CHEM 4320/5320: BIOCHEMISTRY 1

University of Arkansas Little Rock

LibreTexts™

CHEM 4320/5320: Biochemistry 1

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 03/24/2026

TABLE OF CONTENTS

Licensing

1: Properties of the Twenty Common Amino Acids

- 1.1: Nomenclature of Amino acids
- 1.2: Structure of Amino Acids
- 1.3: Properties of Amino Acids
  - 1.3.1. Charged Nature of Amino Acid
  - 1.3.2. Stereochemistry of Amino Acids
- 1.4: Reactions of Amino Acids
  - 1.4.1 Acid-base Chemistry of Amino Acids
  - 1.4.2. Acid-Base Reactions of Amino Acids

2: Proteins Structure: from Amino Acid Sequence to Three Dimensional Structure

- 2.1: The Structure of Proteins
- 2.2: Protein Sequencing
- 2.3: Protein Structural Determination
- 2.4: Protein Folding and Prions
- 2.5: Denaturation of proteins
- 2.6: Amino Acids and Proteins (Exercises)

3: Methods of Protein Purification and Characterization

- 3.1: Protein Purification
- 3.2: Cell Disruption
- 3.3:Cell Fractionation and Centrifugation
- 3.4: Chromatography
  - 3.4.1. Affinity Chromatography
  - 3.4.2. Gel Exclusion Chromatography
  - 3.4.3. Ion Exchange Chromatography
- 3.5: Electrophoresis
  - 3.5.1: Blotting
  - 3.5.2: Exercises

4: Overview of Hemoglobin and Myoglobin

- 4.1: Myoglobin, Hemoglobin, and their Ligands
- 4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin
- 4.3: Exercises

5: Michaelis-Menten Enzyme Kinetics, Inhibitors, pH optima; Bi-Substrate Reactions

- 5.1: Catalytic Efficiency of Enzymes
- 5.2: Enzyme Parameters

- 5.3: Michaelis-Menten Kinetics
- 5.4: Enzyme Inhibition
- 5.5: Temperature, pH, and enzyme concentration on the rate of a reaction
- 5.6: Multi-Substrate Sequential Mechanisms
- 5.7: Double displacement reaction

6: Classification and Catalytic Strategies of Enzymes

- 6.1: Serine proteases
- 6.2: Transititon State Analogs and Catalytic Antibodies
- 6.3: Restriction Endonuclease

7: Regulation of Enzyme Activity

- 7.1: Control of Metabolism Through Enzyme Regulation
- 7.2: Amino Acids, Proteins, and Enzymes (Summary)
- 7.3: Exercises

8: Carbohydrate Structures, Stereochemistry, and Glycosides

- 8.1: Carbohydrates Fundamentals
- 8.2: Monosaccharides
- 8.3: Disaccharides
- 8.4: Oligosaccharides
- 8.5: Polysaccharides
- 8.6: Exercises
- Index

10: Pyruvate Dehydrogenase Links Glycolysis to Krebs Cycle

- 10.1: The Krebs Cycle (Citric Acid Cycle)
- 10.2: Exercises

11: Electron Transport Chain and Oxidative Phosphorylation

- 11.1: ETC and Oxidative Phosphorylation
- 11.1: Exercises

12: The Flow of Genetic Information: from DNA to RNA and Proteins

- 12.1: The Structure of DNA
- 12.2: DNA Replication
- 12.3: DNA Repair
- 12.4: RNA
  - 12.4.1 Types of RNA
  - 12.4.2. RNA - Transcription
  - 12.4.3. Regulation of Transcription
- 12.5: The Genetic Code
- 12.6: Translation
- 12.7: Exercises

13: Integrated chapter (HIV)

- 13.1: Envelope glycoprotein GP120
- 13.2: HIV-1 protease (PR)
- 13.3: HIV vaccine
- 13.4: Exercises

9: Glycolysis and Gluconeogenesis

- 9.1: Glycolysis - Reaction and Regulation
- 9.2 Gluconeogenesis: Reaction and regulation
- 9.3: Exercises

Index

Glossary

Detailed Licensing

Licensing

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

CHAPTER OVERVIEW

1: Properties of the Twenty Common Amino Acids

Amino acids are exactly what they say they are! They are compounds containing an amino group, $-NH_2$, and a carboxylic acid group, -COOH. The biologically important amino acids have the amino group attached to the carbon atom next door to the -COOH group. They are known as 2-amino acids and are also known as alpha-amino acids.

Topic hierarchy

1.1: Nomenclature of Amino acids
1.2: Structure of Amino Acids
1.3: Properties of Amino Acids
1.3.1. Charged Nature of Amino Acid
1.3.2. Stereochemistry of Amino Acids
1.4: Reactions of Amino Acids
1.4.1 Acid-base Chemistry of Amino Acids
1.4.2. Acid-Base Reactions of Amino Acids

1: Properties of the Twenty Common Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

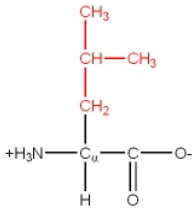
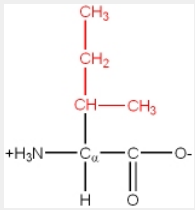
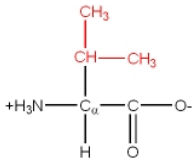
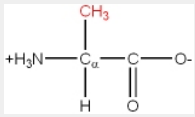
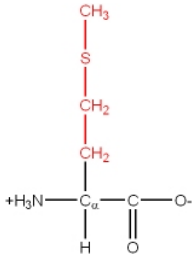
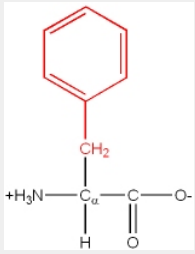
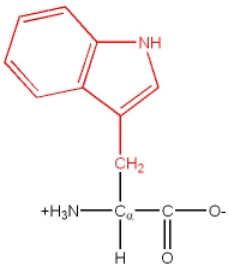
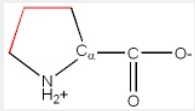
1.1: Nomenclature of Amino acids

