release channel for the GPCR-mediated tastes.

New information has provided insight into how specific taste qualities are fine-tuned to recognize their partner ganglionic neurons in the brain. Lee et al discovered semaphorin proteins, 7A and 3A, as the physical links between sweet and bitter TRCs, respectively, and their partner ganglion neurons in the brain. It remains to be determined what physical links exist between umami

TRCs and their corresponding neurons in the brain. Delineating the underlying molecular basis for this interaction would provide a further understanding of purinergic transmission in the taste system. In addition, whether these mechanisms are relevant for kokumi sensation has not been investigated, despite CaSR having distinct expression in TRCs and significant functional synergy with other prominent taste qualities (sweet, umami, and salty). Moreover, there is still debate regarding the recognition of kokumi as a sixth taste entity; consequently, the calcium-sensing receptor (CaSR) is not yet included in the nomenclature for any subtypes of taste GPCRs, although it would best fit with Type 1 taste receptors.

13.8.4: Type 1 Taste G Protein-Coupled Receptors (Sweet and Umami)

Type 1 taste receptors (TAS1Rs) belong to the class C GPCRs, which possess a large N-terminal extracellular domain (ECD) fused to the heptahelical seven transmembrane domain (TMD). The ECD is further divided into two ligand-binding domains (LBD1 and LBD2), having a bilobed structure called a Venus flytrap domain (VFT) due to its resemblance to this shape. Except for $GABA_B$ receptors, a cysteine-rich domain (CRD) connects the VFT to the TMD.

In contrast to other receptors from this class C of GPCRs, such as the metabotropic glutamate receptor (mGluR) or γ-aminobutyric acid type B receptors (GABABRs), which function as homo- or heterodimers, respectively, the TAS1Rs function as obligatory heterodimers. The distinct expression pattern of TAS1R1 and TAS1R2 in different subsets of murine cells led to the idea that they could detect two different taste profiles. However, following the discovery of the TAS1R3 subtype, it was clear that when TAS1R1 heterodimerizes with TAS1R2, the receptor detects sweet taste substances. On the other hand, when heterodimerized with TAS1R3 (TAS1R1/TAS1R3), it is responsible for detecting umami or amino acid tastes. Please refer to Figure 13.8.4*A* for the basic structure of sweet and umami receptors.

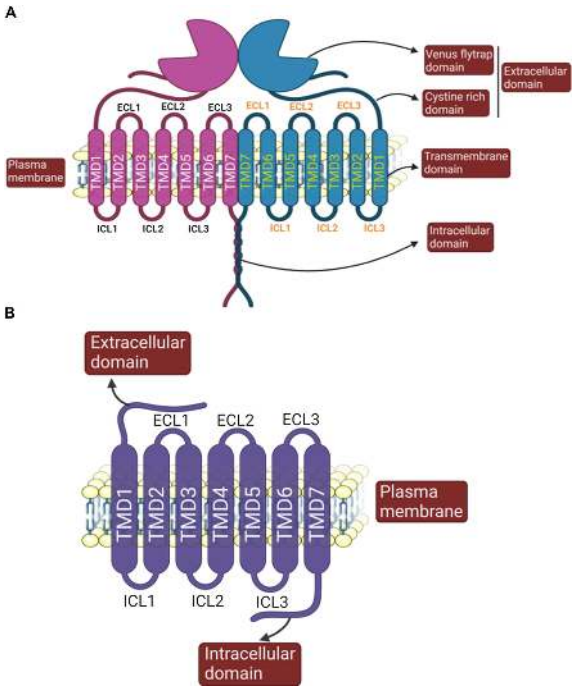


Figure 13.8.4: depicting basic structural features of sweet/umami/kokumi dimeric receptor **(A)** and monomeric bitter receptor **(B)** (created with BioRender.com).

Figure 13.8.5 shows an interactive iCn3D model of Taste receptor type 1 member 2 (TAS1R2) AlphaFold model (uniprot Q8TE23)

Figure 13.8.5: Taste receptor type 1 member 2 (TAS1R2) AlphaFold model (uniprot Q8TE23). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...mkhYb2uKoR68b9

The gray is the predicted transmembrane helices. The cyan is the intracellular domain. The blue is the extracellular domain. The predicted model of the structure has high confidence except for the yellow/orange at the distal end of the extracellular domain. Key residues in the ligand binding domain are shown as sticks, CPK colors, and labeled.

13.8.5: Sweet Taste Signal Transduction Mechanisms

The TAS1R2/TAS1R3 receptor recognizes a wide variety of sweet substances, including natural sugars, artificial sweeteners, amino acids, and proteins, as summarized in Table 13.8.3. This was demonstrated in studies using heterologous expression systems as well as knockout mice for TAS1R2 and/or TAS1R3 subtypes that showed a blunted response to sugars, sweeteners, and D-amino acids, confirming the TAS1R2/TAS1R3 heterodimer as the main sweet taste receptor *in vivo*.

TABLE 3 | Agonists of sweet taste receptors along with their EC50 values.

Agonists	Nature	Binding pocket	EC50 (mM)	References
Sucrose	Natural carbohydrate	VFT (TAS1R2 and TAS1R3)	62	(Li et al., 2002; Servant et al., 2010; Zhang et al., 2010; Zhang
				et al., 2003)
Aspartame	Peptide	VFT (TAS1R2)	0.75	(Li et al., 2002; Liu et al., 2011; Masuda et al., 2012)
Neotame	Peptide	VFT (TAS1R2)	5	(Li et al., 2002; Masuda et al., 2012)
Cyclamate	Sulfamate	TMD (TAS1R3)	3.1	(Xu et al., 2004; Jiang et al., 2005c)
Brazzein	Protein	CRD (TAS1R3)	0.08	(Li et al., 2002; Jiang et al., 2004; Ide, et al., 2009; Masuda et al.,
				2012)
Thaumatin	Protein	CRD (TAS1R3)	0.005	Masuda et al., 2012; Jiang et al., 2004
Monellin	Protein	VFT (TAS1R3), VFT (TAS1R2)	0.01	Koizumi et al., 2007; Jiang et al., 2004
Neoculin	Protein	VFT (TAS1R2)	0.001	(Jiang et al., 2004; Koizumi et al., 2007)
Saccharin	N sulfonyl amide	VFT (TAS1R2)	0.19	(Li et al., 2002; Masuda et al., 2012; DuBois, 2016)
Suosan, cyanosuasan	Arylurea	VFT (TAS1R2)	ND	(Tinti and Nofre, 1991; Du Bois, 2016)
SC-45647	Guanidinoacetic acid	VFT (TAS1R2)	0.3	(DuBois, 1995; Sanematsu et al., 2014)
Sucralose	Halogenated carbohydrate	VFT (TAS1R2 and TAS1R3)	0.06	(Li et al., 2002; Masuda et al., 2012)
Acesulfame K	Sulfamate ester	VFT (TAS1R2)	0.54	(Li et al., 2002; Masuda et al., 2012)
Perillartine	Oxime, ethoxyphenyl urea, alkoxyaryl urea,	TMD (TAS1R2)	15	(Li et al., 2002; Servant et al., 2010)
Dulcin	Ethoxyphenyl urea	TMD (TAS1R2)	0.01	(Servant et al., 2010)
S819	Alkoxyaryl urea	TMD (TAS1R2)	0.025	(Zhang et al., 2008)
D-tryptophan	Amino acid	VFT (TAS1R2)	2.09	(Li et al., 2002; Masuda et al., 2012)
Xylitol, sorbitol	Polyols	VFT (TAS1R2)	ND	(Mahalapbutr et al., 2019)
Maltotriose, acarbose	Oligosaccharide,	ND	ND	(Pullicin et al., 2017;

Agonists	Nature	Binding pocket	EC50 (mM)	References
	pseudotetrasaccharide			Pullicin et al., 2019)

Where VFT, Venus flytrap domain; TMD, transmembrane domain; ND, not determined.

Table **13.8.3**: Agonists of sweet taste receptors and their EC50 values.

The sweet receptor couples to heterotrimeric Gα-gustducin, which includes Gβ3 and Gγ13, as mice lacking Gα-gustducin showed a reduced response to sweet substances, either natural or artificial. Moreover, a point mutation in the C-terminal region of gustducin (G352P) (critical for its receptor interaction) results in the loss of its ability to activate taste GPCRs while keeping other functions intact. Further, G352P acts as a dominant negative to block heterotrimeric G protein interaction with taste receptors. It disrupts the responses to sweet and bitter compounds in wild-type (WT) and null mice. In addition, the G352 mutant further reduces any residual sweet/bitter taste responses of the null mice by acting as a "βγ sink" to bind all unbound βγ-subunits and remove them from the viable pool of G protein heterotrimers available to the receptor. These observations confirm the essential requirement of Gα-gustducin in sweet and bitter taste transduction.

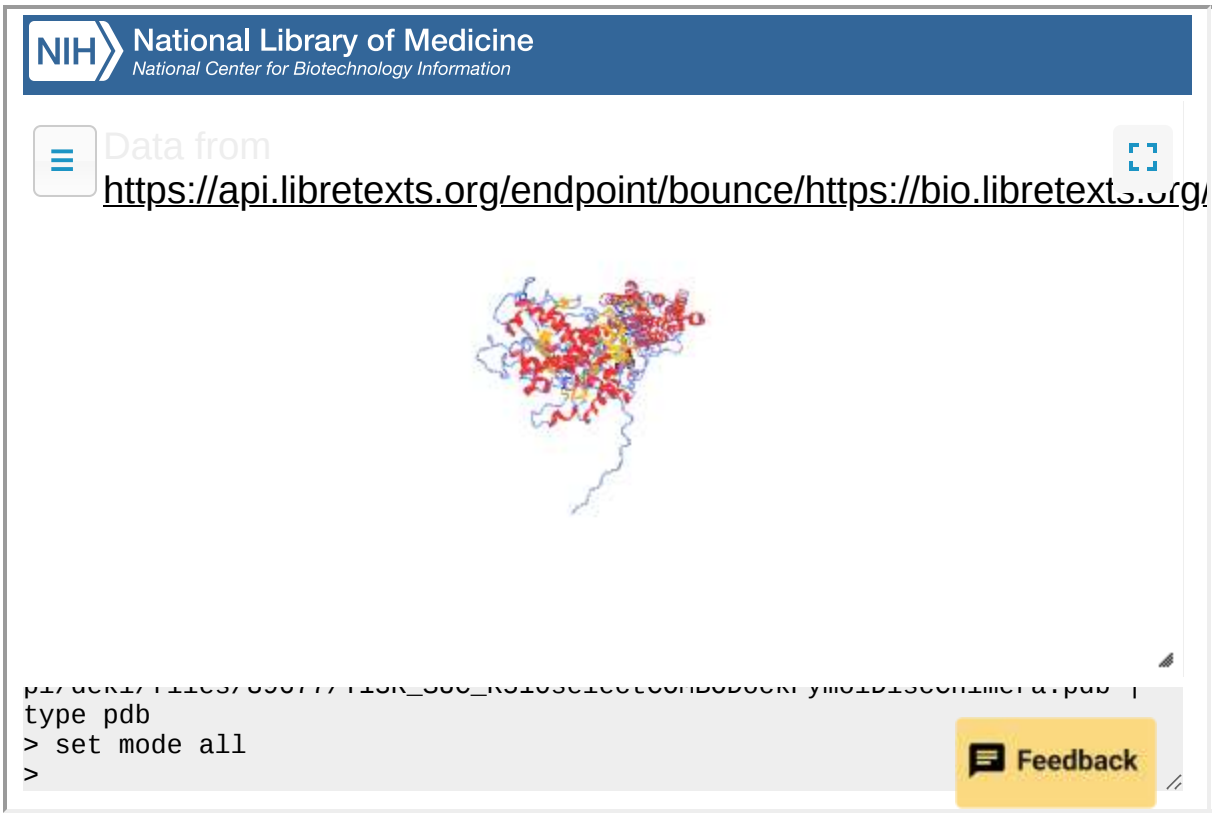
In addition to the Gα-gustducin pathway, sweet taste transduction occurs *via* two additional signaling pathways involving different secondary messengers. The first one involves cAMP, and the second one involves IP3. Normally, sugars elevate the cAMP level, while sweeteners stimulate IP3 production. Sucrose or other sugars bind to either TAS1R2 or TAS1R3 and recruit Gαs protein, leading to increased cAMP levels, which initiate the influx of cations through ion channels. Alternatively, cAMP activates protein kinase A, leading to TRC cell depolarization and an influx of calcium ions and neurotransmitter release. Sweetener binding to the TAS1R2/TAS1R3 heterodimer recruits Gα-gustducin proteins that stimulate PLCβ2, which in turn hydrolyzes phosphatidylinositol 4,5-bisphosphate (PIP2) to diacylglycerol (DAG) and IP3. IP3R3-induced Ca^{2+} release from ER stores activates TRPM5, leading to an action potential and subsequent release of neurotransmitters. See Figure **13.8.3** for details.

Interestingly, confirmed that TRPM4 also mediates taste signaling independent of TRPM5, and knocking out both channel proteins (TRPM4/5) abolishes the sweet, umami, and bitter taste response completely. This revealed another layer of complexity to sweet signal transmission. This in-depth mechanistic research has increased our understanding of sweet and bitter receptors. It also presents a challenge to dissect the taste signal transmission pathways for umami and kokumi.

Recent Updates: September 16, 2024

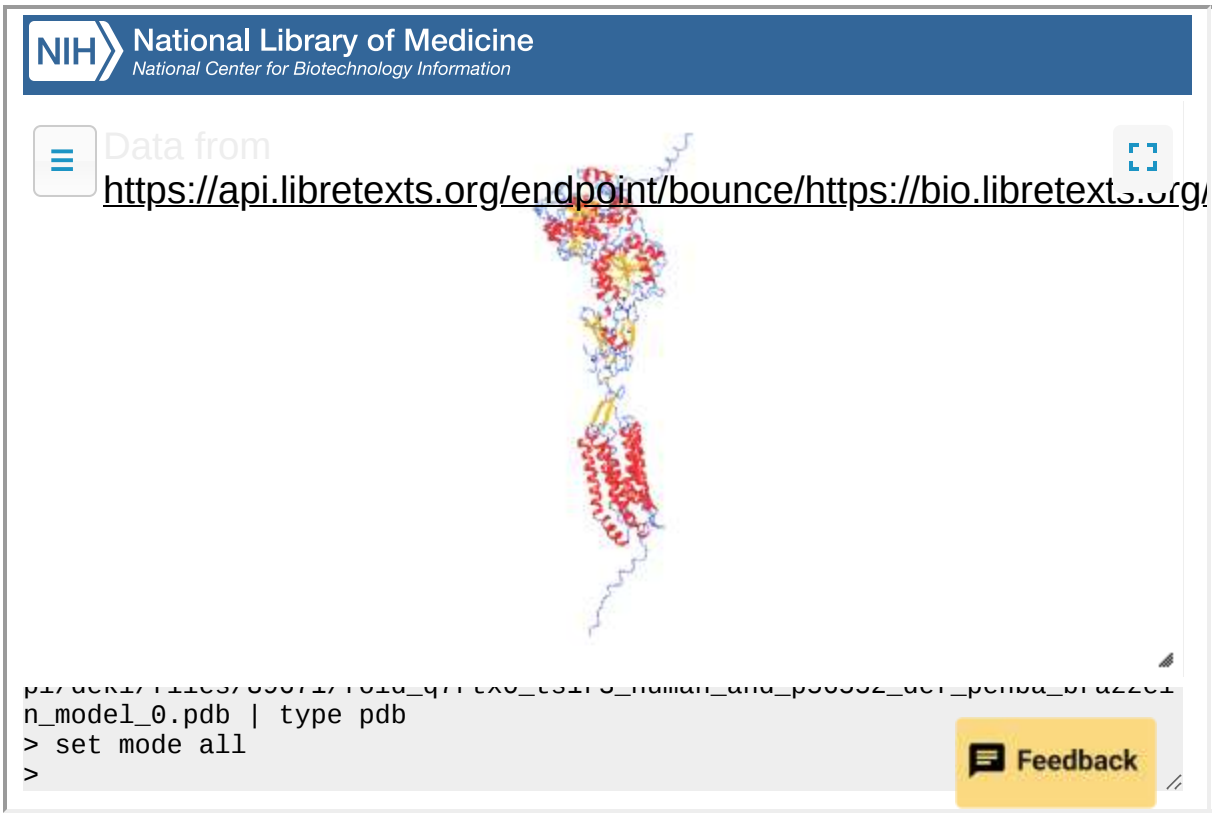
Sweet Proteins

Figure **13.8.*x*** shows an interactive iCn3D model of Taste receptor type 1 member 3 (TAS1R3, also abbreviated T1R3) AlphaFold model (UniProt **Q7RTX0**) docked with sucrose. The PDB file was obtained using SwissDock and AutoDock Vina within SwissDock for docking.



Sucrose is shown in spacefill bound about halfway down the T1R3 receptor.

Now let's compare this docked structure to Figure **13.8.x** below, which shows an interactive iCn3D model of Taste receptor type 1 member 3 (TAS1R3, also abbreviated T1R3) AlphaFold model (uniprot **Q7RTX0**) docked with protein called **brazzein** (recommended name **Defensin-like protein)** from Pentadiplandra brazzeana. Abramson, J et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature (2024).



Brazzein (54 amino acids, molecular weight 6500) is bound in the same region as the spacefill model of sucrose in the first iCn3D model. The protein is digested in the gut, so it theoretically has no negative metabolic consequences compared to sucrose and non-protein artificial sweeteners. This protein is 2000x sweeter than sugar and is used in some commercial products. The PDB file was obtained using AlphaFold Server (beta) to dock the two proteins after inputting their amino acid sequences (an extraordinary feat possible since 2023 but now easily available online for free!).

Figure 13.8.6 below shows an interactive iCn3D model showing the interaction interface between TAS1R3 (T1R3) and brazzein.

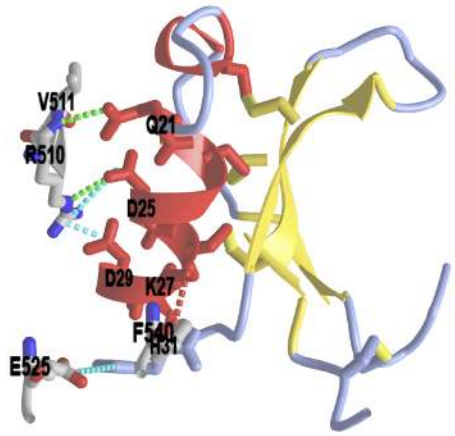


Figure 13.8.6:

For an interactive iCn3D model of this rendering:

- download this file to your computer's download folder (easiest in Chrome)
- open iCn3D

- In the top menu bar, choose **File, Open File, iCn3D PNG Image,** and navigate to your download folder to select and open the file.

Note the diverse types of noncovalent interactions between amino acids from the receptor and brazzein (in CPK colors).

Compare this with the interactions between docked sucrose and TAS1R3 (T1R3) shown in the figure below. Docking of the TAS1R3 (T1R3) AlphaFold structure and sucrose was performed using AutoDock Vina through Swiss Dock. Figure 13.8.7

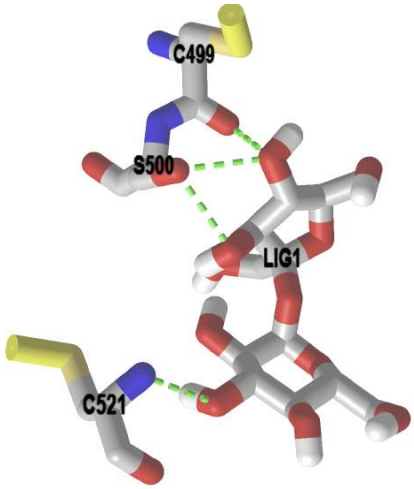


Figure 13.8.7: For an interactive iCn3D model, download this file to your computer's download folder (easiest in Chrome) and follow the instructions above.

Note that the bound ligand (sucrose) interactions occur in the same region of T1R3 as for bound brazzein, but involve just hydrogen bonding.

13.8.5.1: Structural, Molecular, and Conformational Changes of Sweet Receptor

Since the sweet taste receptor has not yet been crystallized, determining the sweetener binding site's structure and activation mechanism has been challenging. Based on homology with other class C GPCRs (mGluRs and $GABA_B$Rs), multiple studies propose similar activation mechanisms for the sweet receptor. The many sweet agonists and their diverse binding sites across receptor domains (VFT, TMD, and CRD) (see Table 13.8.3) may explain its complex yet broadly tuned nature. For example, a single residue in VFT (I60) of TAS1R3 of the TAS1R2/TAS1R3 heteromer is required for a saccharin preference in inbred mouse strains.

Several studies utilizing homology and computational modeling based on the crystal structure of mGluR and $GABA_B$Rs have predicted structural and functional aspects of orthosteric and allosteric binding sites for the sweet receptor. They reported that both VFT regions undergo ligand-dependent conformational changes and intersubunit interactions between ECDs, further stabilizing heterodimer formation for subsequent downstream signaling. The binding of orthosteric agonists to VFT of TAS1R2 leads to major conformational changes that form a TMD6/TMD6 interface between TMDs of TAS1R2 and TAS1R3, which is consistent with the activation process observed biophysically on the mGluR2 homodimer. The initial role of the bound agonist is to pull the bottom part of VFT3 (VFT of TAS1R3) toward the bottom part of VFT2 (VFT of TAS1R2) to transmit this movement from VFT2 (where agonists bind) through the VFT3 and the CRD3 (VFT and CRD of TAS1R3) to the TMD3 (TMD of TAS1R3). This facilitates G protein coupling and downstream signaling. The CRDs are crucial in this streamlined relay of structural changes, where disulfide bonds provide rigidity to the CRD and amplify the mechanical constraints that help attain an active conformation. This is empirically supported by a study in which a single mutation (A537P) in the CRD of TAS1R3 abolished the response to all sweeteners, indicating that the CRD3 must couple ligand binding in VFT2 to the conformational changes required in TMD3 for receptor activation.

Trafficking and cell surface expression are also crucial factors for sweet taste transduction. Molecular modeling with mutagenesis scanning revealed specific regions consisting of hydrophobic residues in ECD (site II; at the tip of CRD) and TMD regions (site IV; includes TMD6 and the cytoplasmic base of TMD5) of the TAS1R2 subunit to be important for dimerization with TAS1R3. Moreover, the CRD region and ECL2 domain of the transmembrane region seem important for surface co-expression of the

TAS1R2/TAS1R3 dimer. In particular, the cytosolic C-terminus portion of the CRD region of TAS1R2 needs to be properly folded for coexpression and trafficking. This reflects the difficulty in expressing these receptors consistently in mammalian cell lines.

13.8.5.2: Positive Allosteric Modulation of Sweet Receptor

Class C GPCRs pose an ideal target for allosteric modulation, either positive (PAM) or negative (NAM). PAMs show little or no agonist activity but significantly enhance agonist activity. Sweet taste is a major global target of the food industry, and non-caloric sweeteners are highly sought after to exploit a huge commercial market. Novel PAMs (see Table **13.8.4**) for the sweet heteromer were reported that were not sweet on their own but significantly enhanced the sweetness of sucralose or sucrose. Agonist binding to the VFT region of TAS1R2 facilitates a closed conformation, constituting an active state of the sweet receptor, while its open conformation represents an inactive state. Molecular modeling and mutagenesis studies revealed that these PAMs follow a similar binding mode to that reported for umami PAMs (IMP and GMP). They bind near the opening of the binding pocket of the VFT region adjacent to their agonists, through Van der Waals and hydrogen bonding interactions, and utilize several critical residues for their activity. Although these residues are not in direct contact with any receptor-bound sweetener, mutation of some of them (K65, Y103, L279, D307, and R383) diminishes the response to sweeteners, suggesting that these residues normally stabilize the closed conformation. The initial closing of the VFT region by agonist binding and further stabilization of the closed conformation by subsequent binding of SE modulators occur in two steps. First, by interacting with the ECD region of TAS1R2, and second, by strengthening the hydrophobic interactions between the two lobes of ECD and lowering the free energy needed for their closure.

TABLE 4 | Sweet taste receptors' positive allosteric regulators with concentration (used in cell-based assays in studies) and negative allosteric modulators with their IC50 values.

Positive allosteric modulators (PAMs)	Nature	Binding pocket	Conc. (mM)	References
SE1, SE2, SE3	Undisclosed	VFT (TAS1R2)	0.05	(Servant et al., 2010; Zhang et al., 2010)
Neohesperidin dihydrochalcone (NHDC)	Flavonoid	TMD (TAS1R3)	0.25	(Jiang et al., 2005c; Winnig et al., 2007)
Unnatural tripeptides (several)	Biaryl derivative tripeptides	ND	2 – 20	Yamada et al., 2019
Sodium, cholesterol	Cation, lipid	TMD (TAS1R2)	ND	Perez-Aguilar et al., 2019
NAMs	—	—	IC50 (mM)	—
Lactisole	Carboxylic acid salt	TMD (TAS1R3)	0.041	(Jiang et al., 2005c)
(2-(2,4-dichlorophenoxy)propionic acid)	Carboxylic acid salt	TMD (TAS1R3)	0.006	(Nakagita et al., 2019)
Gymnemic acid	Triterpenoid glycoside	TMD (TAS1R3)	6.9	(Sanematsu et al., 2014)
Clofibric acid	Herbicide	TMD (TAS1R3)	1.4	(Maillet et al., 2009; Kochem and Breslin, 2017)
Amiloride	Diuretic	TMD (TAS1R2)	0.87	(Imada et al., 2010; Zhao et al., 2018)
Umami compounds: MSG, Glu-Glu, Glu-Asp	Peptides	VFT (TAS1R2)	ND	(Shim et al., 2015)

Where VFT, Venus flytrap domain; TMD, transmembrane domain; ND, not determined.

Table **13.8.4**: Sweet taste receptor's positive allosteric regulators with concentration (used in cell-based assays in studies) and negative allosteric modulators with their IC50 values.

Several unnatural tripeptides with a novel core biaryl structure were found as potential sweet enhancers using a high-throughput chemical screening approach and heterologous expression of the TAS1R2/TAS1R3 heteromer. This study divided the potential molecule into three parts: "head and linker," which are essential for its sweet enhancer activity, while the "tail" determines the activity level. This approach provided some useful inputs toward the synthesis of potent PAMs. Firstly, an amine incorporated at the α-position of the carbonyl group in the tail structure interacts with the TAS1R2 subunit, increasing allosteric activity. Secondly, additional hydrophobic substitutions in the tail structure increased allosteric activity in the molecule. Lastly, the distance between the head and linker and the insertion of an amide bond is crucial for its synthesis. Although their binding characteristics and allosteric mechanisms are unknown, these observations provide a starting point for identifying and synthesizing new sweet PAMs in the future.

Unlike an agonist, which binds to the extracellular domain, small-molecule PAMs can also bind to the transmembrane domain in class C GPCRs. For example, the flavonoid sweetener, neohesperidin dihydrochalcone (NHDC) binds to TMD regions to enhance the agonist-induced sweet response. It interacts with a receptor binding pocket in the TMD of TAS1R3 and requires seventeen critical residues in TMDs and extracellular loop 2 for its allosteric activity. These residues also contribute to cyclamate and lactisole binding sites. Among seventeen residues, eight alter receptor activation by NHDC (Q637$^{3.29}$, S640$^{3.32}$, H641$^{3.33}$, Y699$^{4.60}$, W775$^{6.48}$, F778$^{6.51}$, L782$^{6.55}$, and C801$^{7.39}$) and influence lactisole-mediated inhibition. Similarly, nine of the seventeen residues (Q637$^{3.29}$, H641$^{3.33}$, H721^{ex2}, S726$^{5.39}$, F730$^{5.43}$, W775$^{6.48}$, F778$^{6.51}$, L782$^{6.55}$, and C801$^{7.39}$) mediate activation by cyclamate. In contrast, six (Q637$^{3.29}$, H641$^{3.33}$, W775$^{6.48}$, F778$^{6.51}$, L782$^{6.55}$, and C801$^{7.39}$) influence receptor inhibition by lactisole as well as receptor activation by cyclamate [superscript refers to the nomenclature suggested for class C GPCRs, where the first number denotes TMD region. The second number denotes the residue position from the most conserved residue.

Notably, three critical residues in TMD6 (W775$^{6.48}$, F778$^{6.51}$, L782$^{6.55}$) and one in TMD7 (C801$^{7.39}$) of TAS1R3 were found crucial for allosteric binding, as their mutation to alanine altered the receptor's sensitivity to NHDC and cyclamate, as well as to the inhibitor lactisole. Therefore, TMD6 and TMD7α helices of TAS1R3 are integral to allosteric modulation of the sweet receptor, implicating them in TAS1R2 and TAS1R3 subunit interactions and indicating an important role for this structural region in the conformational changes involved in receptor activation. Furthermore, these residues are conserved across mammalian species.

13.8.5.3: Negative Allosteric Modulation of Sweet Receptor

Like PAMs, negative allosteric modulators (NAM) such as lactisole and gymnemic acid bind to the TMD region of TAS1R3 and inhibit sweet substance-induced responses. Lactisole, an aralkyl carboxylic acid, inhibits the sweet and the umami receptor response in humans and presents a rare opportunity to study the structural cross-talk between these two taste qualities. Using heterologous expression and mutagenesis, Jiang et al. reported that lactisole's sweet inhibition might be mediated by its binding to TMD3, TMD5, and TMD6 of TAS1R3 and induce a conformation change that restricts the movement required to stabilize the active state. Residues A733$^{5.46}$ in TMD5, L798$^{7.36}$ in TMD7, and R790^{ex3} in extracellular loop 3 were crucially important for sensitivity to lactisole in humans. These observations were confirmed in a recent study where 2-(2,4-dichlorophenoxy)propionic acid (2,4-DP) was found to be a more potent antagonist and utilize the same residues as well as four additional ones (H641$^{3.37}$, H734$^{5.43}$, F778$^{6.53}$, and Q794$^{7.32}$) in binding to TAS1R3. Moreover, the (S)-isomer of both compounds was found to be more strongly bound to the TMD of TAS1R3 and be a more effective inhibitor [lactisole; (S)-lactisole IC50, 20 µM, while (R)-lactisole exerted no inhibition at this concentration.; 2,4-DP: (S)-isomer was 10-fold more effective than (R)-2,4-DP. The (S)- lactisole isomer interacts with the TMD via its carboxyl group and stabilizes in only one orientation in the binding pocket, which does not allow for very strong binding. In contrast, (S)-2,4-DP binds through two moieties simultaneously, a carboxyl group and an aromatic ring with two Cl^- groups, and stabilizes in several different orientations through hydrophobic interactions that allow stronger binding, resulting in stronger negative allosteric modulation.

These observations provide information about the relevance of structural modification in NAM compounds that could affect their interaction with the receptor. Although TMDs of TAS1R3 are the most likely regions responsible for allosteric modulation, TMDs and VFT regions of TAS1R2 cannot be ruled out completely. For example, the diuretic amiloride binds to TAS1R2 (TMD3, TMD5, TMD7) and inhibits the sweet response in a species-dependent manner. Further, the umami compound [monosodium glutamate (MSG)] and peptides (Glu-Asp, Glu-Glu) bind to the VFT region of TAS1R2 and inhibit the sweet-induced response. These

observations suggest that both subunits are important for the allosteric activity of TAS1R2/TAS1R3, and further structural studies are required to design novel sweet allosteric modulators.

13.8.6: Umami Taste Signal Transduction Mechanisms

In contrast to four well-known basic human tastes (sweet, bitter, salty, and sour), umami or 'savory taste' is relatively recent and was introduced in the early 2000s by Kikuna Ikeda as a new seasoning element in food. The main stimulus for the umami taste is the amino acid L-glutamate, present in the diet mainly as MSG. Glutamate was first extracted from konbu/kombu (dried kelp of *Fucus vesiculosus*) and described as having a "unique taste" and "very different from other tastes". The terminology "umami" comes from the Japanese word "umai" meaning "delicious." Moreover, the taste of umami is also produced by food such as mushrooms and soy sauce that contain amino acids (L-aspartate), peptides, and synthetic ingredients similar to glutamate and some organic acids, as shown in Table **13.8.5**.

TABLE 5 | Umami receptor agonists with their EC50 values and other pharmacological properties.

Agonist	Nature	EC50 (mM)	Binding pocket	References
L-amino acids (glutamate, aspartate, alanine, serine, asparagine, arginine, histidine, threonine, glutamine) L-theanine	Amino acids Amino acid (plant origin)	3 (glutamate), ND for others ND	VFT (TAS1R1) VFT (TAS1R1)	(Li et al., 2002; Nelson et al., 2002; Zhang et al., 2008; Toda et al., 2013) (Narukawa et al., 2014)
VFT, venus flytrap domain; ND, not determined.				

Table **13.8.5**: Umami receptor agonists with their EC50 values and other pharmacological properties.

The umami receptor (TAS1R1/TAS1R3) is a heteromeric member of the class C GPCRs, whereas most other receptors of this class exist as homodimers. TAS1R1/TAS1R3 is the predominant umami taste receptor, and the TAS1R1 subtype is critical for sensing umami taste, as its deletion abolishes the response to umami taste stimuli. However, TAS1R1/TAS1R3 is not the only receptor capable of detecting umami ligands. Studies using heterologous expression, afferent nerve recordings, and behavioral experiments have confirmed that metabotropic glutamate receptors 1 and 4 (taste-mGluR1 and taste-mGluR4) also sense umami stimuli. Notably, TAS1R3 knock-out mice show a strongly diminished response to glutamate and sweet stimuli, and taste cells isolated from these mice respond to IMP and glutamate, which is abolished in the presence of mGluR antagonists (Pal Chaudhry et al., 2016). TAS1R1/TAS1R3 is not only activated by glutamate, but this activation is strongly enhanced in the presence of 5′-ribonucleotides (inosine 5′ monophosphate; IMP), a response that is a hallmark of umami taste.

The main transduction components following the activation of TAS1R1/TAS1R3 are similar to those for sweet taste, i.e., α-gustducin (and γ13/β1 or β3), PLCβ2, IP3R, and TRPM4/5. Cyclic nucleotides may also contribute to the transduction of umami taste in TRCs. When taste tissue is stimulated with umami, its cyclic AMP level is decreased. However, the consequence of decreased cAMP in TRCs has not yet been fully elucidated. Both α-transducin and α-gustducin are involved in umami taste signal transduction, as mice lacking the gene for one of these proteins showed a reduced response to this taste. In the taste palate, fungiform papillae, *α*-gustducin, and *α*-transducin activate PDE that reduces cAMP levels. Ligand binding to the TAS1R1/TAS1R3 heterodimer releases Gβγ subunits to stimulate PLCβ2, which hydrolyzes PIP2 to DAG and IP3. IP3 then activates IP3R3, releasing calcium ions from intracellular compartments, as shown in Figure **13.8.3**. Calcium ions activate TRPM5 and TRPM4 channels, leading to an influx of sodium ions, subsequent cell membrane depolarization, and finally the release of ATP, which activates ionotropic purinergic receptors located in sensory fibers. This pathway was confirmed when mice devoid of TRPM5, TRPM4, PLCβ2, and IP3R3 showed a reduced response to umami taste perception following glutamate stimuli.

13.8.6.1: Structural, Molecular, and Conformational Changes of Umami Receptor

In the last decade, several in-depth modeling and mutagenesis approaches have improved structural and molecular understanding of the umami receptor. The VFT regions of both subunits of TAS1R1/TAS1R3 comprise orthosteric and allosteric ligand binding sites for umami stimuli.

Mutagenesis and molecular modeling studies reveal that the cognate agonist glutamate binds in the VFT region of the TAS1R1 subunit of TAS1R1/TAS1R3 and stabilizes the closed active receptor conformation. Moreover, four residues in the TAS1R1 VFT region (S172, D192, Y220, and E301) showed no detectable response to glutamate when they were mutated to alanine, suggesting that they are critical for glutamate binding. The glutamate binding and stabilization of the closed conformation of TAS1R1 activate the downstream signaling pathway, while TAS1R3 remains in an open (inactive) conformation. Therefore, closure of the VFT is the key event that sensitizes umami taste receptor signal transduction. Apart from glutamate, other L amino acids were also found to elicit functional responses by binding to the corresponding VFT region of TAS1R1. Six residues that contributed to the acidic amino acid agonist (L-glutamate and L-alanine) responses have been identified (S148, R151, A170, E174, A302, and D435).

13.8.6.2: Allosteric Modulation of Umami Receptor

Due to significant advancements in understanding and the application of umami taste in the food industry, its allosteric modulators are being sought after. Several allosteric umami ligands have been discovered with varying potency, only a few of which have been characterized at the molecular level. The best-characterized umami PAMs, the 5′-ribonucleotides: inosine 5′-monophosphate (IMP) and guanosine 5′-monophosphate (GMP), interact with the VFT region of the TAS1R1 subunit to enhance the glutamate-induced response that is the hallmark of umami taste, as shown in Table 13.8.6. IMP and GMP binding sites in the VFT are adjacent to those for glutamate binding. The mutation of four residues (H71, R277, S306, and H308) abolished the IMP/GMP-induced glutamate response, suggesting their involvement in the allosteric binding of these nucleotides. Structurally, IMP and GMP stabilize the closed form of the TAS1R1 VFT region through electrostatic interactions and coordinate the positively charged residues that act as pincers. The ability of IMP and GMP to interact with the VFT region (as opposed to the TMD region) represents a unique mechanism of positive allosteric regulation within class C GPCRs.

TABLE 6 | Umami receptor allosteric modulators with concentrations used in cell-based assays and other pharmacological properties.

Allosteric modulators	Nature	Conc. (mM)	Binding pocket	References
IMP/GMP	Nucleotide	1	VFT (TAS1R1)	(Li et al., 2002; Nelson et al., 2002; Zhang et al.,
				2008)
Cyclamate	Sodium	8	TMD (TAS1R3)	(Xu et al., 2004)
	cyclohexylsulfamate			
Methional (3-methylsulfanylpropanal)	—	0.12	TMD (TAS1R3)	(Toda et al., 2018)
Lactisole (2-4-methoxyphenoxy propionic acid)	Carboxylic acid salt	5	TMD (TAS1R3)	(Xu et al., 2004)
Clofibric acid (4-chlorophenoxy)-2-methylpropanoic	Herbicide acid	4	TMD (TAS1R3)	(Maillet et al., 2009; Kochem and Breslin, 2017)
acid				

Where VFT, Venus flytrap domain; TMD, transmembrane domain.

Table 13.8.6: Umami receptor allosteric modulators with concentrations used in cell-based assays and other pharmacological properties.

In contrast to IMP and GMP, which bind to the TAS1R1 extracellular domain, the well-known flavor compound methional and its analogs bind to the TMD region and allosterically regulate the umami receptor in a species-dependent manner. Importantly, methional utilizes several distinct residues in different TAS1R1 transmembrane domains (TMD2-7) to act as a PAM in the human umami receptor, yet it behaves as an NAM in the mouse counterpart. This unusual phenomenon provided an opportunity to study the mechanisms of both positive and negative modulation in TAS1R1 simultaneously.

Construction of chimeric receptors between human (h) and mouse (m) and their functional analysis demonstrated that the TMD of TAS1R1 is the key domain for switching methional's PAM/NAM activities. Point mutation substitutions between these species identified four residues (h/m; F768/L769, N769/H770, S799/T800, and S802/G803) that are collectively required to switch PAM/NAM activities. A similar mode of allosteric regulation and PAM/NAM mode switching has been reported for mGluR5, suggesting this is an unusual and distinct phenomenon of the class C GPCRs. Further, alanine scanning mutagenesis in TAS1R1 of the corresponding residues vital for the activity of other taste inhibitors (sweetener inhibitors; NHDC and cyclamate; sweet and umami taste inhibitors; lactisole) revealed three residues required for PAM (W697$^{4.50}$, F728$^{5.40}$, and F732$^{5.44}$) and a single residue (F642$^{3.40}$) for NAM. These results suggest that methional's PAM and NAM activities are conferred by residues distinct from those required for the PAM/NAM switch. Knowing that methional is an important part of food seasoning globally, these observations could help maximize its use in enhancing flavors, amino acids, and nucleotides.

Despite PAMs being a central focus for umami allosteric modulation, there has also been considerable research on negative allosteric modulation, where lactisole emerged as a prominent NAM of the umami receptor, TAS1R2/TAS1R3. Because umami and sweet receptors share the TAS1R3 subunit, findings from studies on sweet receptor lactisole binding are relevant. A comprehensive study on the sweet receptor identified critical residues within the TMD regions (S640$^{3.32}$, H641$^{3.33}$ in TMD3 and F778$^{6.51}$, L782$^{6.55}$ in TMD6) of TAS1R3 required for the lactisole binding pocket and showed a large effect on sensitivity to lactisole. Because lactisole shares structural similarities with two other classes of compounds: fibrates and phenoxy-herbicides, researchers studied them to search for novel sweet/umami inhibitors. The lipid-lowering drug, clofibric acid, inhibits the TAS1R3 umami receptor-mediated *in vitro* and *in vivo* response, as shown in Table **13.8.6**. Like lactisole, clofibrate inhibits the umami taste from glutamate by binding with a similar affinity to TAS1R1/TAS1R3. However, its specificity against the umami receptor must still be validated alongside other umami taste receptors (mGluR1, mGluR4, or NMDA).

13.8.7: TYPE 2 TASTE G PROTEIN-COUPLED RECEPTORS (BITTER RECEPTORS)

Type 2 taste GPCRs are represented by bitter taste receptors with a distinct subset of bitter-sensing cells in type II TRCs. Notably, 25 bitter taste receptors (TAS2Rs) are reported to be expressed in humans. Much work has been done to explore the diversity among TAS2Rs and their agonists in taste biology. Some TAS2Rs (TAS2R3, TAS2R5, TAS2R13, TAS2R50) are narrowly tuned to structurally similar bitter compounds, whereas others are broadly tuned (TAS2R10, TAS2R14, TAS2R46), responding to several bitter compounds. Initially, it was believed that each bitter-sensitive type II TRC expressed every TAS2R isoform, but other studies suggest that TAS2Rs can be expressed differentially, allowing for possible discrimination among bitter compounds. Please refer to Figure **13.8.4*B*** for the basic structure of the bitter receptor.

13.8.7.1: Bitter Taste Signal Transduction Mechanisms

The bitter taste is the most complex of all the five basic tastes and protects against the ingestion of toxic substances by eliciting an innate aversive response across species. The TAS2Rs that mediate bitter taste perception are among ∼50 TAS2Rs identified in mammals, and 25 are known to be expressed in humans. The TAS2R family is the most diverse and binds to a wide range of agonists compared to other taste GPCRs.

TAS2Rs are distinctive among class A GPCRs in that many bind agonists with low apparent affinity in the micromolar range, rather than the nanomolar range. The activation of TAS2Rs by harmless, minute amounts of bitter compounds, such as those contained in most vegetables, would limit the availability of food resources appearing safe for consumption and therefore could negatively affect survival. Hence, the concentration ranges at which bitter taste receptors are activated are well-balanced to allow species to maintain a healthy diet yet avoid ingesting spoiled food containing strongly bitter ligands.

Hundreds of bitter compounds have been reported to evoke bitterness and activate human bitter receptors in different cell-based assays. These bitter agonists include plant-derived and synthetic compounds such as peptides, alkaloids, and many other substances. Many compounds activate some TAS2Rs, whereas others show strict specificity for a single bitter compound.

Interestingly, TAS2R31, TAS2R43, and TAS2R46 have around 85% sequence homology, but they bind to different agonists, reinforcing the idea that each TAS2R might have a unique ligand-binding pocket.

The canonical TAS2R signal transduction cascade signaling molecules shared among bittersweet and umami receptors include the heterotrimeric G protein subunits (Gα-gustducin, Gβ3, and Gγ13), a phospholipase C (PLCβ2), an inositol trisphosphate receptor (InsP3R), and the TRPM5 ion channel. Upon receptor activation by bitter ligands, the G protein α-gustducin dissociates from its βγ subunits. The latter activates PLCβ2, releasing Ca^{2+} from IP3-sensitive Ca^{2+} stores, resulting in Na^{+} influx through TRPM5 channels. This Na^{+} influx depolarizes the cells and causes the release of neurotransmitter ATP through gap junction hemichannels or CALHM1 ion channels, as shown in Figure **13.8.3**.

13.8.7.2: Structural, Molecular, and Conformational Changes of Bitter Receptors

Classification of TAS2Rs has always been ambiguous because they were initially considered a distinct family or grouped with the frizzled receptors. Still, most recent analyses support their classification with Class A GPCRs. The ability of bitter taste receptors to interact with numerous structurally diverse substances compared to other GPCRs is remarkable. It includes many drugs/antibiotics, polyphenols, bacterial metabolites, salts, and metal ions. Therefore, exploring the criteria for identifying highly heterogeneous bitter compounds with pronounced selectivity has become a significant research area. Some of these studies rely solely on in silico homology/computational modeling, and others on *in vitro* genetic modification and functional assay systems.

As a group of over ∼50 receptor subtypes, TAS2Rs recognize structurally diverse agonists. Some are broadly tuned (TAS2R46, TAS2R14, TAS2R10, and TAS2R43) and recognize diverse agonists, while others (TAS2R1, TAS2R4, TAS2R7) show strong selectivity and narrow tuning. The agonist binding cavity in most bitter GPCRs is located deep within their transmembrane domain (TMD), except TAS2R7, which resides on the extracellular surface. TAS2Rs are also distinct in containing highly conserved TMD regions, with thirteen key residues and two motifs (LXXXR in TMD2 and LXXSL in TMD5) that are absent in class A GPCRs, and may reflect their different activation mechanisms. LXXSL plays a structural role by stabilizing the helical conformation of TMD5 at the cytoplasmic end and a functional role by interacting with residues in intracellular loop 3 (ICL3), which is important for proper receptor folding and function. Moreover, mutation of the conserved residues in LXXSL and LXXXR motifs results in protein misfolding and poor surface expression.

The initial study highlighting bitter taste receptors' structure–activity relationship was performed with receptors belonging to a subfamily of closely related TAS2Rs. By physically swapping the extracellular loop 1 (ECL1) between TAS2R43 and TAS2R31, chimeric TAS2R31/TAS2R43 (ECL) gained responsiveness to the compound n-isopropyl-2methyl-5-nitrobenzenesulfonamide (IMNB), whereas the reverse chimera TAS2R31 (ECL)/TAS2R43 lost responsiveness to IMNB. Although this report supports an important contribution of residues located within the transmembrane region of the investigated receptors, the extracellular loops appear important for agonist selectivity. This empirical finding contrasts with earlier computational studies, which predicted the agonist binding site to lie within the helical bundle of TAS2Rs without particular contacts between extracellular loops and docked agonists.

13.8.7.3: Bitter Receptor Ligand Binding Pocket

The emergence of TAS2Rs as the most broadly tuned taste receptors might give the impression that their specific interaction with numerous agonists is due to several binding pockets that accommodate subgroups of bitter compounds. However, structure–function analysis of TAS2Rs (except for TAS2R7) has demonstrated the presence of only a single agonist binding pocket comprising the upper parts of TMD2, TMD3, TMD5, TMD6, and TMD7. Their broad tuning and recognition of such a broad spectrum of agonists might most likely be attributed to an additional extracellular binding site called a “vestibular site,” in addition to the orthosteric site, as reported for TAS2R46. This two-site architecture offers more ligand recognition points than a single one and thus might help in selecting the appropriate agonists. Moreover, the presence of the vestibular site may also help to discriminate among the wide spectrum of bitter ligands.

Although broadly tuned receptors (TAS2R46, TAS2R31, and TAS2R43) have high homology in amino acid sequence, their agonist profiles only slightly overlap, which suggests the involvement of key residues at different positions in agonist specificity. Consequently, when strychnine-interacting positions in TAS2R46 (residues differ at this position in TAS2R31, TAS2R43) were exchanged between these two receptors, not only was the strychnine responsiveness transferred to the recipient receptor (TAS2R31, TAS2R43), but also sensitivity to additional TAS2R46 agonists (absinthin and denatonium). Sensitivity to activation by aristolochic

acid was lost in the mutant receptors. This experimental evidence supports the presence of a common agonist binding pocket and agrees with other studies on TAS2R16, TAS2R14, and TAS2R7 receptors.

Recent studies used homology modeling and mutagenesis to elucidate the nature of the ligand-binding pocket in TAS2R7, TAS2R14, and TAS2R16 receptors. They reported that the binding pocket is flexible and wide open to accommodate molecules of diverse sizes and shapes, thus permitting chemical modifications among agonists. Although the molecular basis for the promiscuity of bitter receptors is attributed to their apparent flexible, spacious binding site, future work is required to elucidate the contact points between TAS2Rs' binding site residues and their agonists in terms of additional binding locations.

13.8.7.4: Bitter Receptors Ligand Binding Domain and Amino Acid Residues

Most of the TAS2R studies are based on molecular modeling, mutagenesis, and heterologous expression systems, which suggest that the ligand binding pocket is formed by several key residues in most TMDs (TMD1, TMD2, TMD3, TMD5, TMD6, and TMD7), except for TMD4.

Studies show similarities and differences regarding residues and positions involved in agonist-receptor interactions. However, most of them agree that besides position $N^{3.36}$ in TMD3 (superscript as per Ballestros-Weinstein nomenclature for class A GPCRs) and other residues ($L^{3.32}$, $L^{3.33}$, and $E^{3.37}$) in its proximity, play a role in agonist activation of several broadly tuned TAS2Rs (TAS2R1, TAS2R16, TAS2R30, TAS2R38, TAS2R46). In contrast, for the narrowly tuned TAS2R7, one position in TMD3 ($H94^{3.37}$) and another in TMD7 ($E264^{7.32}$) were crucial for metal ion binding. Mutagenesis and molecular modeling revealed that these two residues contribute to the metal ion binding pocket in TAS2R7. Moreover, metal ions bind distinctively to residues lining the binding pocket, and interestingly, the presence of calcium in the assay solution appears to affect the TAS2R7 response to metal ions. It is not clear how calcium affects metal ion binding to TAS2R7, but it might work cooperatively with certain ions and not others. Future studies focusing on structural interactions between the receptor and metal ions will provide further insights into how they activate the receptor.

Two studies of TMD2 suggest that position N2.61 is critical for binding in TAS2R1 and TAS2R46. Likewise, in TMD7, position 2657.39 is implicated in binding to TAS2R46 (E265) and TAS2R1 (I263). In TMD5, position H5.43 is implicated in binding in TAS2R16 and E5.46 in TAS2R1, while in TMD7, position $E^{7.32}$ is crucial for metal ion binding. These residues represent putative contact points for agonist interaction and form a pattern of being spaced one helical turn from each other.