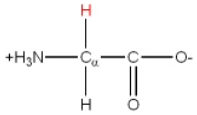
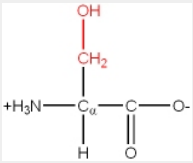
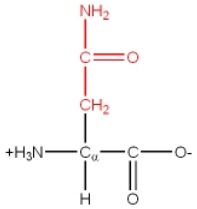
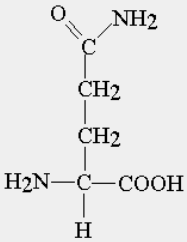
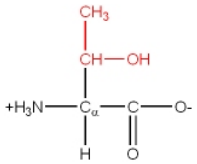
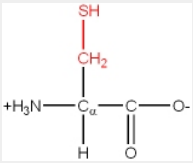
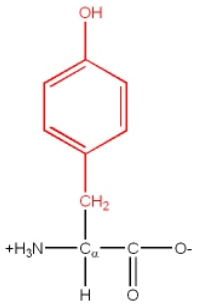
Common amino acids

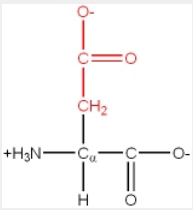
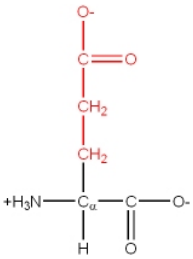
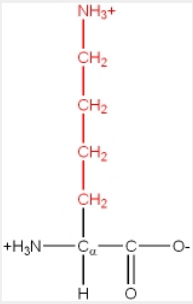
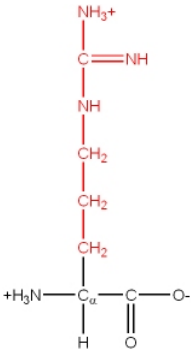
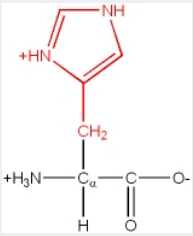
There are 20 common amino acids. They are composed of C, H, O, N and S atoms. They are structurally and chemically different, and also differ in size and volume. Some are branched structures, some are linear, some have ring structures. One of the 20 common amino acids is actually an imino acid. A typical grouping of their chemical nature is as follows:

- Nonpolar (hydrocarbons and one sulfur-containing amino acid). Dispersion forces and hydrophobic effects predominate in their interactions. They cannot H-bond with water and these side chains have a characteristic hydrophobic effect in water.
- Polar uncharged. Contain functional groups that can H-bond with water and other amino acids. Include C, H, O, N and S atoms.
- Acidic. Contain a carboxylic acid functional group with a negative charge at neutral pH. Can H-bond with water, can form ionic interactions, and can also serve as nucleophiles or participate in acid-base chemistry.
- Basic. Nitrogen containing bases (e.g. guanidino, imidazole or amino groups) with a net positive charge at neutral pH. Can serve as proton donors in chemical reactions, and form ionic interactions.

The amino acids have a name, as well as a three letter or single letter mnemonic code:

Type	Name	R-group Structure
Nonpolar	Leucine Leu, L	
	Isoleucine Ile, I	
	Valine Val, V	
	Alanine Ala, A	
	Methionine Met, M	
	Phenylalanine Phe, F	
	Tryptophan Trp, W	
	Proline Pro, P	

Type	Name	R-group Structure
	Glycine Gly, G (note: sometimes included in polar group)	
Polar, uncharged	Serine Ser, S	
	Asparagine Asn, N	
	Glutamine Gln, Q	 glutamine
	Threonine Thr, T	
	Cysteine Cys, C	
	Tyrosine Tyr, Y	

Type	Name	R-group Structure
Acidic	Aspartic acid Asp, D	
	Glutamic acid Glu, E	
Basic	Lysine Lys, K	
	Arginine Arg, R	
	Histidine His, H	

Uncommon amino acids

In addition to the 20 common amino acids, there are several uncommon ones found:

- Hydroxylysine and hydroxyproline. These are found in the protein collagen. Collagen is a fibrous protein made up of three polypeptides that form a stable assembly, but only if the proline and lysine residues are hydroxylated. (requires vitamin C for reduction of these amino acids to hydroxy form)
- Thyroxine, an iodinated derivative of tyrosine, found in thyroglobulin (produced by thyroid gland; requires iodine in diet)

- g-carboxyglutamic acid (i.e. glutamic acid with two carboxyl groups) found in certain blood clotting enzymes (requires vitamin K for production)
- N-methyl arginine and n-acetyl lysine. Found in some DNA binding proteins known as histones

Amino acid derivatives not found in proteins

Some amino acids are made that are not intended for incorporation into proteins, rather they have important functionalities on their own

1. Serotonin (derivative of tryptophan) and g-amino butyric acid (a derivative of glutamic acid) are both neurotransmitters
2. Histamine (derivative of histidine) involved in allergic response
3. Adrenaline (derivative of tyrosine) a hormone
4. Various antibiotics are amino acid derivatives (penicillin)

The Essential Amino Acids

Humans must include adequate amounts of 9 amino acids in their diet.

- Histidine
- Isoleucine
- Leucine
- Lysine
- Methionine (and/or cysteine)
- Phenylalanine (and/or tyrosine)
- Threonine
- Tryptophan
- Valine

These "essential" amino acids cannot be synthesized from other precursors. However, cysteine can partially meet the need for methionine (they both contain sulfur), and tyrosine can partially substitute for phenylalanine. Two of the essential amino acids, **lysine** and **tryptophan**, are poorly represented in most plant proteins. Thus strict vegetarians should ensure that their diet contains sufficient amounts of these two amino acids. 19 of the 20 amino acids listed above can exist in two forms in three dimensions.

Contributors

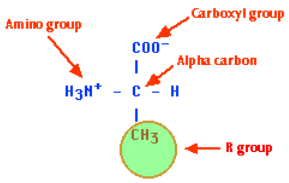
- Mike Blaber (Florida State University)
- John W. Kimball. This content is distributed under a Creative Commons Attribution 3.0 Unported (CC BY 3.0) license and made possible by funding from The Saylor Foundation.