Recent mutagenesis studies performed in broadly tuned TAS2R14 with agonists (aristolochic acid, picrotoxinin, thujone) found several residues in TMDs to be involved in agonist binding. However, in contrast to TAS2R10 and TAS2R46, mutation of TAS2R14 did not result in a complete loss of function for all agonists but a varied reduction in responsiveness or selectivity toward agonists. Among several mutants, only the mutation of W89A resulted in a complete loss of responsiveness against picrotoxinin, while others showed more subtle agonist-selective changes. This indicates that TAS2R14 is not streamlined for the most sensitive detection of selected agonists, but rather tailored to detect numerous diverse agonists, with comparatively lower apparent affinity.

The binding characteristics of bacterial acyl homoserine lactones (AHLs) on TAS2Rs (TAS2R4, TAS2R14, and TAS2R20) suggest the presence of a single orthosteric site situated close to the extracellular surface and reinforce the significant role of the extracellular loop structure (ECL2) in TAS2R ligand binding and activation. The crucial AHL binding residues in TAS2R4 and TAS2R14 are predominantly located in ECL2, while in TAS2R20, they are present in TMD3 and TMD7 helices. The ECL2 residues, N165 in TAS2R4, and R160 and K163 in TAS2R14 were crucial for lactone binding. In contrast, TAS2R20 residues W88 (TMD3) and Q265 (TMD7) are essential for agonist binding. In addition, the hydrophobic amino acids in the three TAS2Rs are important in directing the orientation of the hydrophobic acyl chains of lactones that facilitate receptor activation.

The transmembrane domain in GPCRs is composed mainly of hydrophobic amino acids accommodated in the plasma membrane. Therefore, the hydrophobic properties of the receptor binding pocket are important for any membrane-accessible agonist. Hydrophobic residues in TMD3 and TMD7 of TAS2R16 are important in forming a wide ligand-binding pocket that accommodates larger ligands like the β-glycosides. By using salicin analogs as TAS2R16 novel agonists (differing structurally from salicin in β-glucoside core constituents), several critical residues were identified that are required for signaling. Interestingly, these were identical to the residues critical for salicin signaling, except for W261, which was not required for activation by the analog 4-NP-β-mannoside. Importantly, all these residues are in the receptor's TMD helices or intracellular face, consistent with classical GPCR signal transduction. These results suggest that larger ligands bind to the wide binding pocket of TAS2R16 on the extracellular side,

and then their signal is transduced via conserved residues on the intracellular side. This can account for the broad spectrum of ligand recognition conferred by TAS2R16.

Unlike broadly tuned receptors, narrowly tuned ones like TAS2R7 show two different types of critical residues in ligand binding. The first type includes D86, W170, and S181, which are agonist-independent. Their mutation significantly reduces the ability of TAS2R7 to bind an agonist. A second group, consisting of D65 and W89, is selective for quinine and enhances binding to a specific category of ligand.

Despite the variation in the amino acid type and location important for agonist binding among receptors of the bitter family, for the most part, ligand binding pockets are present on the extracellular surface of TMDs or ECL2. The function of the residues at these binding pockets is dictated by multiple factors, including the type of ligand, the movements in TMDs, and the associated movement of ECL2 to accommodate the ligand. Structure–function studies have identified a conserved KLK/R motif in the intracellular carboxyl-terminal domain of 19 TAS2Rs that is critical for cell surface expression, trafficking, and receptor activation.

13.8.7.5: Agonist, Antagonist Binding and Modulation of Bitter Receptors

In simple pharmacological terms, an antagonist is a ligand that inhibits the biological response induced by an agonist and does not induce any response of its own. In contrast, a ligand that reduces the constitutive/basal activity of a GPCR is considered an inverse agonist. An antagonist acts as a competitive inhibitor to block receptor activity. Many agonists have been identified for bitter receptors, but few antagonists have been found, as described in Table 13.8.7. Finding an antagonist/inhibitor for bitter taste would help understand the TAS2R signal transduction mechanism, but it also has potential use in foods to overcome unwanted bitterness in consumer products. Such bitter blockers have been proposed to increase the palatability of bitter-tasting food and beverages, increase compliance in taking bitter-tasting drugs, especially children's formulations, and reduce or prevent off-target drug effects in extra-oral tissues.

TABLE 7 | Bitter taste receptor inhibitors with their IC50 values and other pharmacological properties.

Antagonist	Mode of action	Bitter receptors	Tested agonists	IC50 (μM)	References
GIV3727or 4-(2,2,3-trimethylcyclopentyl) butanoic acid	Competitive orthosteric inhibitor	31	acesulfameK	6.4	(Slack et al., 2010)
		43	Aristolochic acid	11.33	
		4	Colchicine	108	
		40	Cohumulone	6.24	
Gamma-aminobutyric acid (GABA)	Orthosteric inhibitor	4	Quinine	3.2	(Pydi et al., 2014b)
3β-hydroxydihydrocostunolide (3HDC)	ND	46	Absinthin	14.1	(Slack et al., 2010; Brockhoff et al., 2011)
			Andrographolide	4.9	
			Denatonium	6.8	
			Picrotoxinin	4.7	
			Strychnine	15.3	
3-hydroxypelenolide(3HP)	ND		Absinthin	57.8	(Brockhoff et al., 2011)
			Andrographolide	44.5	
			Denatonium	51.4	
			Picrotoxinin	22.9	
			Strychnine	84.9	
Probenecid	Allosteric inhibitor	16	Salicin	292	(Greene et al., 2011)
Sakuranetin	ND	31	Saccharin	5.5	(Fletcher et al., 2011)
6-Methoxysakuranetin	ND	31	Saccharin	10.2	(Fletcher et al., 2011)
Jaceosidin	ND	31	Saccharin	11.7	(Fletcher et al., 2011)
6,3′-dimethoxyflavanone	ND	39	Epicatechin gallate (ECG)	4075	(Roland et al., 2014)
			Denatonium	240	
6-Methoxyflavanone	ND	39	Epicatechin gallate (ECG)	479	(Roland et al., 2014
N,N-bis(carboxymethyl)-	ND	4	Quinine	0.059	(Pydi et al., 2014b)

Antagonist	Mode of action	Bitter receptors	Tested agonists	IC50 (μM)	References
l-lysine(BCML)					
(±) abscisic acid (ABA)	ND	4	Quinine	34.4	(Pydi et al., 2015)
ND, not determined.					

Table **13.8.7**: Bitter taste receptor inhibitors with IC50 values and other pharmacological properties.

To date, ∼12 bitter inhibitors have been reported to interact with only 10 TAS2Rs subtypes, as described in Table **13.8.5**. They do so by binding to transmembrane domains similar to agonists. GIV3727 (4-(2,2,3-trimethylcyclopentyl) butanoic acid) was the first TAS2R antagonist discovered and well-characterized structurally that acts as an orthosteric competitive antagonist for TAS2R31. It competes with the acesulfame K agonist both *in vitro* and *in vivo*. GIV3727 is moderately selective because it inhibits multiple bitter receptors, including TAS2R4, TAS2R40, and TAS2R43. Homology modeling revealed that the -COOH group in GIV3727 is important for ligand-receptor interactions, as its replacement with an ester or the corresponding alcohol abolished its antagonist activity. Moreover, a mutagenesis study in TAS2R31 and TAS2R43 revealed residues K265$^{7.39}$ and R268$^{7.39}$ in TMD7 to be crucial for its antagonistic activity. Similarly, another non-selective inhibitor, probenecid (p-(dipropylsulfamoyl) benzoic acid) was found to act as NAM of TAS2R16 activity and inhibits TAS2R38 and TAS2R43 as well. Two point mutations, P44T and N96T, in TMD3 of hTAS2R16 were found to suppress probenecid's ability to inhibit salicin activity significantly. Hydrophobicity seems important for their pharmacological activity, as observed for both probenecid and GIV3727. The sesquiterpene lactone, 3β-hydroxydihydrocostunolide (3HDC), is an interesting bitter blocker as it acts as a competitive antagonist of TAS2R46, TAS2R30, TAS2R40, yet activates TAS2R4, TAS2R10, TAS2R14, and TAS2R31 as an agonist.

Similarly, various flavanones were also noted as antagonists for TAS2R31 and TAS2R39 with varying efficacy. Taken together, most of the currently known antagonists are non-selective, and there is an urgent need for studies that focus on selective antagonists of major broadly tuned TAS2Rs (such as TAS2R10, TAS2R14, TAS2R16, and TAS2R46). To target bitterness in terms of food industry needs, potential peptide inhibitors from different protein sources such as hen protein hydrolysates (inhibits TAS2R4, TAS2R7, TAS2R14) and beef proteins (inhibits TAS2R4) are reported to be effective. Several umami glutamyl peptides isolated from soybeans have been found to act as non-competitive allosteric inhibitors of TAS2R16 against the salicin-induced response.

13.8.7.6: Constitutive Activity of Bitter Receptors

A phenomenon in GPCR activity is constitutive activity, essentially an active state occurring without an agonist, which has been demonstrated in more than 60 GPCRs. It is the production of a second messenger or downstream signaling by a receptor in a ligand-independent manner. The constitutive activity provides another possibility for taste inhibitor discovery using inverse agonists. Inverse agonists can inhibit both agonist-dependent and agonist-independent activity, while antagonists can inhibit only agonist-dependent activity. Interestingly, some mutations in GPCRs can lead to constitutive activity, and receptors with this characteristic (including constitutively active mutants or CAM) are important tools to investigate new bitter inhibitors. Although constitutive activity has not been observed naturally in TAS2Rs, when induced by mutation, these receptors provide a useful means to investigate the relationship between an active receptor conformation and inverse agonist pharmacology.

Molecular modeling and functional assays report five CAMs critical residues for TAS2Rs, one in TMD7 (S285$^{7.47}$) and four others in intracellular loop 3 (H214A, Q216A, V234A, and M237A). Of the five CAMs, only TAS2R4 with the H214A mutation shows a 10-fold increase in constitutive activity. This histidine residue is highly conserved in most TAS2Rs. Mutation of H214 (H214A) helped in finding two new inverse agonists (GABA and ABA) as summarized in Table **13.8.7**. Similar pharmacological approaches can be used to generate mutants of all TAS2Rs to screen for their inverse agonist/bitter taste blockers. However, for better characterization and interpretation of TAS2Rs, future *in vivo* studies should be performed to understand the functional relevance of these CAMs. At the same time, it is worth noting that the potential presence of endogenous agonists makes it difficult to determine the true constitutive activity of GPCRs, including TAS2Rs.

13.8.8: Kokumi Sensation Signal Transduction

In addition to the five basic tastes, sensations beyond these add another dimension to taste perception. One such example is "**kokumi**," which is distinct from the other five tastes in that it does not have a taste but instead induces a sensation of "mouthfulness," depth, thickness, and aftertaste in the flavors. Although this flavor has been used historically and is well recognized in Japanese cuisine, it was first characterized by researchers who isolated a kokumi taste substance from water extracts of garlic and onion and identified γ-glutamylcysteinylglycine or glutathione (GSH) as the main active ingredient of kokumi flavor. GSH is abundant in food-grade yeast extract and has been used to make foods more flavorful.

Kokumi signal transduction was unknown until **kokumi-calcium sensing receptors (CaSR)** expression was reported in a subpopulation of taste cells in mice and rats, which suggested it could function as a taste receptor for calcium and amino acids. However, its apparent role in kokumi stimuli detection was not confirmed. Ohsu et al. for the first time reported that kokumi peptides (GSH, γ-Glu-Val-Gly, and various γ-glutamyl peptides (see Table 13.8.8) signal through CaSR and can synergize with sweet, salty, and umami taste qualities to impart an augmented kokumi sensation, i.e., increased depth of flavor, which was further complemented by later studies. By using heterologous expression systems and human sensory analysis, these studies demonstrated that kokumi peptides impart kokumi sensation to sweet, salty, and umami taste via CaSR, as the kokumi component was specifically suppressed in the presence of the CaSR-specific NAM NPS-2143. Maruyama et al. identified a distinct population of taste cells expressing CaSR in mouse lingual tissue that did not express either sweet or umami receptors to validate this idea further. Notably, these cells are specifically responsive to kokumi substances and elicit a Ca^{2+} response to focally applied kokumi stimuli in mouse lingual slices. Moreover, this response was inhibited in the presence of NPS-2143. These findings support the idea that CaSR mediates kokumi sensation effects in TRCs

TABLE 8 | Kokumi sensation receptor agonists, allosteric modulators with concentrations used in cell-based assays.

Ca^{2+}	Orthosteric agonist/cation	1[a]	VFT
Mg^{2+}	Orthosteric agonist/cation	10[a]	VFT
Gd_{2+}	Orthosteric agonist/cation	0.02[a]	VFT
Al^{2+}	Orthosteric agonist/cation	0.5[a]	VFT
Sr^{2+}	Orthosteric agonist/cation	0.5[a]	VFT
Mn^{2+}	Orthosteric agonist/cation	0.5[a]	VFT
Ni^{2+}	Orthosteric agonist/cation	0.5[a]	VFT
Ba^{2+}	Orthosteric agonist/cation	0.2[a]	VFT

Ca^{2+}	Orthosteric agonist/cation	1[a]	VFT	Ca^{2+}
Spermidine	Orthosteric agonist/polyamine	0.002[a]	VFT	(Nemeth et al., 2018)
Neomycin	Orthosteric agonist/aminoglycoside antibiotic	0.06[a]	VFT	(Katz et al., 1992)
Gentamicin	Orthosteric agonist/aminoglycoside antibiotic	0.15[a]	VFT	Katz et al., 1992)
Kanamycin	Orthosteric agonist/aminoglycoside antibiotic	0.1	VFT	(Katz et al., 1992)
Amyloid β-peptides	Orthosteric agonist/Peptide	0.001–0.04	—	(Ye et al., 1997)
Poly-Lysine	Orthosteric agonist/peptide	0.03 μM[a]	VFT	(Brown et al., 1991; Nemeth et al., 2018)
Poly L-arginine	Orthosteric agonist/peptide	0.004 μM[a]	VFT	Brown et al., 1991; Nemeth et al., 2018)
Lysozyme	Agonist/protein	0.59[a]	ND	(Yamamoto et al., 2020)
Thaumatin	Agonist/protein	0.07[a]	ND	(Yamamoto et al., 2020)
Aromatic L-amino acids (Trp, Phe, His, Ala, Ser)	PAMs	10	VFT	(Conigrave et al., 2000; Mun et al., 2004; Geng et al., 2016)
Anions (SO_4^{2-})	NAM	10	VFT	(Geng et al., 2016)
Cinacalcet	PAM/phenylalkylamine	0.051 μM[a]	TMD	(Miedlich et al., 2002; Petrel et al., 2004; Nemeth et al., 2004)
Calindol	PAM/phenylalkylamine	0.31 μM[a]	TMD	Miedlich et al., 2002; Petrel et al., 2004)

NPS R-568	PAM/phenylalkylamine	0.5 μM[a]	TMD	(Miedlich et al., 2002; Petrel et al., 2004)
NPS R-467	PAM/phenylalkylamine	0.01	TMD	(Miedlich et al., 2002; Petrel et al., 2004)
γ-Glu-Val-Gly	PAM/Peptide	0.041 μM[a]	—	(Ohsu et al., 2010)
γ-Glu-Cys-Gly (Glutathione)	PAM/Peptide	76.5 μM[a]	VFT	(Ohsu et al., 2010; Wang et al., 2006
γ-Glu-Ala	PAM/Peptide	3.65 μM[a]	ND	(Wang et al., 2006; Ohsu et al., 2010)
γ -Glu-Val	PAM/Peptide	1.34 μM[a]	ND	(Wang et al., 2006; Ohsu et al., 2010)
γ -Glu-Cys	PAM/Peptide	0.45 μM[a]	VFT	(Ohsu et al., 2010; Wang et al., 2006)
γ -Glu-α-aminobutyryl-Gly (Opthalmic acid)	PAM/Peptide	0.018 μM[a]	ND	(Ohsu et al., 2010)
NPS2143	NAM	0.0003 (IC50)	TMD	(Gowen et al., 2000; Petrel et al., 2004)
Calhex 231	Mixed PAM/NAM	0.1–1 μM (PAM); 3–10 μM (NAM)	TMD	(Petrel et al., 2003; Petrel et al., 2004; Gregory et al., 2018)

Where VFT, Venus flytrap domain; TMD, transmembrane domain; ND, not determined. a shows EC50 value.

Table 13.8.8: Kokumi sensation receptor (kokumi-calcium sensing receptors (CaSR)) agonists, allosteric modulators with concentrations used in cell-based assays.

More recently, kokumi peptides have been found to have an extraoral physiological role in the gastrointestinal tract, where they stimulate the secretion of hormones (cholecystokinin and glucagon-like peptide 1 by activating CaSR. However, future studies with tissue-specific deletion of CaSR in taste buds would help delineate its role in taste physiology.

CaSR involvement in taste is a relatively recent discovery, but its central role in extracellular calcium homeostasis in mammals is well recognized. Diverse ligands activate CaSR, including cations (Ca^{2+} and Gd^{3+}), peptides, polyamines, and amino acids, as listed in Table 13.8.8. Unlike other taste modalities (sweet, bitter, and umami), CaSR–ligand binding and recruitment of G protein result in the activation of an intricate, amplifying signaling network that initiates numerous intracellular functions. The functional diversity of CaSR results from its ability to activate multiple Gα proteins (Gq/11, Gi/o, G12/13, and Gs), which subsequently affect multiple signaling pathways related to the pathophysiology of parathyroid hormone secretion, cancer, and metastasis.

Kokumi substrates activate CaSR and transmit their signal through Gαq/11 proteins, which further activate PLCβ, releasing intracellular Ca^{2+} stored through activation of IP3 receptor channels in the ER. Whether the kokumi pathway strictly relies on Gαq/11 protein or can also use Gα-gustducin, like other taste modalities, for downstream signaling, is still unknown, as shown in Figure 13.8.3. The growing number of reports on kokumi flavor signal transduction sheds light on its potential use as a flavor enhancer.

13.8.8.1: Structural, Molecular, and Conformational Changes of Kokumi Receptor

The kokumi-calcium-sensing receptors (CaSR) belong to the class C GPCR. Within this class, CaSR and metabotropic glutamate receptors (mGluRs) are known to function as disulfide-linked homodimers, as shown in Figure 13.8.4A. Structurally, the human CaSR is similar to sweet and umami taste receptors but differs in being a homodimer instead of a heterodimer. The ECD of CaSR not only senses nutrients (Ca^{2+}, L-Phe, and polypeptides, as summarized in Table 13.8.8) and allows ligands to modulate CaSR cooperatively, but is also required for its dimerization. The binding of Ca^{2+} and other ligands to the ECD changes the conformation

of the seven transmembrane domains, causing alterations in the intracellular loops and the intracellular domain (ICD), further triggering downstream signaling pathways. The ICD is relatively diverse among species and helps to control CaSR signaling in multiple ways by modulating receptor expression, trafficking, and desensitization.

Homology modeling, mutagenesis, and heterologous expression revealed distinct and closely located binding sites for Ca^{2+} and aromatic L-amino acids, in VFT and the cleft of the VFT, respectively. Notably, four putative Ca2+ binding sites of varying affinity have been predicted in the VFT of the CaSR, and the interaction between site 1 and the other three sites plays a central role in positive cooperativity in sensing Ca^{2+}. Besides Ca^{2+}, aromatic L amino acids (L-Trp, L-Phe) also activate the CaSR by binding adjacent to the VFT region through three serine and one threonine residue (S169/S170/S171/T145). Interestingly, the double mutation T145/S170 selectively impaired L amino acid (Phe, Trp, His) sensing of CaSR, while Ca^{2+} sensing remained intact.

The recent crystal structure of the entire extracellular domain of CaSR identified four novel Ca^{2+} binding sites in each protomer of the homodimer, including one at the homodimer interface, which does not correspond to any of the sites reported previously by Huang et al. It is unclear why these additional calcium-binding sites were not found in earlier studies. This might be due to the different expression systems, crystallization conditions, and analysis methods. The conditions of the more recent studies may have stabilized an active conformational state in which these calcium sites become available. Among these four Ca^{2+}-binding sites, site 4 seems most relevant to receptor activation as it directly participates in the active CaSR conformation. Moreover, a previously reported natural mutation G557E reduced the potency of Ca^{2+}, possibly by affecting backbone conformation, thereby weakening the affinity of Ca^{2+} for this site. This confirms that a Ca^{2+} ion at site 4 stabilizes the active conformation of the receptor by facilitating homodimer interactions between the membrane-proximal LBD2 region and CRD of CaSR.

The most interesting aspect of Ca^{2+} and L-amino acid interplay was reported by Zhang et al., who studied L-Phe binding characteristics by monitoring intracellular $[Ca^{2+}]_i$ oscillations in living cells and performing molecular dynamic simulations. Their findings supported a previous observation that the L-Phe binding pocket is adjacent to the Ca^{2+} binding site 1. Importantly, by binding to this site, L-Phe influences all Ca^{2+} binding sites in the VFT region and enhances CaSR functional cooperativity through positive heterotropic cooperativity to Ca^{2+}. Moreover, the dynamic communication of L-Phe at its predicted binding site in the hinge region with the Ca^{2+} binding sites not only influences the adjacent Ca^{2+} binding site 1 but also globally enhances cooperative activation of the receptor in response to alterations in extracellular Ca^{2+}.

The crystal structures of the entire ECD region of CaSR in the resting and active conformations have provided additional information about the dynamics between calcium and L-amino acid binding. Most importantly, using L-Trp, the study provided direct evidence that L-amino acids are CaSR co-agonists, acting concertedly with Ca^{2+} to achieve full receptor activation. Several lines of evidence support this contention: 1) L-Trp binds at the interdomain cleft of the VFT, which is a canonical agonist-binding site for class C GPCRs and shares a common receptor-binding mode with the endogenous agonists (amino acids or their analogs) of mGluR and $GABA_B$ receptors. 2) L-Trp interacts with both LBD1 and LBD2 in ECD to facilitate its closure, a crucial first step during CaSR activation. In contrast, no Ca^{2+} ion is found at the putative orthosteric agonist-binding site to induce domain closure. 3) Mutations of L-Trp-binding residues (S147A, S170A, Y218A, and E297K) severely reduced Ca^{2+} induced IP accumulation and intracellular Ca^{2+} mobilization , indicating that L-Trp is required for a Ca^{2+} induced receptor response. Notably, extracellular Ca2+ above a threshold level is required for amino-acid-mediated CaSR activation, and amino acids increase the receptor's sensitivity toward Ca^{2+}. Taken together, amino acids and Ca^{2+} ions act jointly to trigger CaSR activation.

Knowing that aromatic L-amino acids (Trp, Phe, His) are important tastants in kokumi flavor, CaSR becomes more relevant for taste biology. Moreover, the kokumi tripeptide, glutathione (GSH), and glutamyl peptide are suggested to bind allosterically to CaSR at the same site as L-amino acids and enhance its activity in the presence of 0.5–1 mM free calcium, thereby acting as a positive allosteric modulator. In addition, an ECD crystal structure might help to explain the structural and molecular details of the GSH binding pocket, such as the nature of critical residues and their binding characteristics. Given recent reports of calcium emerging as a taste modifier, it would be worth investigating how GSH and Ca^{2+} operate in kokumi human perception.

To summarize, the CaSR is a classic GPCR with seven transmembrane helices and is in subfamily class C along with taste receptors T1R1 and T1R3. As its name implies, it's activated by Ca^{2+} and kokumi peptides. It's expressed in taste cells. The peptide γ-EVG acts as a positive allosteric modulator (PAM). Figure **13.8.8** shows the Cryo-EM map of the human CaSR/γ-EVG complex.

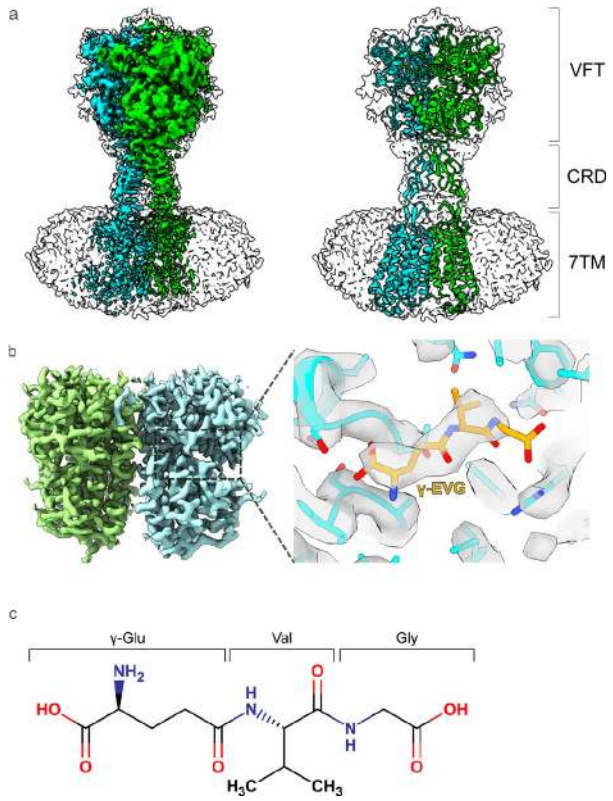


Figure 13.8.8: Cryo-EM map of the CaSR/γ-EVG complex. Yamaguchi, H., Kitajima, S., Suzuki, H. *et al*. Cryo-EM structure of the calcium-sensing receptor complexed with the kokumi substance γ-glutamyl-valyl-glycine. *Sci Rep* **15**, 3894 (2025). https://doi.org/10.1038/s41598-025-87999-1. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

Panel (a) Overall density map of CaSR (left) and cartoon of the model of CaSR in an active form (PDB: 7m3G) (right). CaSR is composed of a VFT domain, a CRD, and a 7TM domain. (**b**) Focus-refined map of the VFT domain of CaSR. Stick model of γ-EVG (orange) is overlaid with the cryo-EM density. (**c**) Chemical structure of γ-EVG. *Cryo-EM* cryo-electron microscopy, *CaSR* calcium-sensing receptor, *γ-EVG* γ-glutamyl-valyl-glycine, *VFT* Venus flytrap, *CRD* cysteine-rich domain, *7TM* seven-transmembrane.