1.1: Nomenclature of Amino acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.2: Structure of Amino Acids

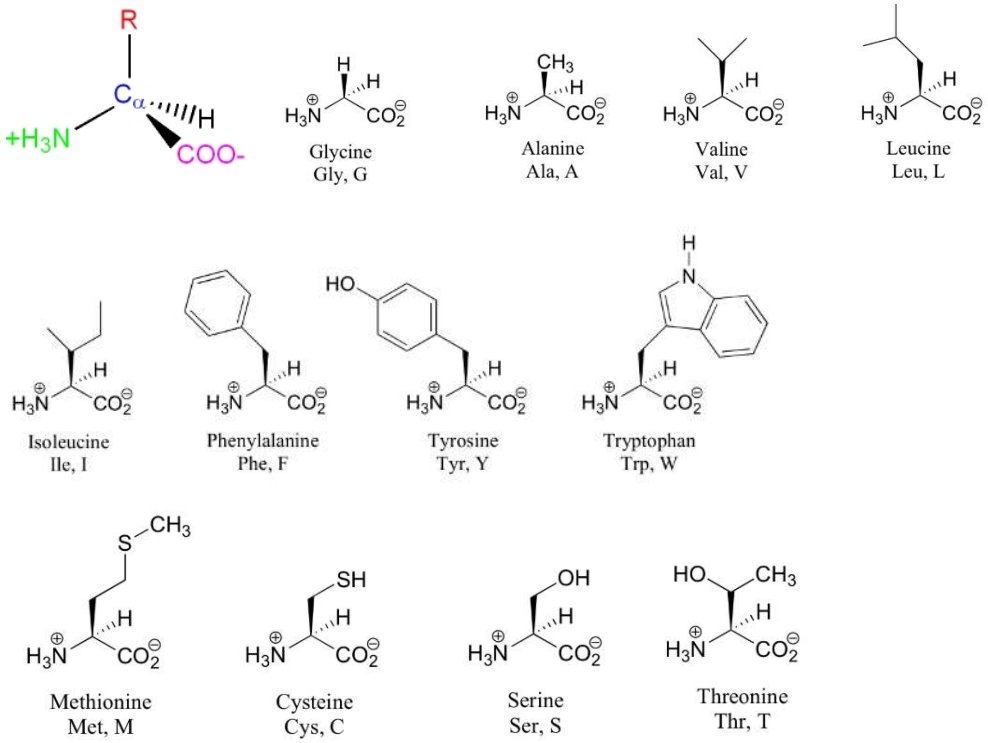
Key structural features

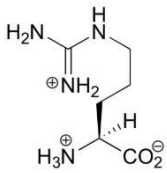
Amino acids are the building blocks (monomers) of proteins. 20 different amino acids are used to synthesize proteins. The **shape** and other properties of each protein is dictated by the precise sequence of amino acids in it.



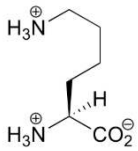
Each amino acid consists of an "**alpha carbon atom"** is tetrahedral and chiral (i.e. each of the four functional groups are different) to which is attached

- a hydrogen atom
- an amino group (hence "amino" acid)
- a carboxyl group (-COOH). This gives up a proton and is thus an acid (hence amino "acid")
- one of 20 different "R" groups. It is the structure of the R group that determines which of the 20 it is and its special properties. Amino acids are also known as "residues". The amino acid shown here is **Alanine.**

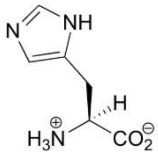




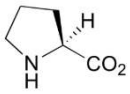
Arginine
Arg, R



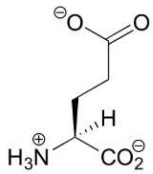
Lysine
Lys, K



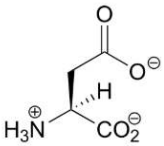
Histidine
His, H



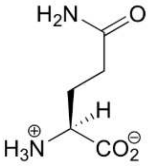
Proline
Pro, P



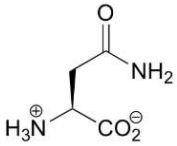
Glutamate
Glu, E



Aspartate
Asp, D

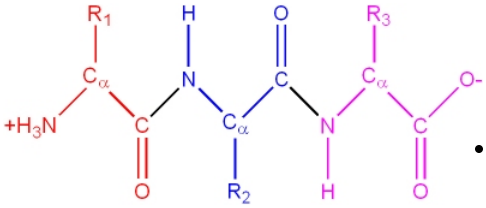


Glutamine
Gln, Q



Asparagine
Asn, N

Amino acid monomers are chemically linked to form linear polymers known as proteins.



Note: this is drawn "flat" for clarity, but the Ca are still tetrahedral (the H atom on the Ca is also not shown in this diagram)

Key structural features:

- A "peptide bond" is formed by a condensation reaction between the carboxylic acid of one amino acid with the amino group of the next amino acid
- The amino acid R1, at the "amino terminus" of the polymer is the "first" amino acid. The residue (R3 in the above diagram) at the carboxyl terminal is known as the "last" amino acid. These termini define the directionality of the protein.

Contributors

- Mike Blaber (Florida State University)
- Organic Chemistry With a Biological Emphasis by Tim Soderberg (University of Minnesota, Morris)

1.2: Structure of Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

SECTION OVERVIEW

1.3: Properties of Amino Acids

1.3.1. Charged Nature of Amino Acid

1.3.2. Stereochemistry of Amino Acids

1.3: Properties of Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.3.1. Charged Nature of Amino Acid

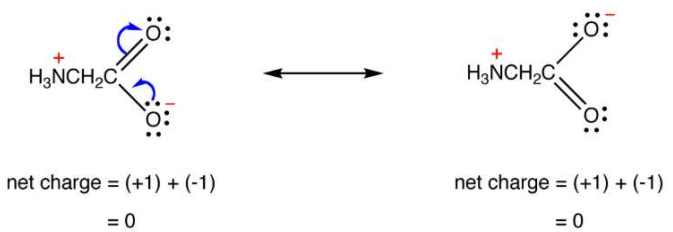
Amino acids are the building blocks used to make proteins and peptides. The different amino acids have interesting properties because they have a variety of structural parts which result in different polarities and solubility.

Each amino acid has at least one amine and one acid functional group as the name implies. The different properties result from variations in the structures of different R groups. The R group is often referred to as the amino acid "side chain".

Zwitterion

Amino acid physical properties indicate a "salt-like" behavior. Amino acids are crystalline solids with relatively high melting points, and most are quite soluble in water and insoluble in non-polar solvents. In solution, the amino acid molecule appears to have a charge which changes with pH. An intramolecular neutralization reaction leads to a salt-like ion called a **zwitterion**. The accepted practice is to show the amino acids in the zwitterion form:

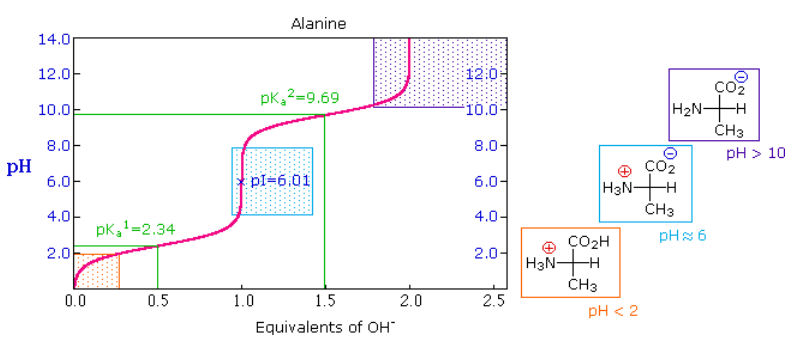
1. The carboxyl group can lose a hydrogen ion to become negatively charged.
2. The amine group can accept a hydrogen ion to become positively charged.



Since amino acids, as well as peptides and proteins, incorporate both acidic and basic functional groups, the predominant molecular species present in an aqueous solution will depend on the pH of the solution. In order to determine the nature of the molecular and ionic species that are present in aqueous solutions at different pH's, we make use of the **Henderson - Hasselbalch Equation**, written below. Here, the pK_a represents the acidity of a specific conjugate acid function (HA). When the pH of the solution equals pK_a, the concentrations of HA and A^- must be equal (log 1 = 0).

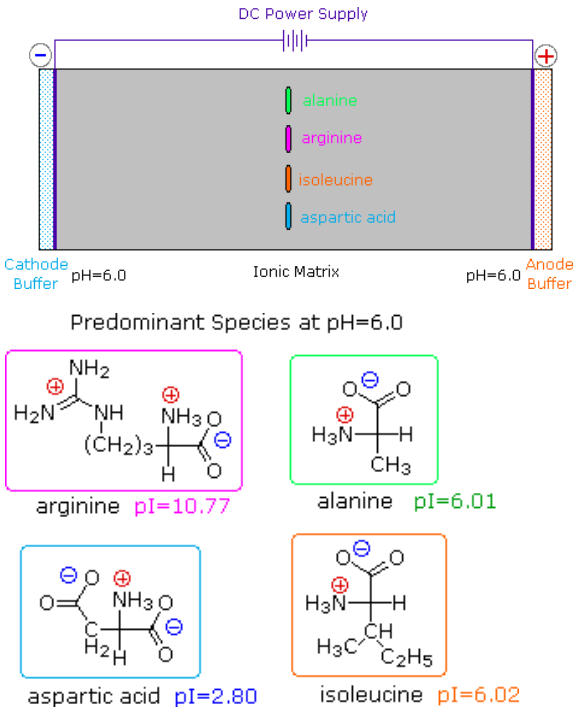
Henderson-Hasselbalch Equation: $pK_a = pH + \log \frac{[HA]}{[A^-]}$

The titration curve for alanine, shown below, demonstrates this relationship. At a pH lower than 2, both the carboxylate and amine functions are protonated, so the alanine molecule has a net positive charge. At a pH greater than 10, the amine exists as a neutral base and the carboxyl as its conjugate base, so the alanine molecule has a net negative charge. At intermediate pH's, the zwitterion concentration increases, and at a characteristic pH, called the **isoelectric point** (**pI**), the negatively and positively charged molecular species are present in equal concentration. Starting from a fully protonated state, the pK_a's of the acidic functions range from 1.8 to 2.4 for $-CO_2H$, and 8.8 to 9.7 for $-NH_3^{(+)}$. The isoelectric points range from 5.5 to 6.2. Titration curves show the neutralization of these acids by added base, and the change in pH during the titration.



The distribution of charged species in a sample can be shown experimentally by observing the movement of solute molecules in an electric field, using the technique of **electrophoresis**. For such experiments an ionic buffer solution is incorporated in a solid matrix layer, composed of paper or a crosslinked gelatin-like substance. A small amount of the amino acid, peptide or protein sample is placed near the center of the matrix strip and an electric potential is applied at the ends of the strip, as shown in the following diagram. The solid structure of the matrix retards the diffusion of the solute molecules, which will remain where they are inserted, unless acted upon by the electrostatic potential. In the example shown here, four different amino acids are examined simultaneously

in a pH 6.00 buffered medium. Note that the colors in the display are only a convenient reference, since these amino acids are colorless.



At pH 6.00 alanine and isoleucine exist on average as neutral zwitterionic molecules, and are not influenced by the electric field. Arginine is a basic amino acid. Both base functions exist as "onium" conjugate acids in the pH 6.00 matrix. The solute molecules of arginine therefore carry an excess positive charge, and they move toward the cathode. The two carboxyl functions in aspartic acid are both ionized at pH 6.00, and the negatively charged solute molecules move toward the anode in the electric field.

The Isoelectric Point

As defined above, the isoelectric point, **pI**, is the pH of an aqueous solution of an amino acid (or peptide) at which the molecules on average have no net charge. In other words, the positively charged groups are exactly balanced by the negatively charged groups. For simple amino acids such as alanine, the pI is an average of the pK_a's of the carboxyl (2.34) and ammonium (9.69) groups. Thus, the pI for alanine is calculated to be: (2.34 + 9.69)/2 = 6.02, the experimentally determined value. If additional acidic or basic groups are present as side-chain functions, the pI is the average of the pK_a's of the two most similar acids. To assist in determining similarity we define two classes of acids. The first consists of acids that are neutral in their protonated form (e.g. CO_2H & SH). The second includes acids that are positively charged in their protonated state (e.g. $-NH_3^+$). In the case of aspartic acid, the similar acids are the alpha-carboxyl function (pK_a = 2.1) and the side-chain carboxyl function (pK_a = 3.9), so pI = (2.1 + 3.9)/2 = 3.0. For arginine, the similar acids are the guanidinium species on the side-chain (pK_a = 12.5) and the alpha-ammonium function (pK_a = 9.0), so the calculated pI = (12.5 + 9.0)/2 = 10.75.