Figure 13.8.9 shows an interactive iCn3D model of Cryo-EM Structure of calcium sensing receptor domain of CaSR in complex gamma-glutamyl-valyl-glycine as a kokumi substance γ-glutamyl-valyl-glycine or γ-EVG (**9J7I).**

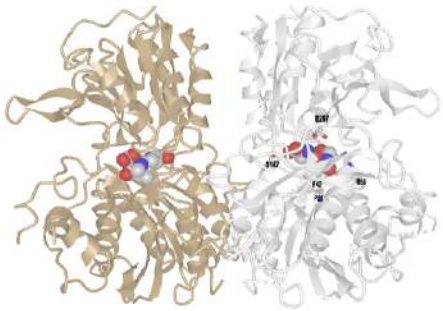


Figure 13.8.9: Taste receptor type 1 member 2 (TAS1R2) AlphaFold model (uniprot Q8TE23). (Copyright; author via source). Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...QVzVoftzAz3337.

The protein is showin as a homodimer. Peptides (spacefill) are bound to the orthosteric site. The grey subunit contains a bound γ-EVG. Key residues in binding the γ-EVG (Pro39, Phe42, Arg66, Ser147, and Glu297) are labeled as sticks. The brown subunit has

a GVG peptide bound. Peptides are shown as colored spheres.

13.8.8.1.1: Allosteric Modulation of Calcium-Sensing Receptor

CaSR is classically known to be involved in the pathophysiology of parathyroid and renal-related diseases by sensing calcium ions in the extracellular fluid. Research on related therapeutic applications has identified several classes of PAMs and NAMs that modulate CaSR agonist sensitivity. More recently, this has been applied to kokumi taste signal transduction.

13.8.8.2: Endogenous Modulators (L-amino Acids, Anions, and Glutathione Analogs)

Several studies based on molecular modeling and mutagenesis report L-amino acids (L-Phe, L-Tyr, L-His, and L-Trp) as PAMs because they enhance the Ca^{2+}-induced response of CaSR. Aromatic L-amino acids bind in the VFT domain and require a highly conserved five-residue binding motif (S147, S170, D190, Y218, and E297). Among these residues, E297 was identified through the natural mutation E297K as essential for structural and functional activity, as summarized in Table **13.8.8**.

As recently identified NAMs, anions SO_4^{2-} and PO_4^{3-} are important modulators of the Ca^{2+}-induced response. They bind in the VFT region and act as moderate NAMs for CaSR activity. Based on anomalous difference maps, four anion-binding sites were identified in the inactive and active CaSR ECD structures. Sites 1 and 3 are located above the interdomain cleft in LBD1, while site 4 lies in the LBD2 region. Sites 1 and 3 appear to stabilize the inactive conformation.

In contrast, site 2, present in both active and inactive conformations, appears important for receptor function as mutations in its residues (R66H, R69E, and S417L) abolished the Ca^{2+}-induced response. In addition, each protomer structure contains one Ca^{2+} ion and three SO_4^{2-} ions, contributing to the receptor's structural integrity. Taken together, anions along with Ca^{2+} and amino acids are involved in an intricate interplay for CaSR activation to maintain conformational equilibrium between inactive and active states.

As positive allosteric modulators, γ glutamyl peptides including glutathione (γGlu-Cys-Gly) and its analogs (see Table **13.8.8**) are predicted to have overlapping binding sites with L-amino acids in the VFT region. Kokumi peptides that activate CaSR resemble amino acids in having free α-amino and free α-carboxylate groups because they contain both amide bond formation between the γ-carboxylate group of L-glutamate and the α-amino group of its neighboring Cys residue. However, compared to amino acids, glutathione analogs have much larger side chains and are more potent activators of CaSR. Nonetheless, free sulfhydryl is not required for CaSR activation.

The crystal structure of ECD enables mapping the GSH binding site and investigating how GSH binding works in synergy with Ca^{2+} to modulate the kokumi sensation. NPS2143, the sole kokumi NAM identified to date, has been reported to inhibit kokumi taste sensation to GSH and its analogs. This provides an opportunity to screen for novel kokumi-enhancing molecules in a cell-based assay.

13.8.8.3: Synthetic Drugs as Allosteric Ligands of Calcium-Sensing Receptor

Because of its pathophysiological importance, various synthetic PAMs and NAMs of CaSR have been identified and are in clinical use. The allosteric modulation of CaSR by synthetic drugs has been recently reviewed. Since the 1990, the terms calcimimetics and calcilytics have been used for drugs that mimic or antagonize the effect of extracellular Ca2+ on CaSR activity, respectively. Pharmacologically, a calcimimetic activates the CaSR and includes agonists (type I) and allosteric ligands (type II). Most type I calcimimetics are inorganic or organic polycations (e.g., Mg^{2+}, Gd^{3+}, neomycin).

In contrast, type II calcimimetics are small naturally occurring molecules (aromatic amino acids or GSH) or synthetic drugs and peptides (NPS R-568, cinacalcet). Type II calcimimetics (like aromatic amino acids) bind in the ECD, while others (e.g., NPS R-568, NPS R-467) bind in the TMD of the CaSR. Calcilytics are thus small organic molecules that appear to act as NAMs and bind in the TMD of the receptor.

Homology modeling and mutational studies show that both PAMs and NAMs have overlapping but non-identical binding sites in TMD and can partially allosterically modulate CaSR activity in the complete absence of the ECD. Still, their potencies vary among structurally different compounds (see Table **13.8.8**). Several residues reportedly critical for allosteric modulation, $W818^{6.48}$, $F821^{6.51}$ (TMD6) and $E837^{7.39}$, $I841^{7.43}$ (TMD7), $R680^{3.28}$, $F684^{3.32}$, $F688^{3.36}$ (TMD3) impair calcimimetic and calcilytic induced CaSR signaling. Nevertheless, subtle differences in ligand–receptor interactions drive negative vs. positive modulation of CaSR

signaling by NPS2143 or cinacalcet and NPSR-568, respectively. The details of CaSR allosteric modulation by synthetic drugs are out of the scope of the current review; for a comprehensive explanation, refer to these studies.

13.8.8.4: Conclusion

Taste GPCR research has advanced rapidly over the past two decades, providing a more thorough understanding of receptor molecular pharmacology and signal transduction pathways. Except for the kokumi receptor ECD, high-resolution crystal structures for any taste receptor would be a major step toward designing novel and potent surrogate taste receptor ligands and selective antagonists. This has been a challenge due to low taste GPCR functional heterologous expression, appropriate post-translational modifications, high conformational flexibility, and low detergent stability. However, significant advancements in structural biology technologies of serial femtosecond crystallography using X-ray free-electron lasers and high-resolution cryo-electron microscopy provide promising tools for understanding conformational dynamics and visualizing the receptor activation process with high spatial and temporal resolution. The physiological relevance of taste GPCRs will be further advanced through *in vivo* studies to help provide information on potential synergies in taste signal transduction mechanisms, particularly among bitter, umami, sweet, and kokumi receptors.

13.8.9: Summary

This chapter provides a comprehensive exploration of the molecular and structural underpinnings of taste perception mediated by G protein-coupled receptors (GPCRs). It details how distinct classes of taste receptors—responsible for sweet, umami, bitter, and kokumi sensations—transduce chemical signals from food into neuronal responses, ultimately shaping taste perception.

Taste Receptor Cell Architecture and GPCR Classification

- **Taste Bud Organization:**
  Taste receptor cells (TRCs) are organized into taste buds distributed on the tongue and palate. Type II cells, which express GPCRs, are primarily responsible for detecting sweet, umami, and bitter stimuli, while type III cells mediate sour taste and type I cells may detect salt.
- **Receptor Families:**
  The chapter distinguishes between Type I GPCRs (TAS1Rs) that function as heterodimers—TAS1R2/TAS1R3 for sweet and TAS1R1/TAS1R3 for umami—and Type II GPCRs, which include monomeric bitter receptors (TAS2Rs) and the calcium-sensing receptor (CaSR) associated with kokumi sensation. Structural domains such as the large extracellular Venus flytrap domain (VFT), cysteine-rich domain (CRD), and seven transmembrane domain (TMD) are highlighted as critical for ligand binding and receptor activation.

Signal Transduction Mechanisms

- **Intracellular Pathways:**
  Upon binding of tastants, taste GPCRs activate heterotrimeric G proteins (e.g., Gα-gustducin for sweet, umami, and bitter receptors, and Gα-q/11 for CaSR in kokumi signaling). The released Gβγ subunits stimulate phospholipase Cβ2 (PLCβ2), which catalyzes the production of IP_3 and DAG. IP_3 triggers Ca^{2+} release from endoplasmic reticulum stores, activating downstream ion channels (e.g., TRPM4/5) that depolarize TRCs and induce the release of ATP—a key neurotransmitter in taste signaling.
- **Neurotransmission:**
  The ATP released from taste cells activates purinergic receptors on afferent cranial nerves, transmitting signals to the gustatory cortex and culminating in the perception of taste.

Receptor–Ligand Interactions and Allosteric Modulation

- **Orthosteric Binding:**
  Sweet and umami receptors bind a variety of natural and synthetic ligands in their VFT domains, leading to conformational changes that facilitate dimerization and G protein coupling. In contrast, bitter receptors are broadly tuned to diverse ligands using a single, often spacious binding pocket located within the TMD and extracellular loops.

- **Allosteric Modulation:**
  The chapter describes how positive allosteric modulators (PAMs) and negative allosteric modulators (NAMs) can enhance or inhibit receptor responses without directly activating the receptor. For example, compounds such as neohesperidin dihydrochalcone (NHDC) enhance sweet taste receptor activity, whereas lactisole acts as a NAM by binding to the TMD of TAS1R3, thereby inhibiting both sweet and umami responses. Similar allosteric interactions are explored for the umami receptor and CaSR (kokumi receptor), where endogenous ligands like 5′-ribonucleotides (IMP, GMP) and glutathione (GSH) modulate receptor activity.

- **Molecular Modeling and Mutagenesis:**
  Advances in computational modeling (using tools like AlphaFold and AutoDock Vina) and site-directed mutagenesis have been pivotal in identifying key residues involved in ligand binding and receptor activation. These studies have elucidated how both the extracellular and transmembrane domains contribute to receptor conformation, dimerization, and signal transduction.

Physiological and Pharmacological Implications

- **Taste Perception and Food Science:**
  The integration of multiple signaling pathways allows the taste system to discern complex flavor profiles, influencing dietary choices and food enjoyment. Understanding these mechanisms offers opportunities to design novel taste modifiers—both enhancers and inhibitors—that can improve food palatability and patient compliance with bitter-tasting medications.

- **Therapeutic and Research Challenges:**
  Despite significant progress, the crystallization and high-resolution structural determination of taste GPCRs remain challenging due to their low expression levels, conformational flexibility, and stability issues. Emerging technologies such as cryo-electron microscopy and serial femtosecond crystallography are promising tools that may overcome these barriers, leading to the development of more potent and selective taste receptor ligands.

In summary, this chapter synthesizes current knowledge on the structure and function of taste GPCRs, providing insights into the complex mechanisms by which these receptors detect and transduce taste stimuli. It emphasizes the interplay between receptor architecture, ligand binding, allosteric modulation, and intracellular signaling, highlighting both the physiological relevance and the pharmacological potential of targeting taste receptors in food science and therapeutic applications.

This page titled 13.8: Signal Transduction - Taste (Gustation) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

13.9: Signal Transduction - Temperature

Learning Goals (ChatGPT o3-mini)

- **Explain the Physiological Importance of Thermal Regulation:**
  - Describe how mammals and other animals maintain a nearly constant body temperature despite environmental fluctuations.
  - Identify the roles of external (skin) and internal (visceral and central) thermoreceptors in monitoring and regulating body temperature.
- **Understand Temperature Dependence in Biochemical Processes:**
  - Define the concept of the temperature coefficient (Q_{10}) and explain its relevance in ion channel kinetics and enzyme-catalyzed reactions.
  - Discuss how temperature affects membrane potentials and the kinetics of action potential generation in neurons.
- **Describe the Structure and Function of TREK Channels:**
  - Outline the molecular architecture of TREK channels (TREK1, TREK2, TRAAK), including their four transmembrane segments and two pore domains.
  - Explain how TREK channels are modulated by various stimuli—temperature, pH, membrane stretch, fatty acids, and anesthetics—and how their activation leads to hyperpolarization or depolarization of thermoreceptors.
- **Analyze the Role of TREK Channels in Thermosensation:**
  - Compare the temperature-dependent activation of TREK channels at physiological (37 °C) versus room temperature, and discuss how this influences the firing of thermoreceptors.
  - Evaluate experimental evidence (e.g., knockout (KO) studies, patch-clamp recordings) that demonstrates how TREK channel activity modulates the perception of heat, cold, and thermal pain.
- **Explain the Diversity of TRP Channels in Thermal Sensing:**
  - Identify key TRP channel subtypes that respond to heat (e.g., TRPV1, TRPV2, TRPV3, TRPV4) and cold (e.g., TRPM8, TRPA1) and describe their respective activation thresholds and biophysical properties.
  - Discuss how TRP channels contribute to the perception of noxious versus innocuous thermal stimuli.
- **Compare and Contrast TREK and TRP Channel Mechanisms:**
  - Contrast the ionic conductances and reversal potentials of TREK (potassium channels) versus TRP channels (cationic channels) and explain the opposing effects these channels have on neuronal excitability.
  - Analyze how the balance between the activation of hyperpolarizing TREK channels and depolarizing TRP channels shapes the overall thermosensory response.
- **Discuss the Molecular Origin of Thermosensitivity:**
  - Explain proposed mechanisms by which temperature changes induce conformational rearrangements in TREK and TRP channels.
  - Evaluate how intracellular factors, such as second messengers or membrane tension, might contribute to the temperature sensitivity of these channels.
- **Integrate Experimental Approaches in Thermosensory Research:**
  - Summarize the key techniques (e.g., heterologous expression systems, patch-clamp electrophysiology, knockout mouse models) used to study thermosensitive ion channels.
  - Critically assess how these methods contribute to our understanding of thermal pain and the role of specific channels in thermoregulation.
- **Explore Clinical and Translational Implications:**

- Discuss how alterations in TREK and TRP channel function can contribute to thermal hyperalgesia, cold allodynia, and other sensory disorders.
- Identify potential therapeutic strategies targeting these channels to modulate pain and improve thermal comfort.

Achieving these goals will equip students with a comprehensive understanding of how temperature influences ion channel behavior, the cellular basis of thermoreception, and the interplay between different channel types in generating thermal sensations and pain.

This chapter section is taken in its entirety from: **Ion Channels and Thermosensitivity: TRP, TREK, or Both?** Lamas et al. *Int. J. Mol. Sci.* **2019**, *20*(10), 2371; **https://doi.org/10.3390/ijms20102371.** Creative Commons Attribution (CC BY) license (**http://creativecommons.org/licenses/by/4.0/**).

13.9.1: Introduction

Mammals and other animals spend large amounts of energy maintaining a nearly constant body temperature, irrespective of the environment's temperature. The mechanisms controlling thermal regulation are complex and often rely on negative feedback, where it is first necessary to determine the body and ambient temperature. External receptor cells, found mainly in the skin, can sense the environment's temperature. In contrast, body temperature is sensed by internal receptors expressed by cells in several internal organs. Traditionally, only the skin and core thermoreceptors (spinal cord, hypothalamus) have attracted the attention of researchers. Still, more recently, some very interesting information has emerged regarding visceral thermal receptors, even in humans. Although a hypothesis conceived many years ago, the terminals of receptor neurons are thought to contain branches of nerve fibers without any apparent structural specialization. Indeed, only recently have we begun to understand the molecular basis of thermoreception by cells.

Many biochemical processes, such as chemical reactions and physical processes, like conformational changes, are extraordinarily dependent on temperature. Although these processes generally occur faster at higher temperatures, the relationships can be very complex [3]. If we consider the nervous system (NS), the effects of temperature on the resting membrane potential (RMP) were the first to be studied, as were its effects on the kinetics and speed of compound and single action potentials, long before ion channels were demonstrated.

All neurons and ion channels are affected by changes in temperature, not least because channel gating is generally a temperature-dependent process. However, only some neurons can be called thermoreceptors, and very few ion channel types can be designated thermosensors. Only channels with a temperature coefficient (Q10) ≥2–5 are considered temperature dependent. (Q10 is the ratio of a reaction at two different temperatures that differ by 10 °C. See Chapter 32.11 for more details.) Thermoreceptors are sensitive to temperature changes rather than the value of the temperature itself, probably due to their characteristic strong adaptation. These receptors are classified into two groups depending on whether their discharge frequency increases when heated or cooled, as shown in Figure **13.9.1**. Based on this classification, it is common to speak of four thermal sensations (cold −10 to 15 °C, cool 16–30 °C, warm 31–42 °C, and hot 43–60 °C), whereby cold and hot are potentially noxious and/or painful.

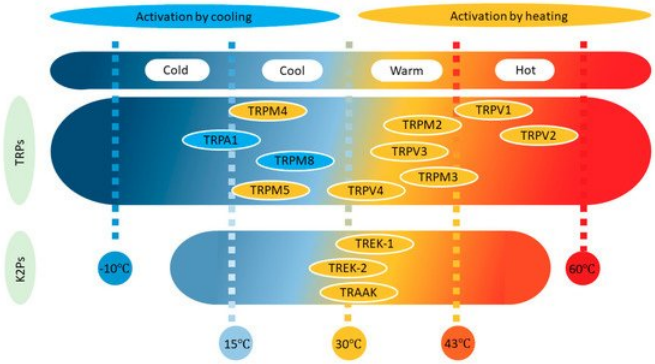


Figure **13.9.1**: Distribution of transient receptor potential (TRP) and TWIK-related potassium (TREK) channels as a function of their temperature threshold. Note that while temperature increases (orange), TREK channels are activated, but TRP channels may also be activated by lowering the temperature (blue).

The modulation of TWIK-related potassium (TREK) channels by temperature has been touched on in several reviews, yet very few have dealt exclusively with this exciting topic. Conversely, after transient receptor potential (TRP) channels sensitive to temperature were discovered, they were studied extensively to understand how thermal stimuli were transduced. Such interest led to the appearance of good reviews covering this issue. In this review, we will focus on the less well-known role of TREK channels in thermosensation, and we will compare the behavior of these channels to that of TRP channels. Other thermosensitive proteins, like the Na/K ATPase and epithelial sodium channels (ENaC), or P2X (ion channels activated by ATP) receptors, have also been described. While these should also receive attention, we consider this to fall beyond the scope of this review. Indeed, cell thermosensitivity seems to be governed by the interplay of several channel types, as reported in hypothalamic neurons.

13.9.2: TREK Channels

The TWIK-related potassium channel (TREK) subfamily belongs to the two-pore domain potassium channel family (K2P). It comprises three members: TREK1, TREK2, and TRAAK (TWIK-related arachidonic acid-activated potassium channel). These are background potassium channels characteristically modulated by several physical and chemical stimuli, such as membrane stretch, pH, unsaturated fatty acids, general anesthetics, and temperature. TREK channels generally display very weak activity at room temperature and normal pressure, even when overexpressed in heterologous systems. However, their activity increases strongly when several different stimuli are applied, including an increase in temperature. From a physiological point of view, it is important to note that at 37 °C, all three members of the TREK subfamily respond to stimuli (pH, membrane stretch, or arachidonic acid), much like they do at room temperature. TREK channels may fulfill a dual role in the transmission of thermal pain. Thus, their strong activation by noxious heat results in an outward current that provokes membrane hyperpolarization and a reduction of thermoreceptor firing, provoking heat-pain relief. Conversely, inhibition of TREK channels by noxious cold should depolarize thermoreceptors and increase their excitability, cooperating in the transduction of noxious cold sensations, as shown in Figure 13.9.2.

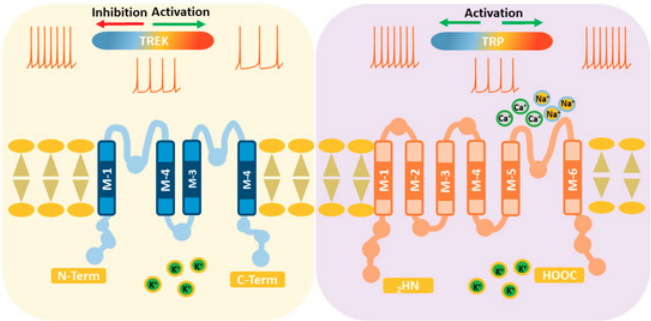


Figure 13.9.2: Proteins forming TREK channels have four transmembrane segments and two pore domains so that they assemble as dimers. Activation of TREK channels by increasing the temperature reduces the excitability of thermoreceptors (left). In contrast, TRP channels possess six transmembrane segments and only one pore domain, so they are tetramers. Activation of TRPs increases the excitability of thermoreceptors, but some are activated by an increase and others by a decrease in temperature (right).

13.9.2.1: TREK1

Soon after their discovery, it was shown that TREK1 channels are strongly and reversibly activated by an increase in temperature when expressed in heterologous systems (cell lines derived from kidney (COS) cells, oocytes). Let's consider that these are mostly voltage-independent channels open at resting potentials. TREK1 channels should function as cold sensors because low temperatures dampen their activity and depolarize these thermoreceptors (see Figure 13.9.2). Many authors have demonstrated that macroscopic TREK1 currents are strongly outwardly rectifying at room temperature. While the outward current is not evident at 12 °C, it is strongly enhanced at 37 °C, and the current progressively increases as the temperature gradually increases. Indeed, the current increases around 7-fold with an increase of 10 °C in the range of 14 to 42 degrees, and importantly, maximal sensitivity (0.9-fold per degree) was reached at nearly physiological temperatures, between 32 and 37 °C. The current induced by heating is also outwardly rectifying, reversing at potentials close to the equilibrium potential for potassium (E_K). In heterologous systems, the activation of TREK1 by temperature may be reversibly inhibited by cAMP, and this inhibition is suppressed by mutation of the C-terminal region that harbors a phosphorylation site for protein kinase A (PKA). Moreover, chicken embryonic atrial myocytes express TREK-like currents, and they have a resting membrane potential of around −20 mV in culture, which increases to −70 mV when the temperature rises to 35 °C, a change that was ascribed to the activation of TREK1/2 channels. In voltage-clamp, the

outward current recorded at +60 mV increased 9-fold. TREK1 and TWIK-related arachidonic acid-activated potassium (TRAAK) channels have been proposed to shut down the firing of hippocampal neurons when the temperature rises too high.

Figure 13.9.a shows an interactive iCn3D model of the mouse temperature sensitive K2P2.1 (TREK-1) potassium channels (6W84)

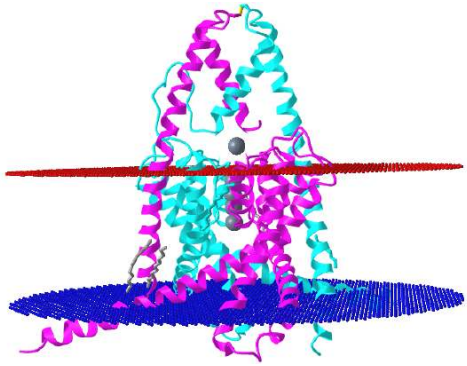


Figure 13.9.a: Mouse temperature sensitive K2P2.1 (TREK-1) potassium channels (6W84). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...U28PvTKakHpHW9

The red dots represent the outer leaflet; The gray spheres are potassium ions.

The threshold for the activation of slowly conducting C-fibers by noxious heat (30–50 °C) recorded in a skin-nerve preparation decreases in TREK1 knockout (KO) mice, and the range of activation of these fibers by heat corresponds closely to the range in which TREK1 is activated (30–45 °C). The number of action potentials in response to a heating ramp (30–50 °C) was higher in the KO mice, although the response of C-fibers to a cooling ramp (32–10 °C) was similar in native and KO mice. Indeed, TREK1 KO mice were hypersensitive to thermal pain up to 50 °C but not at higher temperatures (52–56 °C), indicating that TREK1 channels may be important for the perception of low-threshold but not high-threshold thermal stimuli, for which TRP channels may be more important. Accordingly, the proportion of small diameter, cultured, dorsal root ganglia (DRG) neurons that respond to noxious heat (34%) increases in TREK1 KO (64%) and TREK1/TRAAK KO (74%) mice, as does the proportion of heat-responsive C-fibers in nerve-skin preparations from the single and double KO. TREK1 and TRAAK channels may counteract the stimulatory effect of heat-activated TRP channels in pain-transducing fibers when temperatures increase, such that the overall response may reflect a balance of the activity of these two functionally contrasting channel types. The threshold of thermoreceptors should certainly increase in the presence of TREK channels as temperatures increase.