Contributors

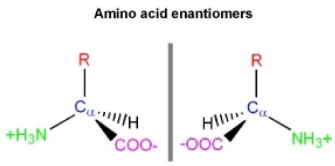
Charles Ophardt (Professor Emeritus, Elmhurst College); Virtual Chembook

- Gamini Gunawardena from the OChemPal site (Utah Valley University)

1.3.1. Charged Nature of Amino Acid is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

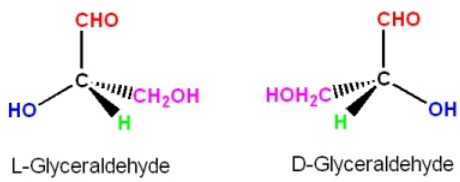
1.3.2. Stereochemistry of Amino Acids

With the exception of glycine, all the 19 other common amino acids have a uniquely different functional group on the central tetrahedral alpha carbon (i.e. C_α). The C_α is termed "chiral" to indicate there are four different constituents and that the Ca is asymmetric. Since the C_α is asymmetric there exists two possible, non-superimposable, mirror images of the amino acids:

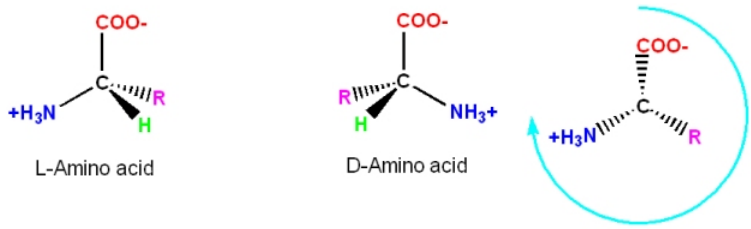


The D, L system

Glyceraldehyde contains a chiral carbon, and therefore, there are two enantiomers of this molecule. One is labeled the "L" form, and the other the "D" form. This is the frame of reference used to describe amino acid enantiomers as being either the "L" or "D" form



Even though the two enantiomers would seem to be essentially equivalent to each other, all common amino acids are found in the "L" enantiomer in living systems. When looking down the H-C, a bond towards the C_α there is a mnemonic to identify the L-enantiomer of amino acids (note: in this view the three functional groups are pointing away from you, and not towards you; the H atom is omitted for clarity - but it would be in front of the C)



Starting with the carbonyl functional group, and going clockwise around the C_α of the L-enantiomer, the three functional groups spell out the word CORN. If you follow the above instructions, it will spell out CONR (a silly, meaningless word) for the D-enantiomer

Optical Activity

Enantiomeric molecules have an optical property known as optical activity - the ability to rotate the plane of plane polarized light. Clockwise rotation is known as "**dextrorotatory**" behavior and counterclockwise rotation is known as **"levorotatory"** behavior.

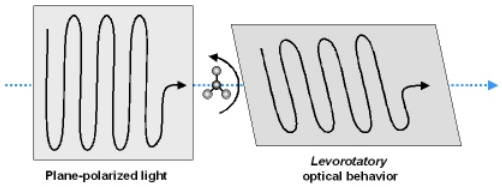


Figure 1.3.2.1: *optical properties of amino acids*

All common amino acids are the L-enantiomer (i.e. their C_α chiral center is the L-enantiomer), based on the structural comparison with L-glyceraldehyde. However, not all L-amino acids are Levorotatory, some are actually Dextrorotatory with regard to their optical activity. To (attempt) to avoid confusion, the optical activities are given as (+) for dextrorotatory, and (-) for levorotatory

- L(+)-alanine (this is the L-enantiomer and it is dextrorotatory)
- L(-)-serine (this is the L-enantiomer and it is levorotatory)

Multiple chiral centers

- Molecules with N chiral centers can exist in 2N isomeric structures
- Isomers that differ in configuration at only one chiral center are called diastereomers

The R,S system of naming chiral centers

A relative ranking of the "priority" of various functional groups is given as:

$$SH > OH > NH_2 > COOH > CHO > CH_2OH > CH_3 > H \tag{1}$$

- A chiral center has four different functional groups. Identify the functional group with the lowest priority
- View the chiral center down the bond from the chiral center to the lowest priority atom

- don't confuse this with the CORN mnemonic method of identifying the L-amino acid chirality by viewing from the H to the Ca)
- Assign priorities to the three other functional groups connected to the chiral center, using the above ranking
- If the priorities of these other groups goes in a clockwise rotation, the chirality is "R". If the priorities of these other groups goes counterclockwise, the chirality is "S". (Note that this assignment has nothing to do with optical activity, and is not using L-glyceraldehyde as a reference molecule).

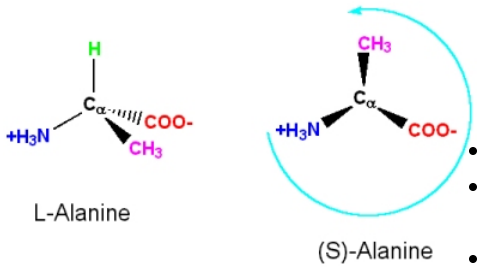


Figure1.3.2.2: Spectroscopic properties of amino acids

This refers to the ability of amino acids to absorb or emit electromagnetic energy at different wavelengths (i.e. energies)

- No amino acids absorb light in the visible spectrum (i.e. they are "colorless").
- If proteins have color (e.g. hemoglobin is red) it is because they contain a bound, non-protein atom, ion or molecule; iron in this case)
- All amino acids absorb in the infrared region (longer wavelengths, weaker energy than visible light)
- Some amino acids absorb in the ultraviolet spectrum (shorter wavelengths, higher energy than visible light)
  - Absorption occurs as electrons rise to higher energy states
  - Electrons in aromatic ring structures absorb in the u.v. spectrum. Such structures comprise the side chains of
  - tryptophan, tyrosine and phenylalanine.

Separation and analysis of amino acid mixtures

The 20 common amino acids differ from one another in several important ways. Here are just two:

- Mass. The smallest amino acid (glycine) has a mass of 57 Da (i.e. g/mol), and the largest (tryptophan) has a mass of 186 Da
- Isoelectric point (pH at which the amino acid has a neutral charge). This is a function of all ionizable groups on the amino acid, including the amino and carboxyl functional groups in addition to any ionizable group on the side chain.