Cooling of DRG neurons in culture from 32 to 20 °C induces a depolarization of about 10 mV and often the firing of action potentials, an effect shown to be due to the inhibition of a background potassium current. Accordingly, the inhibition of a native TREK1-like current may underlie the excitation (depolarization and firing) produced by cold in small, cultured trigeminal ganglion (TG) neurons, as seen in Figure 13.9.2. Interestingly, cold induces subthreshold oscillations in cold-sensitive DRG neurons. Transduction seems rather complex, involving dampening a hyperpolarization-activated cationic current and a permissive role for a slowly inactivating potassium current. Interestingly, the TREK1/TRAAK double knockout (KO) mutant shows a consistent cool allodynia (pain due to a stimulus that does not normally provoke pain), and oxaliplatin, a cancer therapy that causes peripheral nerve neuropathy, exacerbates cold sensitivity in many patients and animals, inducing allodynia to cool temperatures.

Neither the deletion (or knockout - KO) of TREK1 nor TRAAK increases the fraction of small DRG neurons sensitive to noxious cold stimuli (below 20 °C and down to about 10 °C). However, the TREK1/TRAAK KO and the triple TREK1/TREK2/TRAAK KO significantly increased such neurons. Similar results were obtained when recording C-fibers in a skin-nerve preparation, in which case the double KO C-fibers fired more strongly than the single TRAAK KO and wild-type fibers. Oxaliplatin also induced hypersensitivity to noxious cold temperatures, while double and triple KO mice, but not the TREK1 KO mice, are hypersensitive to cold, which is not further affected by oxaliplatin. Hence, deleting two of the three TREK channels appears to be sufficient to reach maximal hypersensitization. The neuroprotective agent riluzole induces an analgesic effect against painful cold in normal animals and in oxaliplatin-pretreated TREK2 KO and TRAAK KO animals. However, riluzole did not affect pain sensitivity in TREK1 KO animals treated with oxaliplatin, in animals treated with the TREK1 inhibitor spadin, or in untreated TREK1 KO mice or triple KO animals. Similarly, a presumed TREK leak outward current recorded in DRG neurons was inhibited by riluzole and fluoxetine at 22 and 30 °C but not at 14 °C, probably because the current was already inhibited at low temperatures. Together, these experiments

suggest that TREK1 channels are essential in perceiving noxious cold and that TREK1 and TRAAK channels work together in sensing cold.

Cell-attached patches demonstrated that the basal activity of expressed TREK1 channels is insignificant at room temperature, gradually increasing as the temperature rises (17-fold for an increase of 20 °C) and with a threshold around 25 °C. The current activated by temperature also displays outward rectification and reverses around the equilibrium potential for K^+, although the single-channel conductance remains unaffected TREK1-like channels naturally expressed in cardiac ventricular myocytes and DRGs and recorded in cell-attached patches, do not open at 24 °C, yet they are very active at 37 °C Surprisingly, temperature increases fail to modulate TREK1 activity in outside-out and inside-out patches, but under the same conditions, TREK1 is still strongly activated by arachidonic acid

TREK1 channels are ideally positioned to act as thermosensors because they are expressed in structures related to thermosensitivity and thermoregulation, such as dorsal root ganglion (DRGs), the trigeminal ganglion (TG), nodose ganglia (NG), or the anterior and preoptic hypothalamus.

13.9.2.2: TREK2

Heterologously expressed TREK2 channels also produce strong outward rectification when recorded in whole-cell configuration at room temperature, which increases greatly at temperatures around 37 °C in several heterologous systems. In COS cells, a small TREK2 current was observed at 0 mV that augmented progressively with a gradual rise in temperature to about 40 °C. Notwithstanding, the response of TREK2 to abrupt changes in temperature was rapid. Importantly, the IVs of the TREK2 current at different temperatures (24 and 37 °C) showed that the effect of temperature was not voltage-dependent: both inward and outward currents increased to the same degree. In this range of temperatures, the current increased 14-fold per 10 °C, indicating a very strong temperature dependence that was even bigger than that of TREK1. Much like TREK1, TREK2 responds to temperature changes around the physiological range, with current activated reasonably well at 37 °C and resting membrane potential (RMP). Most experiments on TREK channels have been carried out at room temperature and 0 mV. However, in the future, these currents should be investigated using more physiological parameters, around a resting potential and 37 °C, providing a more precise idea of their role in the behavior of central neurons.

Figure **13.9.*b*** shows an interactive iCn3D model of the human two-pore domain temperature-sensitive potassium ion channel TREK2 (K2P10.1)(4BW5)

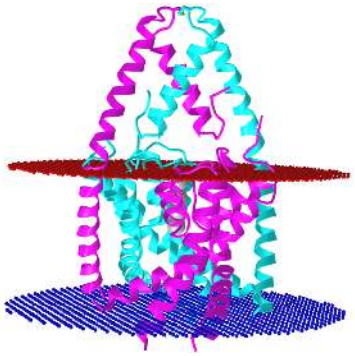


Figure **13.9.*b***: Human two-pore domain temperature-sensitive potassium ion channel TREK2 (K2P10.1)(4BW5). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wSt9GE2pbDwfq5

Cerebellar granule and DRG neurons expressed native TREK2-like channels with weak activity at 24 °C in cell-attached patches. Yet, when the temperature increased to 37 or 41°C they became very active at all voltages (−80 to +80 mV). Moreover, cultured cortical astrocytes have TREK2-like whole-cell outward currents strongly enhanced in the 23–40 °C temperature range. Interestingly, ischemia significantly augmented the outward current provoked by an increase in temperature in these astrocytes In addition, it was recently reported that TREK2 channels contribute about 10 mV to the RMP of DRG neurons at about 30 °C Furthermore, single TREK2 and triple TREK1/TREK2/TRAAK knockout (KO) mice were more sensitive to warm temperatures (40–42 °C) when tested with the tail-flick reflex.

Using a skin-nerve preparation, it was demonstrated that the proportion of heat-sensitive C-fibers and their activity (the number of action potentials) increased in the TREK2 and triple KO mice when temperatures rose to noxious heat levels (ramped from 30 to 50

°C). In contrast, the temperature threshold for firing decreased. At high temperatures (between 40 and 50 °C), the triple but not the single KO fibers were more active than their wild-type counterparts, indicating that TREK2 regulates C-fiber responses at temperatures below 40 °C. In contrast, other TREK channels participate in these responses at higher temperatures. Both KOs suffered hyperalgesia at temperatures around 45 °C, but only the triple KO showed the same behavior above this temperature.

The withdrawal latency in the tail immersion test was reduced in both the TREK2 KO and the triple KO mice when innocuous cooling temperatures were tested (20–25 °C). As such, the KOs show enhanced sensitivity to temperatures in the normal range, and similar results were obtained in a temperature preference test. The percentage of C-fibers responding to moderate cold (30–21 °C) was higher in single and triple KOs when compared to those recorded from the nerve-skin preparation of wild-type mice. Interestingly, the cold threshold for C-fiber firing (21 °C) was lower in the triple KO (24 °C) but not in the TREK2 KO (23 °C) mice. Moreover, oxaliplatin induces mice to spend more time on a hot plate (30 °C) than on a cold plate (20–25 °C) compared to untreated animals, indicating that neuropathic mice have enhanced sensitivity to moderate cold. It has been suggested that TREK2 is implicated in the neuropathic hypersensitivity induced by this drug, and indeed, oxaliplatin almost halved the TREK2 mRNA detected in DRG neurons. Generally, the data suggest that TREK2 channels may be essential for controlling the C-fiber response to cold at moderate temperatures. The tail immersion test showed that triple KO mice were hypersensitive to noxious cold temperatures (15–5 °C), while the single TREK2 KO mice behaved much like the wild-type mice. Moreover, very similar results were obtained in the nocifensive dynamic cold plate test. Accordingly, it was suggested that TREK2 may not be important in noxious cold sensitivity but might be essential for thermoreception at moderately cool temperatures (25–20 °C).

A clear, fast, and reversible increase in activity was also reported for single TREK2 channels in cell-attached patches held at −40 mV when the temperature increased (24 to 37 °C), with a threshold for this increase at 25 °C (from 24 °C) and not affecting the conductance. It should be noted that in these circumstances, the activity of TREK2 single channels was very low at 24 °C. Significantly, neither TASK3 nor TRESK2 showed such dependence on temperature. However, like TREK1, the activity of TREK2 in inside-out patches was not modified by changes in temperature (24 to 42 °C). Finally, it is important to consider that TREK2 channels are expressed strongly in the DRG, TG, and hypothalamus, yet less than TREK1 in the NG.

13.9.2.3: TRAAK

Much like the other family members, TRAAK currents showed a strong open-channel outward rectification when recorded in whole-cell configuration, and these currents increase strongly when the temperature rises (24 to 42 °C). Moreover, in culture, the percentage of small-diameter DRG neurons responding to noxious heat is increased in TRAAK and TRAAK/TREK1 knockout (KO) mice. Consistently in skin-nerve preparations, the percentage of fibers responding to heating (30–50 °C) and the number of action potentials in response to a heating ramp also increase, while the firing threshold is reduced. TRAAK and TRAAK/TREK1 KO mice suffer heat hyperalgesia when evaluated in the tail immersion test at 46–50 °C. Moreover, the double but not the single KO also shows hypersensitivity at higher temperatures (52–56 °C) in the hot plate test.

Knock-out of TRAAK did not modify the percentage of DRG neurons in culture that respond to noxious (12 °C) cold. Moreover, in the cold plate assay, TRAAK KO mice behave like wild-type mice, whereas TREK1/TRAAK KO mice are more sensitive to cooling in the 10 to 20 °C range. The activity of single TRAAK channels heterologously expressed in COS cells and recorded in cell-attached patches at −40 mV was very low at 24 °C. Yet, it increased progressively as the temperature rose from 24 to 37 °C. The threshold for activity was around 30 °C, slightly higher than that reported for TREK1 and TREK2. However, the behavior of TRAAK channels in inside-out patches mimics that of TREK1 and TREK2, such that their activity was not affected by changes in temperature (from 24 to 42 °C). Native TRAAK-like channels in DRG neurons displayed little activity at room temperature, but there was clear activity in all cell-attached patches at 37 °C [36, 48]. Finally, TRAAK channels are clearly expressed in the hypothalamus, TG, and DRG, yet they are only weakly expressed in the NG.

13.9.2.4: Molecular Origin of Thermosensitivity

When first discovered, mouse TREK1 was reported to have four transmembrane segments, two pore domains, and a sequence of 370 aa [37]. The activation of heterologously expressed TREK1 currents by increasing temperature is unaffected by the deletion of the cytoplasmic N-terminal region. By contrast, partial deletion of the C-terminal region (Δ103) or replacement of this region with that of TASK1 strongly dampens the activation of TREK1 by heat. The sensitivity of the TREK1 channels to temperature can be eliminated by mutating helix 1 of the pore (G137I), suggesting that temperature affects the TREK1 and TREK2 channels by manipulating the C-type gate. It was suggested that functional coupling between the C-terminal domain and the C-type gate

through the M4 segment is crucial for the heat sensitivity of the TREK1 channel. Thus, it is tempting to speculate that increasing the affinity of the C-terminal domain for phospholipids of the inner leaflet would increase the activity of TREK1 by heat, as proposed for other stimuli like stretch, PUFAs, phospholipids, or pH. Conversely, dissociating this domain from the membrane would result in TREK1 inhibition. Surprisingly, replacing the C-terminus of TREK2 with that of TASK3 did not reduce the channel's sensitivity to changes in temperature in the range of 24 to 37 °C under similar conditions. However, it became insensitive to pH and arachidonic acid. Heat enhances the activity of TREK1, TREK2, and TRAAK in whole-cell and cell-attached recordings, yet not in outside- and inside-out patches, indicating that the integrity of the cell, and probably also a second messenger, is necessary for this modulation. The contribution of TREK channels to maintaining the RMP has often been questioned; however, this assertion is mostly based on experiments carried out at room temperature. Thus, new experiments should be performed at physiological temperatures to ascertain the role of these channels in both the RMP and neuronal excitability.

13.9.3: TRP Channels

Six transient receptor potential (TRP) channels are considered thermosensors, four responding to heat and two to cold. Temperature-sensitive TRP channels (Thermo-TRP) are extremely dependent on temperature, showing very high Q10 values (>20).

13.9.3.1: Heat-Sensitive TRP Channels

Four TRP subtypes are activated by increased temperature (see Figure **13.9.1**). Two of them respond to warm stimuli (TRPV4 Warm >27 °C and TRPV3 Warm >34 °C), and the other two to hot-painful stimuli (TRPV1 Hot >43 °C and TRPV2 Hot >52 °C).

TRPV1s are voltage and temperature-dependent channels that display outward rectification when expressed in human embryonic kidney (HEK) cells. This is strongly enhanced by heating (to 48 °C) and by capsaicin.

Figure **13.9.*c*** shows an interactive iCn3D model of human TRPV1 with capsaicin at 48 degrees Celsius in an open state (7LPE)

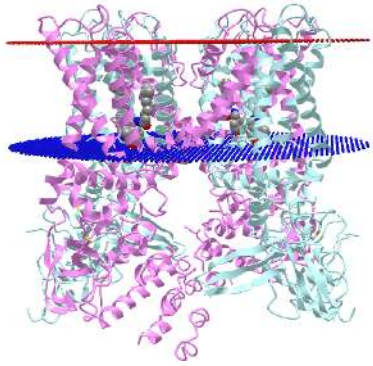


Figure **13.9.*c***: Human TRPV1 with capsaicin at 48 degrees Celsius in an open state (7LPE). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...7Hsf5zpDQoQu99

At room temperature, the current passing through these channels is negligible below 0 mV, but at 42 °C, the channel activates more or less between −100 and +50 mV. These cationic channels are ten times more permeable to Ca^{2+} than to Na^{+} (P_{Ca}/P_{Na} ~10) and are thought to be sensors for noxious heat but not activated by innocuous heat. Indeed, the response to noxious heat in mice lacking TRPV1 (KO) or DRG neurons was weaker. However, other channels may also contribute to the perception of noxious thermal stimuli because heat still evokes receptor activation in several preparations. The NG sensory neurons that innervate the lungs produce an inward current in response to an elevation in temperature (from 23 to 41 °C, with a threshold around 35 °C and a Q10 of about 30 in the range of 35–41 °C) as well as membrane depolarization and action potential firing. This response was ascribed to the presence of TRPV1 channels, even though the participation of TRPV2-4 could not be ruled out. We obtained similar results with NG neurons in culture, although these were slightly more complex because a hyperpolarization was observed before the depolarization and firing (unpublished data). Interestingly, inflammatory mediators like ATP and bradykinin strongly reduce the threshold of TRPV1 activation (30 °C) such that warm temperatures become painful. TRPV1 is strongly expressed in small-diameter sensory neurons of the DRG, TG, and NG and in the hypothalamus, sites where it may exert an important role in thermoreception.

TRPV2 is activated at extremely high temperatures (52 °C), although it is not affected by capsaicin and shows an outwardly rectifying IV curve and a P_{Ca}/P_{Na} ~3. This channel has a Q10 of around 100, and it is thought that the temperatures that activate

TRPV2 are more harmful than those that activate TRPV1. These channels are strongly expressed by myelinated medium-large diameter DRG neurons (Aδ and Aβ), as well as in the hypothalamus and the NG.

TRPV3 channels are activated at warm, close to hot, temperatures (around 34–39 °C, with a Q10 around 6), generating currents with pronounced outward rectification and a P_{Ca}/P_{Na} ~12. They are capsaicin-insensitive channels but are stimulated by camphor, and they are thought to be involved in thermosensation and thermal nociception. Indeed, it has been suggested that TRPV3 channels contribute more to the speed with which mice select a more comfortable temperature than the choice of the temperature value itself. By contrast, TRPV4 channels are more likely to be involved in choosing the preferred temperature from a non-painful range. Interestingly, it was proposed that TRPV3 channels transmit thermal stimuli through skin keratinocytes, which will transmit this information to sensory endings. TRPV3 channels are expressed in sensory DRG and NG neurons, but also in the hypothalamus. Interestingly, they co-localize with TRPV1 in DRG neurons.

TRPV4 are cationic (P_{Ca}/P_{Na} ~6) channels activated at even lower warm temperatures (around 27 °C, with a Q10 of about 10), generating outwardly rectifying currents and responding dynamically to temperature changes in the physiological range. These channels were proposed to play a role in thermosensation and thermoregulation, although some authors could not activate these channels by increasing the temperature. Similarly, some behavioral studies reported a reduced response to temperature changes in TRPV4 knockout (KO) mice, a behavior that was less clear in other studies. Much like TRPV3, the expression of these channels in keratinocytes was proposed to play an important role in transmitting thermal information, which probably contributed to the controversy generated. The sensitivity of this channel to temperature is lost in excised patches, suggesting that it requires a soluble intracellular factor. TRPV4 channels are expressed in DRG, TG, NG, and preoptic/anterior hypothalamic neurons. However, in the hypothalamus, they seem to be expressed in terminals rather than soma, so their role in body thermoregulation is unclear.

TRPM2 (>35 °C), TRPM3 (>40 °C), TRPM4 (>15 °C), and TRPM5 (>15 °C) are channels that can also be activated by warming (**see** Figure **13.9.1**). Yet, they have received less attention, probably because it was initially thought that somatosensory neurons or keratinocytes did not express them. TRPM2 is voltage-insensitive, shows a P_{Ca}/P_{Na} ~1, activates at 35 °C, and has a Q10 of around 15. TRPM3 is expressed broadly, generating an outwardly rectifying current, having a P_{Ca}/P_{Na} between 0.1 and 10, and activating at >40 °C with a Q10 of 7. It is important to say that TRPM3 has been described as part of a triad of TRPs, together with TRPV1 and TRPA1, involved in the transduction of acute noxious heat in mice. The combined ablation of these channels (triple KO) was necessary for the complete reduction of acute noxious sensing; single or double KO combinations resulted in deficits in heat responsiveness, but mice still conserved vigorous withdrawal responses to noxious heat. Heat activation of TRPM2 and TRPM5 was obtained in inside-out patches, suggesting a membrane-delimited mechanism. Interestingly, TRPM2 activation seems to result from the increase in the IV slope, while that of TRPM4 and TRPM5 results from a shift of the activation curve to negative potentials. These last two channels are essentially not permeable to calcium.

13.9.3.2: Cold-Sensitive TRP Channels

Two TRP channels are activated by decreases in temperature (see Figure **13.9.1**). TRPM8 (<25 °C) activates in the cool range while TRPA1 (<18 °C) senses cold-painful temperatures. Similarly, cool fibers (Aδ and C) have activation thresholds at about 30 °C, and cold fibers (C) have activation thresholds <20 °C. Accordingly, two populations of TG neurons were described in terms of their activation threshold when temperatures decreased: 30 °C and 20 °C for low and high thresholds, respectively. In general, cold fibers fire continuously at normal skin temperatures, and they increase their firing frequency when the skin is cooled down, or they shut down when the skin is warmed. In addition, cold fibers can adapt to small decreases in temperature.

TRPM8 channels are voltage-dependent cationic channels permeable to Na^+, K^+, Cs^-, and Ca^{2+} (P_{Ca}/P_{Na} ~3). When expressed in HEK cells and recorded in the whole-cell configuration, they show a voltage-dependent outwardly rectifying current that strongly increases upon cooling from 30 to 15 °C or through the application of menthol.

Figure **13.9.*d*** shows an interactive iCn3D model of TRPM8 ion channel in complex with the menthol analog WS-12 and PI(4,5)P2 (6NR2)

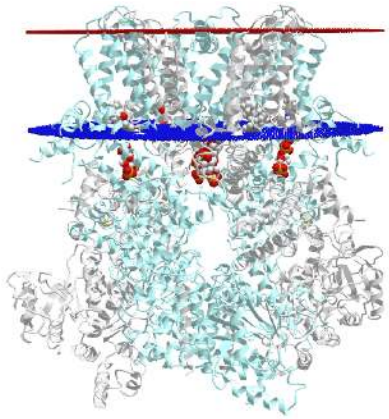


Figure 13.9.d: TRPM8 ion channel in complex with the menthol analog WS-12 and PI(4,5)P2 (6NR2). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...94muQHwYr7eij6

The methanol analog is shown in spacefill and CPK colors. PI(4,5)P2 is shown in spacefill just below the lower cytoplasmic leaflet.

The TRPM8 receptor is a Ca^{2+} cation channel. Cooling compounds like menthol and WS-12 depend on allosteric interactions and membrane phosphatidylinositol 4,5-bisphosphate (PIP_2).

Importantly, basal and cold-stimulated currents reverse around 0 mV and were almost negligible below this potential. Cooling CHO cells expressing TRPM8 (25 to 15 °C) also increases intracellular calcium, and the Q10 in the 25 to 18 °C range is around 24. The effect of temperature is due to an increase in the open probability and a shift in the conductance–voltage relationships along the voltage axis. Similar results were obtained in inside-out macropatch recordings, although the stimulation occurred at lower temperatures, suggesting that the integrity of the cell is important but not indispensable. The role of this channel as a detector of painful cold has been questioned in experiments on KO mice. Still, it is accepted that it is an important cold sensor in vagal, TG, and especially DRG afferents. It was predicted that cold transduction may require the activation and inhibition of several different ion channels (see Figure 13.9.3), such as TRP, TREK, and ENaC channels. If this were the case, TRP channels would probably be more important in the noxious-cold range, whereas TREK channels might participate more strongly in the cool range of temperatures (see Figure 13.9.1). TRPM8 is expressed in small-diameter DRG and TG neurons, presumably thermoreceptors, yet it seems not to co-localize with TRPV1.

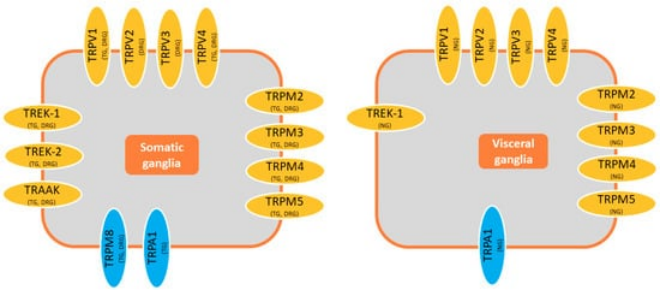


Figure 13.9.3: Most temperature-modulated channels are expressed in both somatic and visceral afferents. The expression of TREK2, TRAAK, and TRPM8 in visceral ganglia appears to be low compared to other channels. However, investigating the colocalization of these channels in the same neuron is necessary to gain a comprehensive understanding of their relative importance in temperature sensing.

TRPA1 is activated by lower temperatures than TRPM8 (<18 °C), and while it would be expected to be involved in cold nociception, this is not that clear. TRPA1 generates an outward rectifying cationic current, both in control conditions and when cold-activated (about 10 °C), with similar permeability for Ca^{2+} and Na^{+} (P_{Ca}/P_{Na} ~1). Cinnamaldehyde can selectively activate currents through this channel in native DRG neurons, as can bradykinin (when co-expressed with BK receptors), strongly suggesting a role in sensing nociceptive stimuli. However, TRPA1 knockout (KO) mice do not seem to have difficulties in sensing cold stimuli through the skin, while the response of TRPM KO mice to cold is significantly dampened. By contrast, about 50% of NG neurons in culture were activated by cooling (<24 °C), mainly through TRPA1 channel activation (increase in $[Ca]_i$, depolarization, and AP firing). Interestingly, about 10% of the NG neurons responded to cold through a TRPA1- and Ca-independent pathway. TRPA1 often co-localizes with TRPV1, and in fact, this could explain the paradoxical hot sensation

experienced with an extremely cold stimulus. Interestingly, most NG neurons are sensitive to cold and heat. TRPA1 is expressed in DRG and TGs, while TRPA1 and TRPM8 are not co-expressed in DRG neurons, but they are in TG neurons. In summary, the data suggest that TRPA1 is the principal ion channel involved in cold sensation in visceral (NG) neurons, while TRPM8 would fulfill the same role in somatic neurons.

13.9.3.3: Molecular Origin of Thermosensitivity

The mechanism by which temperature modulates TRP channels is still unclear, yet several hypotheses have been proposed. Temperature changes could:

- produce a ligand that binds to a receptor and affects the channel;
- produce a structural change in the channel that provokes its opening;
- affect the structure of the membrane, causing changes in tension that would, in turn, affect ion channels.

Because capsaicin induces burning pain, it has been hypothesized that capsaicin and heat may use a common mechanism to activate TRPV1 and produce pain. Both stimuli affect excised patches, and it is generally accepted that TRPV1 is directly activated by noxious heat, so it can be considered a true heat sensor.

The fact that TRPM8 can be activated by cooling in inside-out patches suggests that the mechanism is membrane-delimited, arguing against the participation of a second messenger pathway. Notwithstanding, phospholipase C inhibition strongly dampened the calcium increase provoked by cold stimuli in TRPM8-expressing CHO cells. Cooling activates TRPM8-expressed channels by causing a shift in the voltage dependence of activation to negative values, and the same mechanism is responsible for the activation of TRPV1 by heat, but not the activation of TRPM2. It has been proposed that temperature induces large rearrangements of the protein and thus, a temperature-sensing domain or “temperature sensor” exists in the structure of TRPM8 channels. Much like for TRPM8, inhibition of phospholipase C strongly reduces the increase in calcium provoked by cold stimuli in TRPA1-expressing CHO cells.

13.9.4: Conclusions

There are several important differences between the two main types of thermosensor channels that have been reviewed in this article (see Figure 13.9.2). First, TREKs are potassium channels with negative reversal potentials, such that their activation would result in a reduction in thermoreceptor activity. By contrast, as the reversal potential of TRPs (cationic channels) is close to 0 mV, their activation will increase thermoreceptor excitability. Second, the three TREK channels appear to increase their open probability as temperatures increase. At the same time, there are two possible situations in the case of TRPs: one in which their activity increases as the temperature increases, and another in which activity increases when the temperature decreases (see Figure 13.9.2). Although the activation of these channels generates opposing effects on thermoreceptors (depolarization versus hyperpolarization), the information available to date regarding the participation of TRP and TREK channels in thermosensitivity strongly suggests that both types of channels collaborate and complement each other to generate the sensations of heat, cold, and thermal pain. In support of this hypothesis, TREK, and TRP channels are very often co-expressed in thermoreceptors and other sensory neurons TRP channels are generally accepted as the primary thermosensors; however, several lines of evidence indicate that other channels are necessary to explain the full plethora of mechanisms involved in thermosensation. TREK2 channels appear important in thermoreception at moderate temperatures and sensing innocuous cold but not aversive cold, while TREK1 and TRAAK acting together may be important in sensing painful cold.

13.9.5: Summary

This chapter provides a detailed examination of the molecular and cellular bases for thermosensation and thermal regulation in mammals. It emphasizes how animals maintain constant body temperature via intricate negative feedback systems that involve both peripheral (skin) and internal (visceral and central) thermoreceptors. The discussion integrates the roles of various temperature-sensitive ion channels—particularly two families of channels: the potassium-selective TREK channels and the cationic transient receptor potential (TRP) channels—and explains how their distinct properties contribute to the detection of innocuous and noxious thermal stimuli.

Thermoregulatory Background

- **Homeostasis and Feedback:**
  Animals expend considerable energy to maintain core temperature, relying on a balance between heat production and heat loss. External receptors in the skin and internal receptors in organs (spinal cord, hypothalamus, and visceral tissues) continuously monitor ambient and body temperatures, providing critical inputs for thermoregulatory responses.
- **Temperature-Dependent Biochemistry:**
  Many biochemical reactions and physical processes, such as protein conformational changes and ion channel gating, are highly temperature sensitive. This sensitivity is quantified by the Q_{10} coefficient, with thermosensitive ion channels generally exhibiting Q_{10} values of 2–5 or higher.

TREK Channels as Thermosensors

- **Channel Architecture and Activation:**
  TREK channels, part of the two-pore domain potassium channel (K2P) family, include TREK1, TREK2, and TRAAK. These channels have four transmembrane segments and two pore domains, assembling as dimers. They are normally quiet at room temperature but become strongly activated by stimuli—including increases in temperature—leading to an outward potassium current. This hyperpolarizes thermoreceptors, reducing their excitability.
- **Physiological Role:**
  Experimental data indicate that TREK channels modulate responses to both noxious heat and cold. For instance, TREK1 activity increases significantly from 14 °C to 42 °C, and knockout studies show that loss of TREK channels leads to enhanced thermal pain sensitivity, particularly in the low-threshold range. This suggests that TREK channels help counterbalance excitatory signals from heat-sensitive TRP channels by lowering neuronal excitability during high-temperature exposure, while their inhibition by cold contributes to depolarization and enhanced cold sensitivity.

TRP Channels in Thermosensation

- **Diverse Thermal Responses:**
  Transient receptor potential (TRP) channels are cationic channels with high temperature coefficients (Q_{10} >20) that serve as primary detectors of thermal stimuli. The chapter distinguishes between:
  - **Heat-Sensitive TRPs:**
    TRPV1 (activated >43 °C by noxious heat and capsaicin), TRPV2 (activated at very high temperatures, >52 °C), TRPV3, and TRPV4 (activated by warm temperatures). These channels typically depolarize thermoreceptors by allowing an influx of cations, particularly calcium.
  - **Cold-Sensitive TRPs:**
    TRPM8, which is activated by moderate cooling (<25 °C) and by cooling compounds like menthol, and TRPA1, which responds to painfully cold temperatures (<18 °C). Their activation contributes to the increased excitability of thermoreceptors under cold conditions.
- **Mechanistic Insights:**
  TRP channel activity is influenced by direct temperature-induced conformational changes as well as by shifts in voltage dependence. Their activity, recorded in both whole-cell and patch-clamp configurations, contrasts with TREK channels—resulting in depolarization (excitatory) rather than hyperpolarization.

Integrative Perspectives on Thermosensation

- **Opposing Effects on Neuronal Excitability:**
  TREK channels, with a reversal potential near the potassium equilibrium, tend to hyperpolarize neurons when activated by heat. In contrast, TRP channels, which are non-selective cation channels with reversal potentials near 0 mV, depolarize neurons. The interplay between these channels enables the nervous system to finely tune the sensory responses to a broad range of temperatures—from comfortable warmth to noxious cold or heat.
- **Molecular Origin of Temperature Sensitivity:**
  The chapter discusses hypotheses on how temperature modulates channel activity. For TREK channels, evidence suggests that

the C-terminal domain and its coupling with the C-type gate are critical for heat sensitivity, while for TRP channels, temperature may directly induce large conformational rearrangements in the channel protein itself, possibly in a membrane-delimited manner.

- **Experimental Approaches and Clinical Relevance:**
  Key techniques such as heterologous expression, patch-clamp electrophysiology, and knockout mouse models are used to dissect the roles of TREK and TRP channels in thermosensation. Understanding these mechanisms has implications for pain management, as dysregulation of these channels can lead to thermal hyperalgesia or cold allodynia, highlighting potential targets for analgesic therapies.

In summary, this chapter synthesizes current insights into the molecular mechanisms underlying thermosensation by detailing how TREK and TRP channels operate in concert to regulate neuronal excitability in response to temperature changes. It provides a framework for understanding how these channels contribute to both the perception of thermal stimuli and the modulation of thermal pain, laying the groundwork for future research and therapeutic intervention in temperature-related sensory disorders.

This page titled 13.9: Signal Transduction - Temperature is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

13.10: Signal Transduction - Pressure

Learning Goals (ChatGPT o3-mini)

- **Define Mechanotransduction:**
  - Explain the concept of mechanotransduction and its importance in converting mechanical stimuli into electrochemical signals.
  - Describe the biological processes (e.g., pain sensation, neuronal development, red blood cell volume regulation) that depend on mechanotransduction.
- **Identify Mechanosensitive Ion Channels:**
  - List key families of mechanosensitive ion channels (including TRP, voltage-gated Na^+, K^+, Ca^{2+} channels, TREK channels, and Piezo channels) and understand their general functions in cells.
  - Discuss the unique challenges in studying mechanosensitive channels, particularly those that are multimodal (activated by mechanical forces, chemicals, temperature, and osmolarity).
- **Understand the Discovery and Importance of Piezo Channels:**
  - Summarize the discovery of the Piezo family (Piezo1 and Piezo2) as the long-sought mechanosensitive channels in mammals.
  - Explain the tissue-specific expression patterns of Piezo1 (nonsensory tissues such as lung, bladder, skin) versus Piezo2 (sensory tissues like dorsal root ganglia and Merkel cells).
- **Describe the Structural Features of Piezo Channels:**
  - Outline the overall architecture of Piezo channels, including the three-bladed, propeller-shaped homotrimeric structure.
  - Identify key structural components such as the central cap (CED), anchor region, long intracellular beams, and highly curved peripheral blades.
  - Explain the significance of the unprecedented 38-transmembrane (TM) topology in Piezo1 and how THUs (transmembrane helical units) contribute to the channel structure.
- **Explain the Mechanotransduction Mechanism of Piezo Channels:**
  - Describe the lever-like mechanism by which mechanical forces are converted into channel gating, highlighting the roles of the distal blades, intracellular beam, and pivot residues.
  - Discuss the membrane dome and footprint hypotheses as models to explain how membrane curvature and tension modulate Piezo channel activity.
- **Evaluate the Kinetic and Inactivation Properties of Piezo Channels:**
  - Explain how Piezo channels transition between open, closed, and inactivated states.
  - Analyze factors (e.g., voltage, membrane tension, lipid composition, modulatory proteins) that affect Piezo channel activation and inactivation kinetics.
  - Identify the impact of specific mutations on channel kinetics and relate these to human diseases such as hereditary xerocytosis and lymphatic dysplasia.
- **Understand Pharmacological Modulation:**
  - Summarize the small-molecule activators (e.g., Yoda1, Jedi1/2) and inhibitors (e.g., GsMTx4, Dooku1) of Piezo channels, and their mechanisms of action.
  - Evaluate how pharmacological modulators can alter the mechanical threshold for activation and potential therapeutic applications in mechanosensitivity disorders.
- **Integrate Physiological Functions and Pathophysiology:**
  - Discuss the role of Piezo channels in various tissues (e.g., CNS, blood vessels, erythrocytes, lung, gastrointestinal and urinary tracts, articular cartilage, and sensory ganglia).

- Connect the structural and functional properties of Piezo channels with their contributions to normal physiology and in the context of mechanosensitivity-related diseases.

Achieving these goals will enable students to integrate molecular structure with physiological function, understand the mechanisms underlying mechanotransduction, and appreciate the potential therapeutic implications of modulating Piezo channel activity.

This chapter section is taken in its entirety from Fang, XZ., Zhou, T., Xu, JQ. *et al.* Structure, kinetic properties, and biological function of mechanosensitive Piezo channels. *Cell Biosci* **11,** 13 (2021). https://doi.org/10.1186/s13578-020-00522-z Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/. We added iCn3D molecular models.

13.10.1: Introduction

Mechanotransduction, converting mechanical stimuli into electrochemical signals, is essential for various biological processes, including neuronal cell development, pain sensation, and red blood cell volume regulation. As pivotal mechanosensors in mechanotransduction, mechanosensitive (MS) ion channels have been found in organisms from bacteria to mammals. Extensive studies have revealed a variety of ion channels in eukaryotic cells that can sense various forms of mechanical forces (Table **13.10.1**). These ion channels include transient receptor potential (TRP) channels and voltage-gated Na^+, $K^{+,}$ and Ca^{2+} channels, whose dysfunction may be associated with human genetic diseases. Notably, the MS candidates identified in invertebrates either have no homologs (e.g., TRPN) or no functional conservation (e.g., DEG/ENaC/ASIC) in mammals. Furthermore, most MS candidates (the TRP channel in particular) are activated not only by mechanical stimuli but also by chemicals, temperature, osmolarity, and heat (> 27–34 °C). Defining the molecular details of MS cation channels in mammals is paramount to understanding the mechanotransduction process and finding potentially novel therapeutic strategies for mechanosensitivity disorders.

Channel family	Channel isoforms	Ref.
TRP channels	TRPA1	[6]
	TRPC1	[7]
	TRPC6	[8]
	TRPV1	[9]
	TRPV4	[10]
	TRPM4	[11]
	TRPM7	[12]
	TRPN	[13]
	TRPP2	[14]
K^+ channels	Shaker (Kv1.1)	[15]
	Ca^{2+}-activated K^+ (BK)	[16]
	TREK1/2	[17]
	TRAAK	[18]
	HCN2	[19]
Na^+ channels	Nav1.5	[20]
Ca^{2+} channels	L-type	[21]
	N-type	[22]
	T-type	[23]
Cl^- channels	CFTR	[24]
OSCA protein family	*Sc*CSC1, *Hs*CSC1	[25]
DEG/ENaC superfamily	C.elegans MEC (MEC-4, MEC-10)	[26]
	ASIC	[27]
Other channels	TMC1/2	[28]

In 2010, Coste et al. revealed a novel family of mechanically activated (MA) cation channels in eukaryotes consisting of Piezo1 and Piezo2 channels, which have been proposed as the long-sought-after MS ion channels in mammals. The Piezo1 channel is present in nonsensory tissues, with particularly high expression in the lung, bladder, and skin; by contrast, the Piezo2 channel is predominantly present in sensory tissues, such as dorsal root ganglia (DRG) sensory neurons and Merkel cells. Since their discovery, tremendous effort has been made to reveal the structures and biological functions of Piezo 1 and 2. The partial molecular structure of a Piezo channel was determined by cryo-electron microscopy (cryo-EM). Furthermore, Piezo channels have been linked to various pathological and physiological processes, including erythrocyte volume regulation, cell division, and innate immunity. Moreover, Piezo channel mutations are associated with multiple hereditary human diseases, such as autosomal recessive congenital lymphatic dysplasia, hereditary xerocytosis (a rare, autosomal dominant congenital hemolytic anemia characterized by macrocytic stomatocytosis and decreased red cell osmotic fragility due to a defect in cation permeability), and an autosomal recessive syndrome of muscular atrophy with perinatal respiratory distress. Considerable progress has been made toward characterizing Piezo proteins' structural features, physiological significance, and biophysical properties. Given the importance of Piezo channels in understanding mechanotransduction processes, this review focuses on their structural details, kinetic properties,

and potential functions as mechanosensors. We also briefly review the hereditary diseases caused by mutations in the Piezo genes, which are key to understanding their functions.

13.10.2: Structure of Piezo channels

Piezo proteins have an uncommonly large predicted size of approximately 2500 amino acids and encompass numerous transmembrane (TM) regions. Subsequent research has revealed that the mouse Piezo1 (mPiezo1) channel is an evolutionarily conserved pore-forming ion channel directly gated by membrane stretch. Several published cryo-EM studies have revealed that mPiezo1 exhibits a three-bladed, propeller-shaped homotrimeric structure that includes a central cap, three peripheral blade-like structures on the extracellular side, three long beams on the intracellular side that bridge the blades to the cap, and a TM region between these features, as shown in Figure 13.10.1.

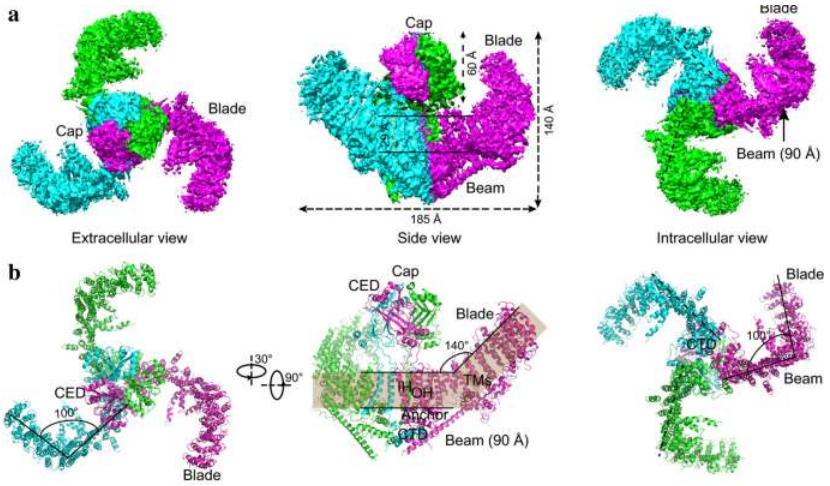


Figure 13.10.1: Cryo-EM structure of the mPiezo1 channel (adapted from Zhao et al.). **a.** Multiple views of the sharpened map of the trimeric channel with the major domains labeled, with the three subunits colored red, green, and blue. **b.** A cartoon model in which the three subunits are red, green, and blue. In the middle panel, the front subunit has been omitted to provide a better view of the curvature of the TMs

13.10.3: **Structure of the Piezo1 channel**

13.10.3.1: Unprecedented 38-TM topology

Piezo channels are predicted to possess an unusually large number of TM regions, ranging from 10 to 40. Zhao et al. recently produced high-resolution structures of mouse Piezo1 (mPiezo1), revealing a unique 38-TM topology in each subunit, as shown in Figure 13.10.2. The two TM regions (TM37 and TM38) closest to the center of the protein are designated as the inner helix (IH) and outer helix (OH), respectively, and enclose the transmembrane pore of the central pore module. The other 36 TM regions (TM1-36) form a curved blade-like structure with nine repetitive folds containing 4 TM regions each, named **transmembrane helical units (THUs).**

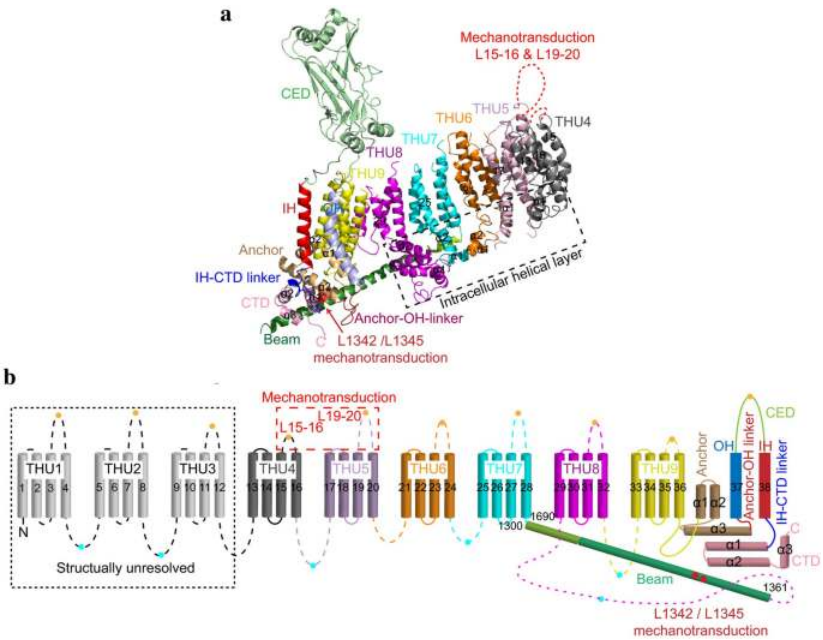


Figure **13.10.2**: A 38-TM topology model and key functional sites in mPiezo1(adapted from Zhao et al.). **a.** A model showing one subunit with individual THUs and featured structural components. Red spheres indicate residues L1342 and L1345 in the beam. **b.** A 38-TM topology model color-coded to match the cartoon model in A

Figure **13.10.*a*** shows an interactive iCn3D model of the mouse mechanosensitive Piezo1 channel (5Z10) (long load time)

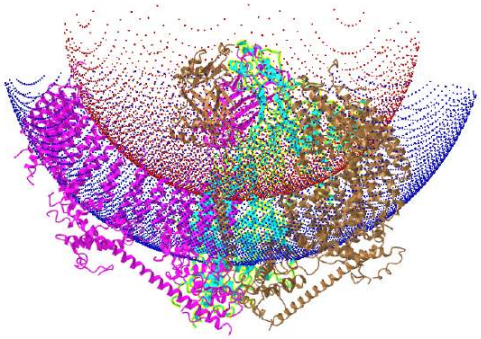


Figure **13.10.*a***: Mouse mechanosensitive Piezo1 channel (5Z10). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ZcVpJczxvaaG79 (long load time)

13.10.3.1: Central cap

Kamajaya and colleagues employed topological prediction modeling. They found that residues 2210 to 2457 in Piezo1 form an extracellular loop following the last TM region from the C-terminus, defined as the C-terminal extracellular domain (CED) **(**Figure **13.10.1)**. The deletion of residues 2218 to 2453 from the Piezo1 protein abolished expression of the central cap, suggesting that this region trimerizes to form the central cap (Figure **13.10.1** and Figure **13.10.3**). Further analysis revealed that the central cap consists of the CED in the form of a trimeric complex that encloses an extracellular vestibule (EV) with openings (Figure **13.10.3***)*.

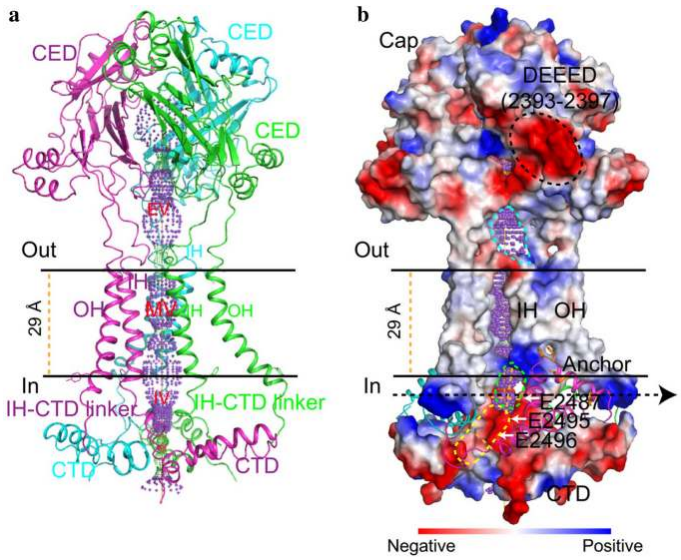


Figure 13.10.3. Structure of the central pore module (adapted from Zhao et al. [35]). **a.** Ribbon diagram of the ion-conduction pore comprising the OH, CED, IH, and CTD from three color-coded subunits. The central solvent-accessible pathway is marked with a dotted mesh generated by the program HOLE. **b.** Pore module presenting the surface electrostatic potentials showing negative (red) and positive (blue) potential. Extracellular and intracellular fenestrations are marked by cyan and green dashed lines, respectively. Yellow dashed lines mark the lateral portal

13.10.3.1: Anchor

A hairpin structure, referred to as the anchor, connects the OH-IH pair to the C-terminal domain (CTD) plane, which moves the OH-CED-IH-containing region of one subunit into the neighboring subunit in a clockwise direction (Figure 13.10.1 and Figure 13.10.3). The anchor comprises three helices (α1, α2, and α3). Helices α1 and α2 were found to organize into an inverted V-shaped structure, which maintains the integrity of the ion-conducting pore (Figure 13.10.2*b*). In parallel with the membrane plane, the long α3 helix links to the OH via a lysine-rich anchor-OH linker that interacts with the polar residue-rich α2–3 turn in the anchor and the glutamate-rich region of the CTD. A few mutations in Piezo1 at locations including KKKK (2182-K2185), T2143, T2142 (T2127 in human Piezo1), R2514, E2523, and E2522, which are located in α3 in the anchor, have been reported to cause severe disease. Additionally, SERCA2, a Piezo-interacting protein, suppresses Piezo1 by acting on the anchor-OH linker. These findings support the structural and functional importance of the anchor region.

13.10.3.1: The long intracellular beam

On the intracellular surface, Piezo1 contains three beam-like structures 90 nm in length that are organized at a 30° angle relative to the membrane plane (Figure 13.10.1 and Figure 13.10.2). Residues H1300-S1362 form the beam structure. The large intracellular THU7-8 loop, which contains approximately 390 residues, might provide the beam with the structural basis for mechanical transmission. Functionally, the three long intracellular beams support the whole TM skeleton and physically bridge the distal THUs to the central ion-conducting pore. When residues 1280 to 1360 (which form this beam structure) were deleted, the resulting mutant protein was absent, suggesting the structural importance of the beam.

13.10.3.1: Highly curved blades

The nine peripheral THUs in each subunit form blade-like structures, with each blade twisted clockwise (Figure 13.10.1*b*). The proximal TM25–TM36 and peripheral TM13-24 interact at a 100° angle, as viewed from 90º relative to the plasma membrane plane, and a 140° angle from a line parallel to the plasma membrane plane. Another important feature of the blades is the L-shaped helical structures formed by TM13, TM17, TM21, TM25, and TM29. Both identifiable structural features appear to be ideal not only for mechanosensation but also for the induction of local membrane curvature. Intriguingly, the peripheral TM13-24 appears to be within a highly curved membrane plane, indicating that the Piezo1 channel can curve the membrane in which it resides. This is consistent with past studies implying that Piezo1 ion channels can be regulated by cellular membrane curvature and tension.

13.10.3.1: The ion-conducting pathway

As pore-forming ion channels, Piezo proteins contain a trimeric ion-conducting channel made up of residues 2,189 to 2,547, which contain the last two TMs (Figure **13.10.3**). The continuous central channel consists of three parts: an EV within the cap region, a transmembrane vestibule (MV) within the membrane, and an intracellular vestibule (IV) underneath the membrane. The EV and IV have an opening that connects to the MVs, which are positioned above and below the membrane. Importantly, DEEED (2393–2397), a patch of negatively charged residues residing in the opening of the extracellular "cap" structure consisting of the CED, is required to ensure efficient ion conduction and determine the selection of cations over anions. Additionally, two critical acidic residues, E2495 and E2496, located at the CTD-constituted IV, may be responsible for divalent calcium ion selectivity, unitary conductance, and pore blockage.

13.10.4: **Structure of the Piezo2 channel**

Like Piezo1 channels, Piezo2 channels are large membrane proteins with over 2,800 residues. However, the Piezo2 channel and Piezo1 channel share approximately only 42% sequence homology. Recent studies have shown that the overall structure of the Piezo2 channel is very similar to that of Piezo1 in that it forms a three-bladed, propeller-like homotrimeric structure comprising a central ion-conducting pore module and three peripheral blades with 38 TMs.