Amino Acid	Mass (Da)	Isoelectric Point	Amino Acid	Mass (Da)	Isoelectric Point
Aspartic Acid	114.11	2.98	Isoleucine	113.16	6.038
Glutamic Acid	129.12	3.08	Glycine	57.05	6.064
Cysteine	103.15	5.02	Alanine	71.09	6.107
Tyrosine	163.18	5.63	Proline	97.12	6.3
Serine	87.08	5.68	Histidine	137.14	7.64
Methionine	131.19	5.74	Lysine	128.17	9.47
Tryptophan	186.12	5.88	Arginine	156.19	10.76
Phenylalanine	147.18	5.91	Threonine	101.11	-
Valine	99.14	6.002	Asparagine	115.09	-
Leucine	113.16	6.036	Glutamine	128.14	-

We can use these differences in physical properties to fractionate complex mixtures of amino acids into individual amino acids

- In looking at the isoelectric point of the different amino acids it seems that they will have different partial charges at a given pH.
- For example, at pH 6.0 some will be negatively charged, and some positively charged.
- For those that are negatively charged, some will be slightly negative, and others strongly negative. Similarly, for those that are positively charged, some will be slightly positive, and others strongly positive
- The charge differences of the amino acids means that they will have different affinities for other cationic or anionic charges

Contributors

The Merck Index, Merck & Co. Inc., Nahway, N.J., 11(1989); CRC Handbook of Chem.& Phys., Cleveland, Ohio, 58(1977)

Mike Blaber (Florida State University)

1.3.2. Stereochemistry of Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

SECTION OVERVIEW

1.4: Reactions of Amino Acids

1.4.1 Acid-base Chemistry of Amino Acids

1.4.2. Acid-Base Reactions of Amino Acids

1.4: Reactions of Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.4.1 Acid-base Chemistry of Amino Acids

Amino acids by themselves have amino (pKa ~9.0-10.5) and carboxyl groups (pKa ~2.0-2.4) that can be titrated. At neutral pH the amino group is protonated, and the carboxyl group is deprotonated. The side chains of acid and basic amino acids, and some polar amino acids can also be titrated:

Amino acid	Functional Group	Side chain pKa
Cysteine	-SH	8.3
Serine	-OH	13
Threonine	-OH	13
Tyrosine	-OH	10.1
Aspartic acid	-COOH	3.9
Glutamic acid	-COOH	4.3
Histidine	Imidazole ring	6.0
Arginine	Guanidino	12.5
Lysine	-NH2	10.5

Physiological pH is near neutral. It would appear that only histidine is of physiological relevance. However, pKa values can be shifted significantly by neighboring charged groups in complex molecular structures.

Reactions of amino acids

- Free amino acids (excluding proline) share similar chemical reactivities due to the common amino and carboxyl groups.
- Different amino acid side chains have different chemical reactivities. Therefore,
- reactivities of different proteins reflects the composition of the unique sequence of amino acids in their structure.

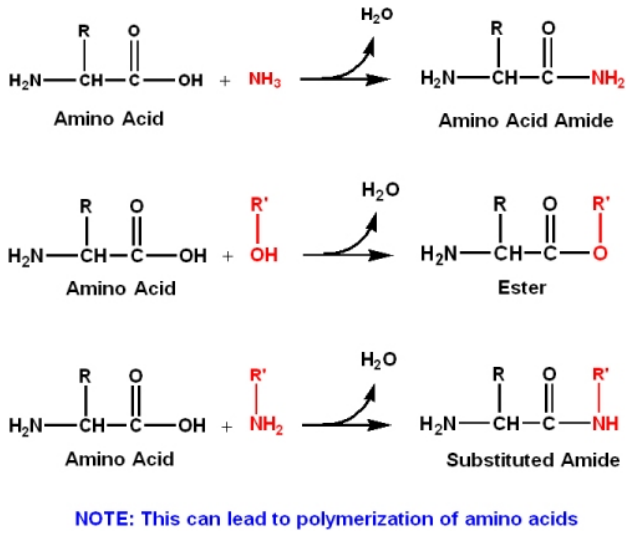


Figure 1.4.1.1: *Some common carboxyl-group reactivities*

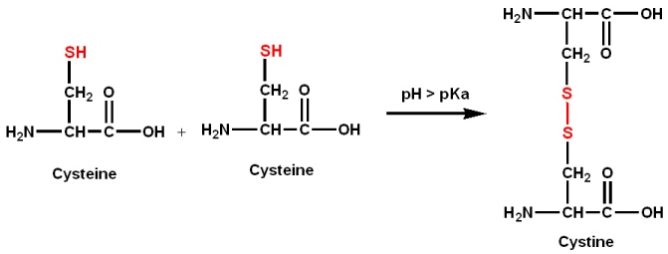


Figure 1.4.1.2: *common side chain reaction involving cysteine*

NOTE: This can covalently link two polypeptide chains in a "disulfide bond" crosslink

Contributors

- Mike Blaber (Florida State University)

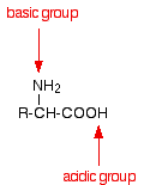
Template:HideTOC

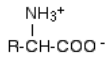
1.4.1 Acid-base Chemistry of Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.4.2. Acid-Base Reactions of Amino Acids

Amino acids as Zwitterions

An amino acid has both a basic amine group ($-NH_2$) and an acidic carboxylic acid group (-COOH). There is an internal transfer of a hydrogen ion from the -COOH group to the $-NH_2$ group to leave an ion with both a negative charge and a positive charge. This is called a zwitterion.



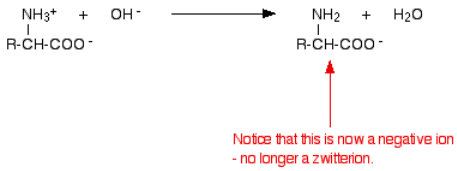


a zwitterion

This is the form that amino acids exist in even in the solid state. If you dissolve the amino acid in water, a simple solution also contains this ion. A zwitterion is a compound with no overall electrical charge, but which contains separate parts which are positively and negatively charged.

Adding an alkali to an amino acid solution

If you increase the pH of a solution of an amino acid by adding hydroxide ions, the hydrogen ion is removed from the $-NH_3^+$ group.