Figure **13.10.*b*** shows an interactive iCn3D model of the mammalian tactile channel PIEZO2 (6KG7) (long load time)

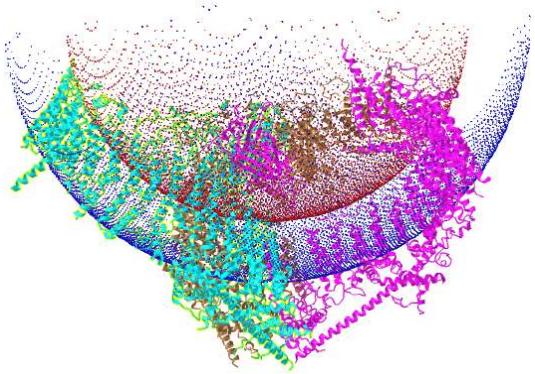


Figure **13.10.*b***: Mammalian tactile channel PIEZO2 (6KG7). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...4FaL71AL23KiA9 (long load time)

In the Piezo2 channel, the charged residues at the interface between the beam and the CTD are required to ensure the normal mechanosensitivity of the channel. Moreover, single-channel recordings indicated that a previously unrecognized intrinsically disordered domain adjacent to the beam acts as a cytosolic plug that limits ion permeation, possibly by clogging the inner vestibule in both Piezo1 and Piezo2. Furthermore, by structurally comparing the Piezo1 and Piezo2 channels, Wang et al. found that the Piezo2 channel has additional constriction sites at L2743, F2754, and E2757 that might serve as a transmembrane gate controlled by the cap domain.

13.10.5: Lever-like mechanotransduction mechanism

Based on the unique topological features of the mPiezo1 channel, a lever-like mechanotransduction mechanism to explain its extraordinary mechanosensitivity was proposed (Figure **13.10.4**). In the mPiezo1 channel, the curved blades composed of THUs can act as a mechanosensor, while the beam structure, with the residues Ll1342 and Ll1345 acting as a pivot, can act as a lever-like apparatus. Coupling the distal blades and the central pore through the lever-like apparatus converts mechanical force into a force used for cation conduction.

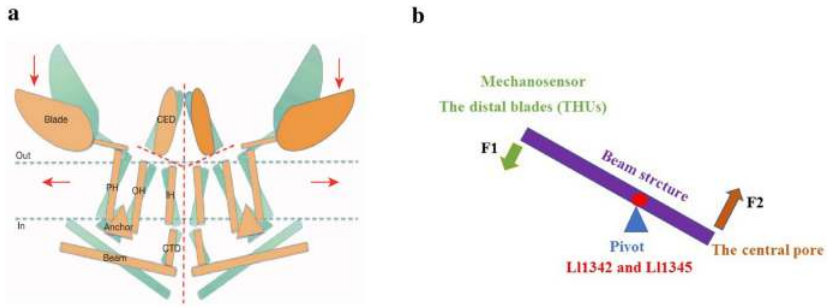


Figure 13.10.4: Model of the lever-like mechanotransduction model. The curved blades can act as a mechanosensor, while the beam structure, with residues Ll1342 and Ll1345 acting as a pivot, can act as a lever-like apparatus. Coupling of the distal blades and the central pore through the lever-like apparatus converts mechanical force into cation conduction. **a** Proposed model of the force-induced gating of Piezo channels. The blue and orange models represent the channel in its closed and open states, respectively. Red dashed lines indicate possible ion-conduction pathways. Adapted from Ge et al. **b** A lever-like mechano-gating model in Piezo1. The blue and red dashed arrows indicate input and output forces, respectively.

Because the lever's pivot is positioned closer to the central pore than to the distal blades, the input force is effectively amplified through the lever-like apparatus. Additionally, a large conformational change in the distal blades is converted into a relatively slight opening of the central pore, allowing cation-selective permeation.

13.10.6: Kinetics properties of Piezo channels

13.10.6.1: Activation mechanisms of Piezo channels

Normal Piezo channel kinetics can be modeled with three states: open, closed, and inactivated; these states have emerged, collectively, as an important mechanism in the Piezo channel function. Studies have proposed that the Piezo1 channel is gated directly by bilayer tension that can be modified by cytoskeletal proteins and linkages to the extracellular matrix (ECM). For example, in overhydrated red blood cells (RBCs), Piezo1-mediated Ca^{2+} influx activates K^+ efflux through the Gardos channel (KCa3.1), leading to water loss and RBC dehydration.

Piezo1 and Piezo2 channels exhibit a three-bladed, propeller-shaped trimeric architecture and can locally deform lipid membranes into a dome-like shape. In addition, changes in the projection area of Piezo channels from closed to open are essential for their mechanosensitivity; this was investigated by calculating the available free energy. Based on these findings, the membrane dome mechanism was proposed and experimentally proved to explain the activation mechanisms of Piezo channels (Figure 13.10.5). Essentially, the dome shape created by Piezo channels in their closed conformation acts as a potential energy source for MS gating. Under tension, lateral membrane tension flattens the Piezo dome, which increases the energy of the membrane-channel system in proportion to the expansion of the projected dome area. Piezo channels then open due to the relative energy difference. This mechanism can account for Piezo channels' highly sensitive mechanical gating with a cation-selective pore. Although the membrane dome mechanism explains the exquisite mechanosensitivity of Piezo channels, it does not consider the shape of the surrounding membrane. Haselwandter et al. proposed the membrane footprint hypothesis, which states that the Piezo1 channel deforms the shape of the membrane outside the perimeter of the channel such that it exhibits a curved membrane footprint, which amplifies the sensitivity of Piezo1 to changes in the membrane tension. Nevertheless, further experiments are needed to test and refine these hypotheses.

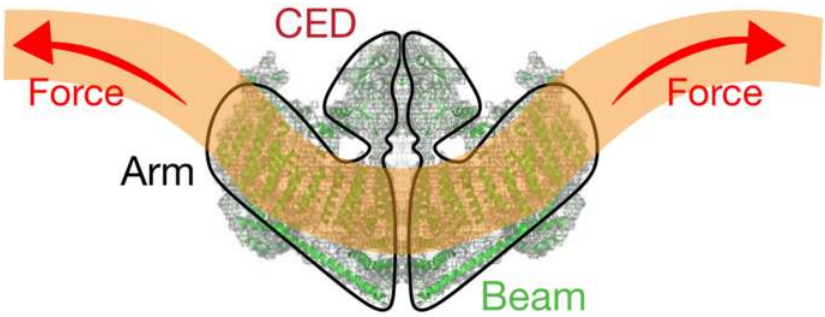


Figure 13.10.5: Model of the membrane doming mechanisms. Changes in membrane curvature lead to a gating force applied to the Piezo1 channel

13.10.6.1: Inactivation kinetics of Piezo channels

Various types of mechanical stimulation trigger Piezo channel activation and sequentially elicit an MA current with rapid decay, even in the presence of continued stimulation, due to rapid inactivation. Coste et al. first described detailed information about the voltage-dependent inactivation kinetics of Piezo channels, characterized as fast at rather negative membrane potentials and slow at rather positive membrane potentials. Additionally, Piezo1 channel inactivation is relatively slow compared with Piezo2 channel inactivation. Several point mutations in Piezo channels have been reported to slow the inactivation process, producing larger cation fluxes and resulting in various human diseases. Given its demonstrated key role in normal channel function, we will review what is known about the inactivation kinetics of Piezo channels next, focusing on the inactivation mechanism.

The available information regarding the structures (residues/domains) and human disease-related point mutations has helped clarify ion channel inactivation mechanisms. Currently, six gain-of-function mutations associated with dehydrated hereditary xerocytosis (DHS) have been found to slow the inactivation rate of Piezo channels (Table 13.10.1), most of which are clustered at the central core region of the Piezo channel structure. This implies that the pore region, which contains an OH, an IH, an extracellular cap domain, and an intracellular CTD, determines the kinetics of inactivation. Further detailed links between structural domains and inactivation kinetics have been investigated. Wu et al. identified that the distinct inactivation kinetics of Piezo1 and Piezo2 channels and characteristic voltage-dependent inactivation appear to be determined by the C-terminal extracellular domains (cap domain). Two potential inactivation gates within the IH and CTD have been considered sufficient for the normal inactivation of the Piezo1 and Piezo2 channels. Recently, three small subdomains within the extracellular cap were shown to confer Piezo channel inactivation individually. These results support the idea that the ion-conducting pore region of Piezo channels is essential for their inactivation properties.

Table 13.10.2: **Mutations in Piezo1 and Piezo2 Associated with Human Diseases:** Full-size table

Interestingly, a slowly inactivating MS current in mouse embryonic stem cells (mESs) that is dependent on the Piezo1 channel has been described. However, heterologous expression of Piezo1 cDNA from mES cells displays fast inactivation kinetics, indicating that additional regulatory mechanisms other than the amino acid sequence determine the slow kinetics of the Piezo1 channel in mES cells. Recently, sphingomyelinase activity has been revealed to be a crucial determinant of Piezo1 inactivation. Various modulators, such as pH, temperature, divalent ion concentrations, alternative splicing, osmotic swelling, membrane lipid composition, co-expression of other membrane proteins, and G-protein-coupled pathways, have also been reported to regulate the Piezo channel kinetics; however, we still know very little about the relationships among these factors and pivotal structural domains.

13.10.6.1: Pharmacological modulators of Piezo channels

Despite the relatively recent discovery of Piezos, progress has been made regarding small-molecule modulators of Piezo1. Piezo1 chemical activators, including Yoda1 and Jedi1/2, were able to open Piezo1 ion channels in the absence of mechanical stimulation. Jedi1/2, a novel hydrophilic Piezo1 chemical activator, acts through the peripheral blades and utilizes a peripheral lever-like apparatus consisting of the blades and a beam to gate the central ion-conducting pore whereas Yoda1 acts as a molecular wedge, facilitating force-induced conformational changes, effectively lowering the channel's mechanical threshold for activation However, the reason why Yoda1 does not efficiently activate the Piezo2 channel is unclear. Specific inhibitors of Piezo1 are in their infancy. As nonspecific inhibitors of the ion pore in stretch-activated ion channels, gadolinium and ruthenium red have also been shown to

block mouse Piezo1 channels with IC50 values of approximately 5 mM. The commonly used toxin inhibitor of mechanosensitive channels, GsMTx4, also inhibited the Piezo1 channel. Still, it might not bind Piezo1 directly, but rather act via modulating local membrane tension near the channel. Dooku1, an analog of Yoda1 without a stimulatory effect, antagonizes Yoda1-evoked activation of Piezo1 and aortic relaxation.

13.10.7: The function of Piezo channels

Piezo channels are expressed in a wide range of mechanically sensitive cells and allow Ca^{2+} influx in response to different external forces, such as fluid flow, pulling, and ultrasonic forces. The biological function of Piezo channels was recently investigated in many studies (Figure **13.10.6**). These studies' results verified Piezo channels' pivotal roles in mechanotransduction under physiological and pathophysiological conditions. Here, we focus on reviewing the biological function of Piezo channels in several different types of MS tissues and cells.

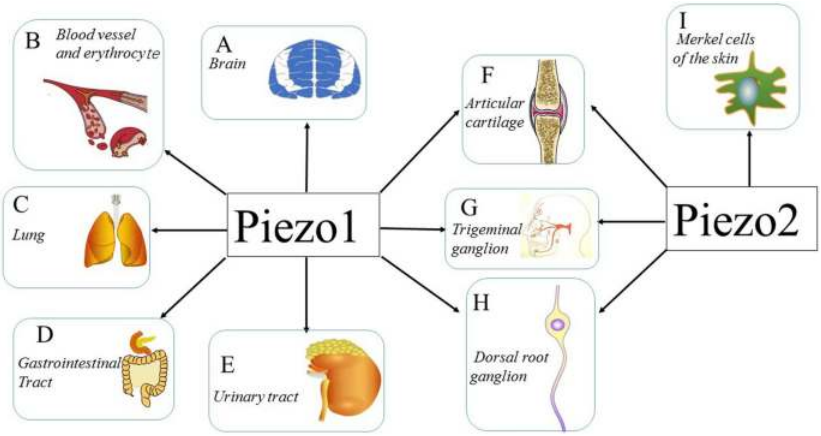


Figure **13.10.6**: Expression and function of Piezo channels. Multiple tissues and cells express Piezo channels, and each of those shown is discussed in this review. Panels **a–e** demonstrate the vital role of the Piezo1 channel in the CNS, blood vessels, erythrocytes, lungs, gastrointestinal tract, and urinary tract. Panels **f–h** illustrate the expression of both the Piezo1 channel and Piezo2 channel in articular cartilage, trigeminal ganglia, and dorsal root ganglia. Panel **i** shows that the Piezo2 channel is expressed in Merkel cells, which are involved in sensing light touch.

Consult the original article for details on the roles of Piezo channels in the systems above.

13.10.8: Summary

This chapter delves into the intricate process of mechanotransduction—the conversion of mechanical stimuli into electrochemical signals—a fundamental mechanism underlying various physiological processes such as pain sensation, neuronal development, and red blood cell volume regulation. The discussion spans the diversity of mechanosensitive ion channels found from bacteria to mammals and emphasizes the importance of understanding their molecular basis for both basic biology and therapeutic innovation.

The chapter begins by outlining the essential role of mechanotransduction in maintaining cellular and organismal homeostasis, highlighting how mechanical forces are detected by specialized receptor cells located in the skin, internal organs, and other tissues. It discusses the impact of temperature on ion channel gating and the specificity of channels with high temperature coefficients (Q_{10} ≥2–5) as critical components of sensory systems.

A major focus is given to the Piezo family of ion channels, discovered as the long-sought mechanosensors in mammals. Piezo1 and Piezo2 channels have been identified with distinct tissue distributions—Piezo1 in nonsensory tissues such as the lung, bladder, and skin, and Piezo2 predominantly in sensory tissues like dorsal root ganglia and Merkel cells. Recent advances using cryo-electron microscopy have revealed that Piezo channels form a three-bladed, propeller-shaped homotrimeric structure with an unprecedented 38-transmembrane topology in each subunit. Key structural features include:

- **Central Cap (CED):** Formed by the C-terminal extracellular domain, it trimerizes to enclose an extracellular vestibule essential for ion conduction.

- **Anchor:** A hairpin structure that connects the inner and outer helices to the cytoplasmic domains, playing a crucial role in stabilizing the ion-conducting pore.
- **Intracellular Beams:** Long, lever-like structures that bridge the peripheral blade domains to the central pore, serving as mechanical levers to amplify input forces.
- **Highly Curved Blades:** Peripheral structures that not only serve as mechanosensors by responding to membrane tension and curvature but also contribute to local membrane deformation.

The chapter explains how the unique architecture of Piezo channels allows them to directly respond to mechanical forces such as membrane stretch and tension. A lever-like mechanism is proposed, whereby the curved blades act as the sensor, and the beams function as a pivot, translating mechanical deformation into the opening of the central ion pore, thereby permitting cation (especially Ca^{2+}) influx.

Furthermore, the kinetic properties of Piezo channels are examined. Piezo channels exhibit distinct states—open, closed, and inactivated—with inactivation kinetics that are modulated by factors such as voltage, membrane lipid composition, and interactions with regulatory proteins. Mutations in key regions of the channel have been linked to human diseases like hereditary xerocytosis and lymphatic dysplasia, underscoring the clinical importance of these channels.

Finally, the chapter touches on pharmacological modulators of Piezo channels. Small-molecule activators like Yoda1 and Jedi1/2 can lower the mechanical threshold for channel activation, while inhibitors such as GsMTx4 and Dooku1 provide insights into channel function and potential therapeutic avenues.

In summary, this chapter integrates structural, kinetic, and pharmacological perspectives to provide a comprehensive understanding of mechanotransduction via Piezo channels. It highlights how these channels' unique molecular features enable them to function as sensitive mechanosensors, crucial for normal physiology and as potential targets for treating mechanosensitivity disorders.

This page titled 13.10: Signal Transduction - Pressure is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski and Patricia Flatt.

- **11.2: Diffusion Across a Membrane - Passive and Facilitated Diffusion** by Henry Jakubowski and Patricia Flatt has no license indicated.

Index

A

acid
- 2.1: Water

adhesins
- 9.6: The Sugar Code and Lectin Decoding
- 9.7: Chapter 7 Problems
- 9.8: Chapter 7 Problems - Answer Key

adhesion
- 2.1: Water

affinity chromatography
- 5.8: Affinity Chromatography

aldoses
- 9.3: Disaccharides

alpha helix
- 4.3: Secondary Structure

amino acid
- 1.7: Types of Biological Molecules
- 3.1: Amino Acids and Peptides

amyloid
- 4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules

anabolic
- 6.3: Endergonic and Exergonic Reactions

anion
- 1.6: The Building Blocks of Molecules

anomeric forms
- 9.3: Disaccharides

atomic number
- 1.6: The Building Blocks of Molecules

ATP energy
- 6.6: Energy in Metabolism

B

base
- 2.1: Water

beta hairpins
- 4.3: Secondary Structure

beta strands
- 4.3: Secondary Structure
- 4.7.2: Secondary Structure and Loops

binding
- 7.5: Binding - The First Step Towards Protein Function

biochemical pathways
- 6.6: Energy in Metabolism

biochemistry
- 9.3: Disaccharides

buffer
- 2.1: Water

C

carbohydrate
- 1.7: Types of Biological Molecules

carbohydrate derivatives
- 9.3: Disaccharides

carbohydrates
- 9.3: Disaccharides

Cardiolipins
- 11.3: Structure and Function - Lipids and Membranes

catabolic
- 6.3: Endergonic and Exergonic Reactions

cation
- 1.6: The Building Blocks of Molecules

Cell Disruption
- 5.2: Cell Disruption

cellulose
- 1.7: Types of Biological Molecules

Centrifugation
- 5.1: Protein Purification

chemical bond
- 1.6: The Building Blocks of Molecules

chitin
- 1.7: Types of Biological Molecules

Chymotrypsin
- 8.2: Mechanisms of Catalysis

cofactor
- 8.1: Regulation of Enzyme Activity

cohesion
- 2.1: Water

collagen
- 4.6: Protein function, domains and cooperativity

competitive inhibition
- 7.2.4: Enzyme Inhibition
- 8.1: Regulation of Enzyme Activity

covalent bond
- 1.6: The Building Blocks of Molecules

cyclic structures
- 9.3: Disaccharides

Cysteine Redox Chemistry
- 3.3: Cysteine Chemistry

D

denaturation
- 1.7: Types of Biological Molecules

Deoxyribonucleic acid (DNA)
- 1.7: Types of Biological Molecules

diastereomers
- 9.3: Disaccharides

disaccharide
- 1.7: Types of Biological Molecules

disaccharides
- 9.2: Monosaccharides
- 9.3: Disaccharides

DNA (Z Form)
- 10.1: Structure and Function - Nucleic Acids

double reciprocal plot
- 7.1: Enzyme Kinetics

E

electron
- 1.6: The Building Blocks of Molecules

electron transfer
- 1.6: The Building Blocks of Molecules

electrophoresis
- 5.4: Electrophoresis

electroporation
- 12.9: Diffusion Across a Membrane - Pores

element
- 1.6: The Building Blocks of Molecules

enantiomers
- 9.3: Disaccharides

endergonic
- 6.3: Endergonic and Exergonic Reactions

energy
- 6: Enzyme Thermodynamics
- 6.6: Energy in Metabolism

entropy
- 6.5: Thermodynamics

enzyme
- 1.7: Types of Biological Molecules

enzyme classification
- 6.1: Enzymes

enzymes
- 6.1: Enzymes

epimers
- 9.3: Disaccharides

epitope
- 9.6: The Sugar Code and Lectin Decoding
- 9.7: Chapter 7 Problems
- 9.8: Chapter 7 Problems - Answer Key

evaporation
- 2.1: Water

exergonic
- 6.3: Endergonic and Exergonic Reactions

F

facilitated diffusion
- 12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion

fat
- 1.7: Types of Biological Molecules

fatty acids
- 12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics

fibrillar proteins
- 4.7.7: Fibrillar Proteins

first law of thermodynamics
- 6.4: The Laws of Thermodynamics
- 6.5: Thermodynamics

Fischer projections
- 9.3: Disaccharides

Fractionation
- 5.3: Fractionation

G

gel electrophoresis
- 5.5: Electrophoresis

Gel Exclusion Chromatography
- 5.6: Gel Exclusion Chromatography

gel filtration chromatography
- 5.6: Gel Exclusion Chromatography

Gibbs energy
- 6.3: Endergonic and Exergonic Reactions

Gibbs Free Energy
- 6.5: Thermodynamics

glycan binding proteins
- 9.6: The Sugar Code and Lectin Decoding
- 9.7: Chapter 7 Problems
- 9.8: Chapter 7 Problems - Answer Key

glycan biochemistry
- 9.3: Disaccharides

glycogen
- 1.7: Types of Biological Molecules

glycolipids
9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls
glycosidic bonds
9.3: Disaccharides

H

Haworth projections
9.3: Disaccharides
helical wheel
4.7.2: Secondary Structure and Loops
Hill coefficient
8.4: Oxygen-Binding Proteins and Allosterism
Hill model
8.4: Oxygen-Binding Proteins and Allosterism
Histidine Tagging
5.10: Histidine Tagging
hormone
1.7: Types of Biological Molecules
hydrogen bond
1.6: The Building Blocks of Molecules
hydrophilic
2.1: Water
hydrophobic
2.1: Water
hydroxynitrile lyase
4.1: The Structure of Proteins- An Overview

I

induced fit model
1.1: Cellular Foundations of Biochemistry
initial velocity method
7.2.2: Kinetics without Enzymes
integrated rate method
7.2.2: Kinetics without Enzymes
ion
1.6: The Building Blocks of Molecules
ion exchange chromatography
5.7: Ion Exchange Chromatography
ionic bond
1.6: The Building Blocks of Molecules
Ionophores
12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion
isoelectric focusing
5.4: Electrophoresis
5.5: Electrophoresis
isotope
1.6: The Building Blocks of Molecules
isozymes
7.2.6: Enzymes and Protein Regulation

K

keratins
4.6: Protein function, domains and cooperativity
ketoses
9.3: Disaccharides
KNF model
8.1: Regulation of Enzyme Activity

L

lectin decoding
9.6: The Sugar Code and Lectin Decoding
9.7: Chapter 7 Problems
9.8: Chapter 7 Problems - Answer Key
lectins
9.6: The Sugar Code and Lectin Decoding
9.7: Chapter 7 Problems
9.8: Chapter 7 Problems - Answer Key
12.6: Membranes and Membrane Proteins
Levinthal paradox
4.4: Protein folding
4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics
4.7.9: Protein Stability - Thermodynamics
4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules
lipids
1.7: Types of Biological Molecules
12.3: Working with Lipids
litmus paper
2.1: Water
lock and key model
1.1: Cellular Foundations of Biochemistry
lysis
5.2: Cell Disruption

M

macromolecule
1.7: Types of Biological Molecules
MAP kinases
13.5: Receptor Tyrosine Kinases (RTKs)
mass number
1.6: The Building Blocks of Molecules
matter
1.6: The Building Blocks of Molecules
membrane
12.6: Membranes and Membrane Proteins
membrane proteins
12.6: Membranes and Membrane Proteins
metabolism
6.6: Energy in Metabolism
Micelle
4.7.3: Tertiary, Quaternary, and Symmetrical Structures
mixed inhibition
8.1: Regulation of Enzyme Activity
Monosaccharide
1.7: Types of Biological Molecules
monosaccharides
9.2: Monosaccharides
9.3: Disaccharides
Morpheein model
8.3: Control of Enzymatic Activity
mutarotation
9.3: Disaccharides
MWC model
8.1: Regulation of Enzyme Activity

N

ncRNA
10.3: Nucleic Acids - RNA Structure and Function
neutron
1.6: The Building Blocks of Molecules
noncompetitive inhibition
8.1: Regulation of Enzyme Activity
nonpolar covalent bond
1.6: The Building Blocks of Molecules
nonreducing sugars
9.3: Disaccharides
Nuclear Hormone Receptors
13.3: Nuclear Hormone Receptors
nucleic acid
1.7: Types of Biological Molecules
nucleic acids
10.1: Structure and Function - Nucleic Acids
nucleotide
1.7: Types of Biological Molecules
nucleus
1.6: The Building Blocks of Molecules

O

octet rule
1.6: The Building Blocks of Molecules
oil
1.7: Types of Biological Molecules

P

passive diffusion
12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion
periodic table of elements
1.6: The Building Blocks of Molecules
pH scale
2.1: Water
phospholipid
1.7: Types of Biological Molecules
photosynthesis
6.6: Energy in Metabolism
pi helix
4.3: Secondary Structure
polar covalent bond
1.6: The Building Blocks of Molecules
Polyacrylamide gel electrophoresis
5.4: Electrophoresis
polypeptide
1.7: Types of Biological Molecules
polysaccharide
1.7: Types of Biological Molecules
polysaccharides
9.4: Polysaccharides
pores
12.9: Diffusion Across a Membrane - Pores
prions
4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules
protein
1.7: Types of Biological Molecules
proteoglycans
9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls
proton
1.6: The Building Blocks of Molecules

Q

Quaternary Structures

4.7.3: Tertiary, Quaternary, and Symmetrical Structures
4.7.4: Secondary Structural Motifs and Domains
4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure
4.7.6: Intrinsically Disordered Proteins
4.7.7: Fibrillar Proteins

R

radioactive isotope

1.6: The Building Blocks of Molecules

Ramachandran plot

4.7.1: Main Chain Conformations
4.7.2: Secondary Structure and Loops

Receptor Tyrosine Kinases (RTKs)

13.5: Receptor Tyrosine Kinases (RTKs)

reducing sugars

9.2: Monosaccharides
9.3: Disaccharides

Ribonucleic Acid (RNA)

1.7: Types of Biological Molecules

S

Saturated fatty acid

1.7: Types of Biological Molecules

Second Law of Thermodynamics

6.4: The Laws of Thermodynamics
6.5: Thermodynamics

secondary structure

4.6: Protein function, domains and cooperativity

secondary structure

4.3: Secondary Structure
4.7.2: Secondary Structure and Loops

serine proteases

8.2: Mechanisms of Catalysis

solvent

2.1: Water

Sonication

5.2: Cell Disruption

spontaneous

6.3: Endergonic and Exergonic Reactions

starch

1.7: Types of Biological Molecules

stereochemistry

9.3: Disaccharides

steroid

1.7: Types of Biological Molecules

subtilisin

8.2: Mechanisms of Catalysis

surface tension

2.1: Water

T

tauopathies

4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules

temperature

2.1: Water

tertiary structure

4.6: Protein function, domains and cooperativity
4.7.3: Tertiary, Quaternary, and Symmetrical Structures
4.7.4: Secondary Structural Motifs and Domains
4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure
4.7.6: Intrinsically Disordered Proteins
4.7.7: Fibrillar Proteins

thermodynamics

6: Enzyme Thermodynamics
6.3: Endergonic and Exergonic Reactions
6.5: Thermodynamics
6.6: Energy in Metabolism

transthyretin

4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules

triglyceride

1.7: Types of Biological Molecules

U

uncompetitive inhibition

8.1: Regulation of Enzyme Activity

unsaturated fatty acid

1.7: Types of Biological Molecules

V

van der Waals interaction

1.6: The Building Blocks of Molecules

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Upper-Level Introductory Biochemistry

Webpages: 146

All licenses found:

- CC BY-SA 4.0: 98.6% (144 pages)
- Undeclared: 1.4% (2 pages)

By Page

- Upper-Level Introductory Biochemistry - *CC BY-SA 4.0*
  - Front Matter - *CC BY-SA 4.0*
    - TitlePage - *CC BY-SA 4.0*
    - InfoPage - *CC BY-SA 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-SA 4.0*
    - Preface from the Authors - *CC BY-SA 4.0*
    - Licensing - *CC BY-SA 4.0*
    - Biochemistry Foundation Concepts and Learning Goals and Objectives - *CC BY-SA 4.0*
    - Embedded Interactive Mathematical Graphs and Models (Vcell and Minisidewider) - *CC BY-SA 4.0*
  - 1: Molecules of Life - *CC BY-SA 4.0*
    - 1.1: Cellular Foundations of Biochemistry - *CC BY-SA 4.0*
    - 1.2: Chemical Foundations - *CC BY-SA 4.0*
    - 1.3: Physical-Chemical Foundations - *CC BY-SA 4.0*
    - 1.4: Genetic Foundations - *CC BY-SA 4.0*
    - 1.5: Essential elements for life - *CC BY-SA 4.0*
    - 1.6: The Building Blocks of Molecules - *CC BY-SA 4.0*
    - 1.7: Types of Biological Molecules - *CC BY-SA 4.0*
    - 1.8: Chapter 1 Questions - *CC BY-SA 4.0*
    - 1.9: Abiotic Origins of Life - *CC BY-SA 4.0*
  - 2: Chemistry of Water - *CC BY-SA 4.0*
    - 2.1: Water - *CC BY-SA 4.0*
    - 2.2: The multiple roles of water - *CC BY-SA 4.0*
    - 2.3: pH and Buffers - *CC BY-SA 4.0*
    - 2.4: Weak Acids and Bases, pH and pKa - *CC BY-SA 4.0*
    - 2.5: Buffering against pH Changes in Biological Systems - *CC BY-SA 4.0*
    - 2.6: Solubility in an aqueous world - noncovalent interactions in depth - *CC BY-SA 4.0*
    - 2.7: Solubility in an aqueous world - The Hydrophobic Effect - *CC BY-SA 4.0*
    - 2.8: Chapter 2 Questions - *CC BY-SA 4.0*
  - 3: Amino Acids and Peptides - *CC BY-SA 4.0*
    - 3.1: Amino Acids and Peptides - *CC BY-SA 4.0*
    - 3.2: Amino Acid Charges - *CC BY-SA 4.0*
    - 3.3: Cysteine Chemistry - *CC BY-SA 4.0*
    - 3.4: 3.1-Structure and Property of the Naturally-Occurring Amino Acids - *CC BY-SA 4.0*
    - 3.5: From Amino Acids to Peptides - *CC BY-SA 4.0*
  - 4: Proteins- Structure and Folding - *CC BY-SA 4.0*
    - 4.1: The Structure of Proteins- An Overview - *CC BY-SA 4.0*
    - 4.2: Levels of Protein Structure - *CC BY-SA 4.0*
    - 4.3: Secondary Structure - *CC BY-SA 4.0*
    - 4.4: Protein folding - *CC BY-SA 4.0*
    - 4.5: Analyses of Protein Structure - *CC BY-SA 4.0*
    - 4.6: Protein function, domains and cooperativity - *CC BY-SA 4.0*
    - 4.7: The Three-Dimensional Structure of Proteins - *CC BY-SA 4.0*
      - 4.7.1: Main Chain Conformations - *CC BY-SA 4.0*
      - 4.7.2: Secondary Structure and Loops - *CC BY-SA 4.0*
      - 4.7.3: Tertiary, Quaternary, and Symmetrical Structures - *CC BY-SA 4.0*
      - 4.7.4: Secondary Structural Motifs and Domains - *CC BY-SA 4.0*
      - 4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure - *CC BY-SA 4.0*
      - 4.7.6: Intrinsically Disordered Proteins - *CC BY-SA 4.0*
      - 4.7.7: Fibrillar Proteins - *CC BY-SA 4.0*
      - 4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics - *CC BY-SA 4.0*
      - 4.7.9: Protein Stability - Thermodynamics - *CC BY-SA 4.0*
      - 4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules - *CC BY-SA 4.0*

- 4.7.11: Biomolecular Visualization - Conceptions and Misconceptions - *CC BY-SA 4.0*
- 4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation - *CC BY-SA 4.0*
- 4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools - *CC BY-SA 4.0*
- 4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24) - *CC BY-SA 4.0*
- 4.7.15: Chapter 4 Questions - *CC BY-SA 4.0*
- 4.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs - *CC BY-SA 4.0*

- 5: Protein Purification - *CC BY-SA 4.0*
  - 5.1: Protein Purification - *CC BY-SA 4.0*
  - 5.2: Cell Disruption - *CC BY-SA 4.0*
  - 5.3: Fractionation - *CC BY-SA 4.0*
  - 5.4: Electrophoresis - *CC BY-SA 4.0*
  - 5.5: Electrophoresis - *CC BY-SA 4.0*
  - 5.6: Gel Exclusion Chromatography - *CC BY-SA 4.0*
  - 5.7: Ion Exchange Chromatography - *CC BY-SA 4.0*
  - 5.8: Affinity Chromatography - *CC BY-SA 4.0*
  - 5.9: High Performance Liquid Chromatography (HPLC) - *CC BY-SA 4.0*
  - 5.10: Histidine Tagging - *CC BY-SA 4.0*
  - 5.11: Quantification of Protein Concentration - *CC BY-SA 4.0*
  - 5.12: Binding - Enzyme Linked Immunosorbant Assays (ELISAs) - *CC BY-SA 4.0*
- 6: Enzyme Thermodynamics - *CC BY-SA 4.0*
  - 6.1: Enzymes - *CC BY-SA 4.0*
  - 6.2: Enzyme Commission Number - *CC BY-SA 4.0*
  - 6.3: Endergonic and Exergonic Reactions - *CC BY-SA 4.0*
  - 6.4: The Laws of Thermodynamics - *CC BY-SA 4.0*
  - 6.5: Thermodynamics - *CC BY-SA 4.0*
  - 6.6: Energy in Metabolism - *CC BY-SA 4.0*
- 7: Enzyme Kinetics - *CC BY-SA 4.0*
  - 7.1: Enzyme Kinetics - *CC BY-SA 4.0*
  - 7.2: Enzyme Activity - *CC BY-SA 4.0*
    - 7.2.1: How Enzymes Work - *CC BY-SA 4.0*
    - 7.2.2: Kinetics without Enzymes - *CC BY-SA 4.0*
    - 7.2.3: Kinetics with Enzymes - *CC BY-SA 4.0*
    - 7.2.4: Enzyme Inhibition - *CC BY-SA 4.0*
    - 7.2.5: Enzymatic Reaction Mechanisms - *CC BY-SA 4.0*
    - 7.2.6: Enzymes and Protein Regulation - *CC BY-SA 4.0*
    - 7.2.7: Ribozymes - RNA Enzymes - *CC BY-SA 4.0*
    - 7.2.8: Cofactors and Catalysis - A Little Help From My Friends - *CC BY-SA 4.0*
  - 7.3: Basic Principles of Catalysis - *CC BY-SA 4.0*
  - 7.4: Derivation of Michaelis-Menten equation - *CC BY-SA 4.0*
  - 7.5: Binding - The First Step Towards Protein Function - *CC BY-SA 4.0*
- 8: Enzyme Regulation - *CC BY-SA 4.0*
  - 8.1: Regulation of Enzyme Activity - *CC BY-SA 4.0*
  - 8.2: Mechanisms of Catalysis - *CC BY-SA 4.0*
  - 8.3: Control of Enzymatic Activity - *CC BY-SA 4.0*
  - 8.4: Oxygen-Binding Proteins and Allosterism - *CC BY-SA 4.0*
- 9: Carbohydrates - *CC BY-SA 4.0*
  - 9.1: Carbohydrates overview - *CC BY-SA 4.0*
  - 9.2: Monosaccharides - *CC BY-SA 4.0*
  - 9.3: Disaccharides - *Undeclared*
  - 9.4: Polysaccharides - *CC BY-SA 4.0*
  - 9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls - *CC BY-SA 4.0*
  - 9.6: The Sugar Code and Lectin Decoding - *CC BY-SA 4.0*
  - 9.7: Chapter 7 Problems - *CC BY-SA 4.0*
  - 9.8: Chapter 7 Problems - Answer Key - *CC BY-SA 4.0*
- 10: Nucleic Acids - *CC BY-SA 4.0*
  - 10.1: Structure and Function - Nucleic Acids - *CC BY-SA 4.0*
  - 10.2: Nucleic Acids - Structure and Function - *CC BY-SA 4.0*
  - 10.3: Nucleic Acids - RNA Structure and Function - *CC BY-SA 4.0*
  - 10.4: Nucleic Acids - Comparison of DNA and RNA - *CC BY-SA 4.0*
  - 10.5: Enzymes for Genetic modifications - *CC BY-SA 4.0*
  - 10.6: DNA Isolation, Sequencing, and Synthesis - *CC BY-SA 4.0*
  - 10.7: Bioinformatics - *CC BY-SA 4.0*
- 11: Lipids - *CC BY-SA 4.0*
  - 11.1: Introduction to lipids - *CC BY-SA 4.0*
  - 11.2: Lipids Aggregates in Water - Micelles and Liposomes - *CC BY-SA 4.0*
  - 11.3: Structure and Function - Lipids and Membranes - *CC BY-SA 4.0*

- 12: Membranes and Membrane Proteins - *CC BY-SA 4.0*
  - 12.1: Basic Concepts in Membranes - *CC BY-SA 4.0*
  - 12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics - *CC BY-SA 4.0*
  - 12.3: Working with Lipids - *CC BY-SA 4.0*
  - 12.4: Transport in Membranes - *CC BY-SA 4.0*
  - 12.5: Problems - *CC BY-SA 4.0*
  - 12.6: Membranes and Membrane Proteins - *CC BY-SA 4.0*
  - 12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion - *CC BY-SA 4.0*
  - 12.8: Diffusion Across a Membrane - Channels - *CC BY-SA 4.0*
  - 12.9: Diffusion Across a Membrane - Pores - *CC BY-SA 4.0*
  - 12.10: Active Transport - *CC BY-SA 4.0*
- 13: Biochemistry of Our Senses - *CC BY-SA 4.0*
  - 13.1: Cell Signaling - *CC BY-SA 4.0*
  - 13.2: Ligand-gated Ion Channel Receptors - *CC BY-SA 4.0*
  - 13.3: Nuclear Hormone Receptors - *CC BY-SA 4.0*
  - 13.4: G-protein Coupled Receptors (GPCRs) - *CC BY-SA 4.0*
  - 13.5: Receptor Tyrosine Kinases (RTKs) - *CC BY-SA 4.0*
  - 13.6: Gated Ion Channels - Neural Signaling - *CC BY-SA 4.0*
  - 13.7: Signal Transduction - Vision and Olfaction - *CC BY-SA 4.0*
  - 13.8: Signal Transduction - Taste (Gustation) - *CC BY-SA 4.0*
  - 13.9: Signal Transduction - Temperature - *CC BY-SA 4.0*
  - 13.10: Signal Transduction - Pressure - *CC BY-SA 4.0*
- Back Matter - *CC BY-SA 4.0*
  - Index - *CC BY-SA 4.0*
  - Glossary - *CC BY-SA 4.0*
  - Detailed Licensing - *CC BY-SA 4.0*
  - Detailed Licensing - *CC BY-SA 4.0*

Detailed Licensing

Overview

Title: Upper-Level Introductory Biochemistry

Webpages: 146

All licenses found:

- CC BY-SA 4.0: 98.6% (144 pages)
- Undeclared: 1.4% (2 pages)

By Page

- Upper-Level Introductory Biochemistry - *CC BY-SA 4.0*
  - Front Matter - *CC BY-SA 4.0*
    - TitlePage - *CC BY-SA 4.0*
    - InfoPage - *CC BY-SA 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-SA 4.0*
    - Preface from the Authors - *CC BY-SA 4.0*
    - Licensing - *CC BY-SA 4.0*
    - Biochemistry Foundation Concepts and Learning Goals and Objectives - *CC BY-SA 4.0*
    - Embedded Interactive Mathematical Graphs and Models (Vcell and Minisidewider) - *CC BY-SA 4.0*
  - 1: Molecules of Life - *CC BY-SA 4.0*
    - 1.1: Cellular Foundations of Biochemistry - *CC BY-SA 4.0*
    - 1.2: Chemical Foundations - *CC BY-SA 4.0*
    - 1.3: Physical-Chemical Foundations - *CC BY-SA 4.0*
    - 1.4: Genetic Foundations - *CC BY-SA 4.0*
    - 1.5: Essential elements for life - *CC BY-SA 4.0*
    - 1.6: The Building Blocks of Molecules - *CC BY-SA 4.0*
    - 1.7: Types of Biological Molecules - *CC BY-SA 4.0*
    - 1.8: Chapter 1 Questions - *CC BY-SA 4.0*
    - 1.9: Abiotic Origins of Life - *CC BY-SA 4.0*
  - 2: Chemistry of Water - *CC BY-SA 4.0*
    - 2.1: Water - *CC BY-SA 4.0*
    - 2.2: The multiple roles of water - *CC BY-SA 4.0*
    - 2.3: pH and Buffers - *CC BY-SA 4.0*
    - 2.4: Weak Acids and Bases, pH and pKa - *CC BY-SA 4.0*
    - 2.5: Buffering against pH Changes in Biological Systems - *CC BY-SA 4.0*
    - 2.6: Solubility in an aqueous world - noncovalent interactions in depth - *CC BY-SA 4.0*
    - 2.7: Solubility in an aqueous world - The Hydrophobic Effect - *CC BY-SA 4.0*
    - 2.8: Chapter 2 Questions - *CC BY-SA 4.0*
  - 3: Amino Acids and Peptides - *CC BY-SA 4.0*
    - 3.1: Amino Acids and Peptides - *CC BY-SA 4.0*
    - 3.2: Amino Acid Charges - *CC BY-SA 4.0*
    - 3.3: Cysteine Chemistry - *CC BY-SA 4.0*
    - 3.4: 3.1-Structure and Property of the Naturally-Occurring Amino Acids - *CC BY-SA 4.0*
    - 3.5: From Amino Acids to Peptides - *CC BY-SA 4.0*
  - 4: Proteins- Structure and Folding - *CC BY-SA 4.0*
    - 4.1: The Structure of Proteins- An Overview - *CC BY-SA 4.0*
    - 4.2: Levels of Protein Structure - *CC BY-SA 4.0*
    - 4.3: Secondary Structure - *CC BY-SA 4.0*
    - 4.4: Protein folding - *CC BY-SA 4.0*
    - 4.5: Analyses of Protein Structure - *CC BY-SA 4.0*
    - 4.6: Protein function, domains and cooperativity - *CC BY-SA 4.0*
    - 4.7: The Three-Dimensional Structure of Proteins - *CC BY-SA 4.0*
      - 4.7.1: Main Chain Conformations - *CC BY-SA 4.0*
      - 4.7.2: Secondary Structure and Loops - *CC BY-SA 4.0*
      - 4.7.3: Tertiary, Quaternary, and Symmetrical Structures - *CC BY-SA 4.0*
      - 4.7.4: Secondary Structural Motifs and Domains - *CC BY-SA 4.0*
      - 4.7.5: Protein with Alpha, Alpha-Beta, Beta and Little Secondary Structure - *CC BY-SA 4.0*
      - 4.7.6: Intrinsically Disordered Proteins - *CC BY-SA 4.0*
      - 4.7.7: Fibrillar Proteins - *CC BY-SA 4.0*
      - 4.7.8: Protein Folding and Unfolding (Denaturation) - Dynamics - *CC BY-SA 4.0*
      - 4.7.9: Protein Stability - Thermodynamics - *CC BY-SA 4.0*
      - 4.7.10: Protein Aggregates - Amyloids, Prions and Intracellular Granules - *CC BY-SA 4.0*

- 4.7.11: Biomolecular Visualization - Conceptions and Misconceptions - *CC BY-SA 4.0*
- 4.7.12: Laboratory Determination of the Thermodynamic Parameters for Protein Denaturation - *CC BY-SA 4.0*
- 4.7.13: Predicting Structure and Function of Biomolecules Through Natural Language Processing Tools - *CC BY-SA 4.0*
- 4.7.14: Predicting Structure from Sequence and Sequence from Structure/Function (New 10/24) - *CC BY-SA 4.0*
- 4.7.15: Chapter 4 Questions - *CC BY-SA 4.0*
- 4.8: Problems - Predicting Protein Structure and Function Using Machine Learning and AI Programs - *CC BY-SA 4.0*

- 5: Protein Purification - *CC BY-SA 4.0*
  - 5.1: Protein Purification - *CC BY-SA 4.0*
  - 5.2: Cell Disruption - *CC BY-SA 4.0*
  - 5.3: Fractionation - *CC BY-SA 4.0*
  - 5.4: Electrophoresis - *CC BY-SA 4.0*
  - 5.5: Electrophoresis - *CC BY-SA 4.0*
  - 5.6: Gel Exclusion Chromatography - *CC BY-SA 4.0*
  - 5.7: Ion Exchange Chromatography - *CC BY-SA 4.0*
  - 5.8: Affinity Chromatography - *CC BY-SA 4.0*
  - 5.9: High Performance Liquid Chromatography (HPLC) - *CC BY-SA 4.0*
  - 5.10: Histidine Tagging - *CC BY-SA 4.0*
  - 5.11: Quantification of Protein Concentration - *CC BY-SA 4.0*
  - 5.12: Binding - Enzyme Linked Immunosorbant Assays (ELISAs) - *CC BY-SA 4.0*
- 6: Enzyme Thermodynamics - *CC BY-SA 4.0*
  - 6.1: Enzymes - *CC BY-SA 4.0*
  - 6.2: Enzyme Commission Number - *CC BY-SA 4.0*
  - 6.3: Endergonic and Exergonic Reactions - *CC BY-SA 4.0*
  - 6.4: The Laws of Thermodynamics - *CC BY-SA 4.0*
  - 6.5: Thermodynamics - *CC BY-SA 4.0*
  - 6.6: Energy in Metabolism - *CC BY-SA 4.0*
- 7: Enzyme Kinetics - *CC BY-SA 4.0*
  - 7.1: Enzyme Kinetics - *CC BY-SA 4.0*
  - 7.2: Enzyme Activity - *CC BY-SA 4.0*
    - 7.2.1: How Enzymes Work - *CC BY-SA 4.0*
    - 7.2.2: Kinetics without Enzymes - *CC BY-SA 4.0*
    - 7.2.3: Kinetics with Enzymes - *CC BY-SA 4.0*
    - 7.2.4: Enzyme Inhibition - *CC BY-SA 4.0*
    - 7.2.5: Enzymatic Reaction Mechanisms - *CC BY-SA 4.0*
    - 7.2.6: Enzymes and Protein Regulation - *CC BY-SA 4.0*
    - 7.2.7: Ribozymes - RNA Enzymes - *CC BY-SA 4.0*
    - 7.2.8: Cofactors and Catalysis - A Little Help From My Friends - *CC BY-SA 4.0*
  - 7.3: Basic Principles of Catalysis - *CC BY-SA 4.0*
  - 7.4: Derivation of Michaelis-Menten equation - *CC BY-SA 4.0*
  - 7.5: Binding - The First Step Towards Protein Function - *CC BY-SA 4.0*
- 8: Enzyme Regulation - *CC BY-SA 4.0*
  - 8.1: Regulation of Enzyme Activity - *CC BY-SA 4.0*
  - 8.2: Mechanisms of Catalysis - *CC BY-SA 4.0*
  - 8.3: Control of Enzymatic Activity - *CC BY-SA 4.0*
  - 8.4: Oxygen-Binding Proteins and Allosterism - *CC BY-SA 4.0*
- 9: Carbohydrates - *CC BY-SA 4.0*
  - 9.1: Carbohydrates overview - *CC BY-SA 4.0*
  - 9.2: Monosaccharides - *CC BY-SA 4.0*
  - 9.3: Disaccharides - *Undeclared*
  - 9.4: Polysaccharides - *CC BY-SA 4.0*
  - 9.5: Glycoconjugates - Proteoglycans, Glycoproteins, Glycolipids and Cell Walls - *CC BY-SA 4.0*
  - 9.6: The Sugar Code and Lectin Decoding - *CC BY-SA 4.0*
  - 9.7: Chapter 7 Problems - *CC BY-SA 4.0*
  - 9.8: Chapter 7 Problems - Answer Key - *CC BY-SA 4.0*
- 10: Nucleic Acids - *CC BY-SA 4.0*
  - 10.1: Structure and Function - Nucleic Acids - *CC BY-SA 4.0*
  - 10.2: Nucleic Acids - Structure and Function - *CC BY-SA 4.0*
  - 10.3: Nucleic Acids - RNA Structure and Function - *CC BY-SA 4.0*
  - 10.4: Nucleic Acids - Comparison of DNA and RNA - *CC BY-SA 4.0*
  - 10.5: Enzymes for Genetic modifications - *CC BY-SA 4.0*
  - 10.6: DNA Isolation, Sequencing, and Synthesis - *CC BY-SA 4.0*
  - 10.7: Bioinformatics - *CC BY-SA 4.0*
- 11: Lipids - *CC BY-SA 4.0*
  - 11.1: Introduction to lipids - *CC BY-SA 4.0*
  - 11.2: Lipids Aggregates in Water - Micelles and Liposomes - *CC BY-SA 4.0*
  - 11.3: Structure and Function - Lipids and Membranes - *CC BY-SA 4.0*

- 12: Membranes and Membrane Proteins - *CC BY-SA 4.0*
  - 12.1: Basic Concepts in Membranes - *CC BY-SA 4.0*
  - 12.2: Membrane Bilayer and Monolayer Assemblies - Structures and Dynamics - *CC BY-SA 4.0*
  - 12.3: Working with Lipids - *CC BY-SA 4.0*
  - 12.4: Transport in Membranes - *CC BY-SA 4.0*
  - 12.5: Problems - *CC BY-SA 4.0*
  - 12.6: Membranes and Membrane Proteins - *CC BY-SA 4.0*
  - 12.7: Diffusion Across a Membrane - Passive and Facilitated Diffusion - *CC BY-SA 4.0*
  - 12.8: Diffusion Across a Membrane - Channels - *CC BY-SA 4.0*
  - 12.9: Diffusion Across a Membrane - Pores - *CC BY-SA 4.0*
  - 12.10: Active Transport - *CC BY-SA 4.0*
- 13: Biochemistry of Our Senses - *CC BY-SA 4.0*
  - 13.1: Cell Signaling - *CC BY-SA 4.0*
  - 13.2: Ligand-gated Ion Channel Receptors - *CC BY-SA 4.0*
  - 13.3: Nuclear Hormone Receptors - *CC BY-SA 4.0*
  - 13.4: G-protein Coupled Receptors (GPCRs) - *CC BY-SA 4.0*
  - 13.5: Receptor Tyrosine Kinases (RTKs) - *CC BY-SA 4.0*
  - 13.6: Gated Ion Channels - Neural Signaling - *CC BY-SA 4.0*
  - 13.7: Signal Transduction - Vision and Olfaction - *CC BY-SA 4.0*
  - 13.8: Signal Transduction - Taste (Gustation) - *CC BY-SA 4.0*
  - 13.9: Signal Transduction - Temperature - *CC BY-SA 4.0*
  - 13.10: Signal Transduction - Pressure - *CC BY-SA 4.0*
- Back Matter - *CC BY-SA 4.0*
  - Index - *CC BY-SA 4.0*
  - Glossary - *CC BY-SA 4.0*
  - Detailed Licensing - *CC BY-SA 4.0*
  - Detailed Licensing - *CC BY-SA 4.0*