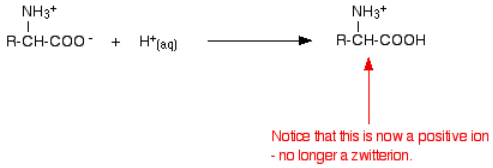
You could show that the amino acid now existed as a negative ion using electrophoresis. In its simplest form, electrophoresis can just consist of a piece of moistened filter paper on a microscope slide with a crocodile clip at each end attached to a battery. A drop of amino acid solution is placed in the center of the paper.

Although the amino acid solution is colourless, its position after a time can be found by spraying it with a solution of ninhydrin. If the paper is allowed to dry and then heated gently, the amino acid shows up as a coloured spot. The amino acid would be found to travel towards the anode (the positive electrode).

Adding an acid to an amino acid solution

If you decrease the pH by adding an acid to a solution of an amino acid, the $-COO^-$ part of the zwitterion picks up a hydrogen ion.

This time, during electrophoresis, the amino acid would move towards the cathode (the negative electrode).



Shifting the pH from one extreme to the other

Suppose you start with the ion we've just produced under acidic conditions and slowly add alkali to it. That ion contains two acidic hydrogens - the one in the -COOH group and the one in the $-NH_3^+$ group. The more acidic of these is the one in the -COOH group, and so that is removed first - and you get back to the zwitterion.

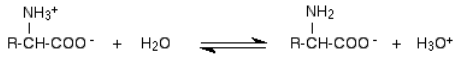
So when you have added just the right amount of alkali, the amino acid no longer has a net positive or negative charge. That means that it wouldn't move towards either the cathode or anode during electrophoresis. The pH at which this lack of movement during electrophoresis happens is known as the isoelectric point of the amino acid. This pH varies from amino acid to amino acid. If you go on adding hydroxide ions, you will get the reaction we've already seen, in which a hydrogen ion is removed from the $-NH_3^+$ group.

You can, of course, reverse the whole process by adding an acid to the ion we've just finished up with. That ion contains two basic groups - the $-NH_2$ group and the $-COO^-$ group. The $-NH_2$ group is the stronger base, and so picks up hydrogen ions first. That leads you back to the zwitterion again.

. . . and, of course, you can keep going by then adding a hydrogen ion to the $-COO^-$ group.

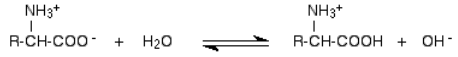
Why isn't the Isoelectric Point of an Amino Acid at pH 7?

When an amino acid dissolves in water, the situation is a little bit more complicated than we tend to pretend at this level. The zwitterion interacts with water molecules - acting as both an acid and a base. As an acid:



The $-NH_3^+$ group is a weak acid and donates a hydrogen ion to a water molecule. Because it is only a weak acid, the position of equilibrium will lie to the left.

As a base:



The $-COO^-$ group is a weak base and takes a hydrogen ion from a water molecule. Again, the equilibrium lies to the left.

When you dissolve an amino acid in water, both of these reactions are happening. However, the positions of the two equilibria aren't identical - they vary depending on the influence of the "R" group. In practice, for the simple amino acids we have been talking about, the position of the first equilibrium lies a bit further to the right than the second one. That means that there will be rather more of the negative ion from the amino acid in the solution than the positive one.

In those circumstances, if you carried out electrophoresis on the unmodified solution, there would be a slight drift of amino acid towards the positive electrode (the anode). To stop that, you need to cut down the amount of the negative ion so that the concentrations of the two ions are identical. You can do that by adding a very small amount of acid to the solution, moving the position of the first equilibrium further to the left. Typically, the pH has to be lowered to about 6 to achieve this. For glycine, for example, the isoelectric point is pH 6.07; for alanine, 6.11; and for serine, 5.68.

Contributors

- Jim Clark (Chemguide.co.uk)

This page titled 1.4.2. Acid-Base Reactions of Amino Acids is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Jim Clark.

2: Proteins Structure: from Amino Acid Sequence to Three Dimensional Structure

Proteins are macromolecules. They are constructed from one or more unbranched chains of amino acids; that is, they are polymers. An average eukaryotic protein contains around 500 amino acids but some are much smaller (the smallest are often called **peptides**) and some much larger (the largest to date is titin a protein found in skeletal and cardiac muscle; one version contains 34,350 amino acids in a single chain!).

Every function in the living cell depends on proteins.

- Motion and locomotion of cells and organisms depends on proteins. [Examples: Muscles, Cilia and Flagella]
- The catalysis of all biochemical reactions is done by **enzymes**, which contain protein.
- The structure of cells, and the extracellular matrix in which they are embedded, is largely made of protein. [Examples: Collagens] (Plants and many microbes depend more on carbohydrates, e.g., cellulose, for support, but these are synthesized by enzymes.)
- The transport of materials in body fluids depends of proteins.
- The receptors for hormones and other signaling molecules are proteins.
- Proteins are an essential nutrient for heterotrophs.
- The transcription factors that turn genes on and off to guide the differentiation of the cell and its later responsiveness to signals reaching it are proteins.
- and many more — **proteins are truly the physical basis of life**.

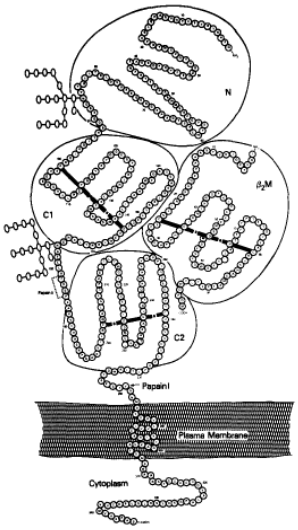


Figure 2.1: *Structure of protein molecules*

The protein consists of two polypeptide chains, a long one on the left of 346 amino acids — it is called the **heavy chain** — and a short one on the right of 99 amino acids. The heavy chain is shown as consisting of 5 main regions or **domains**:

- three extracellular domains, designated here as N (includes the **N-terminal**), C1, and C2;
- a transmembrane domain where the polypeptide chain passes through the plasma membrane of the cell;
- a cytoplasmic domain (with the **C-terminal**) within the cytoplasm of the cell.

Because it is anchored in the plasma membrane of the cell, the heavy chain is called an **integral membrane protein**.

To the right is the protein molecule called **beta-2 microglobulin**. It is not attached to the heavy chain by any covalent bonds, but rather by a number of noncovalent interactions like hydrogen bonds. Proteins associated non-covalently with integral membrane proteins are called **peripheral membrane proteins**.

The dark bars represent disulfide (S—S) bridges linking portions of each external domain (except the N domain). However, the bonds in S—S bridges are no longer than any other covalent bond, so if this molecule could be viewed in its actual tertiary (3D) configuration, we would find that the portions of the polypeptide chains containing the linked Cys are actually close together.

The two objects on the left of the image that look like candlestickrepresent short, branched chains of sugars. The base of each is attached to an **asparagine** (**N**). Proteins with covalently linked carbohydrate are called **glycoproteins**. When the carbohydrate is linked to asparagine, it is said to be "**N-linked**". The presence of sugars on the molecule makes this region hydrophilic.

The amino acids exposed at the surface of the extracellular domains tend to be **hydrophilic** as well. However, most of the amino acids in the transmembrane domain are **hydrophobic**, and the amino acids in the cytoplasmic domain are hydrophilic, which is appropriate for the aqueous medium of the cytosol, but carbohydrate is not found in the intracellular domains of integral membrane proteins.

The regions marked "Papain" represent the places on the long chain that are attacked by the proteinase papain (and made it possible to release the extracellular domains from the plasma membrane for easier analysis). This molecule represents a "single-pass" transmembrane protein; the polypeptide chain traverses the plasma membrane once only. However, many transmembrane proteins pass through several, but always a precisely defined number, of times.

Contributors

Contributed by John W. Kimball

Professor (retired) at Tufts University & Harvard

Contributed by Henry Jakubowski

Professor (Chemistry) at College of St. Benedict/St. John's University

---

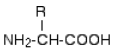
This page titled 2: Proteins Structure: from Amino Acid Sequence to Three Dimensional Structure is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

2.1: The Structure of Proteins

Drawing the amino acids

In chemistry, if you were to draw the structure of a general 2-amino acid, you would probably draw it like this:

However, for drawing the structures of proteins, we usually twist it so that the "R" group sticks out at the side. It is much easier to see what is happening if you do that.

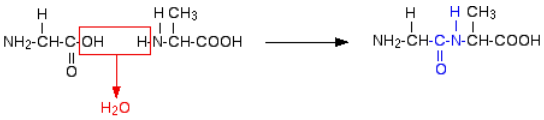


That means that the two simplest amino acids, glycine and alanine, would be shown as:

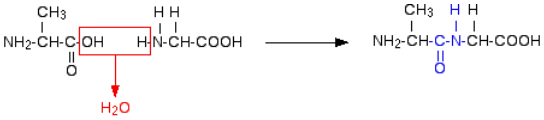
Peptides and polypeptides

Glycine and alanine can combine together with the elimination of a molecule of water to produce a dipeptide. It is possible for this to happen in one of two different ways - so you might get two different dipeptides.

Either:



Or:



In each case, the linkage shown in blue in the structure of the dipeptide is known as a peptide link. In chemistry, this would also be known as an amide link, but since we are now in the realms of biochemistry and biology, we'll use their terms.

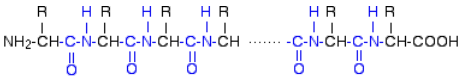
If you joined three amino acids together, you would get a tripeptide. If you joined lots and lots together (as in a protein chain), you get a polypeptide.

A protein chain will have somewhere in the range of 50 to 2000 amino acid residues. You have to use this term because strictly speaking a peptide chain isn't made up of amino acids. When the amino acids combine together, a water molecule is lost. The peptide chain is made up from what is left after the water is lost - in other words, is made up of amino acid residues.

By convention, when you are drawing peptide chains, the $-NH_2$ group which hasn't been converted into a peptide link is written at the left-hand end. The unchanged -COOH group is written at the right-hand end.

The end of the peptide chain with the $-NH_2$ group is known as the N-terminal, and the end with the -COOH group is the C-terminal.

A protein chain (with the N-terminal on the left) will therefore look like this:



The "R" groups come from the 20 amino acids which occur in proteins. The peptide chain is known as the backbone, and the "R" groups are known as side chains.

The primary structure of proteins

The structure of proteins is generally described as having four organizational levels. The first of these is the primary structure, which is the number and sequence of amino acids in a protein's polypeptide chain or chains, beginning with the free amino group and maintained by the peptide bonds connecting each amino acid to the next. The primary structure of insulin, composed of 51 amino acids, is shown in Figure 2.1.1.

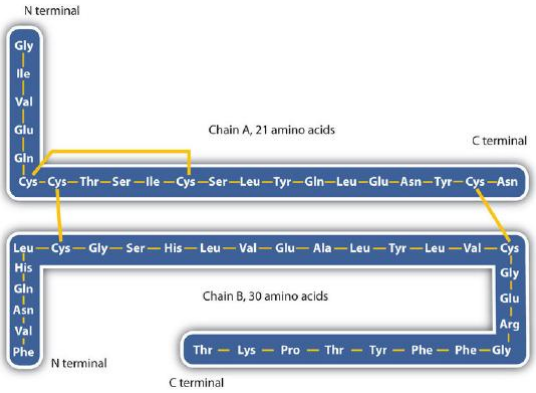


Figure 2.1.1: Primary Structure of Human Insulin. Human insulin, whose amino acid sequence is shown here, is a hormone that is required for the proper metabolism of glucose.

The secondary structure of proteins

A protein molecule is not a random tangle of polypeptide chains. Instead, the chains are arranged in unique but specific conformations. The term secondary structure refers to the fixed arrangement of the polypeptide backbone. On the basis of X ray studies, Linus Pauling and Robert Corey postulated that certain proteins or portions of proteins twist into a spiral or a helix. This helix is stabilized by *intrachain* hydrogen bonding between the carbonyl oxygen atom of one amino acid and the amide hydrogen atom four amino acids up the chain (located on the next turn of the helix) and is known as a right-handed α-helix. X ray data indicate that this helix makes one turn for every 3.6 amino acids, and the side chains of these amino acids project outward from the coiled backbone (Figure 2.1.2). The α-keratins, found in hair and wool, are exclusively α-helical in conformation. Some proteins, such as gamma globulin, chymotrypsin, and cytochrome c, have little or no helical structure. Others, such as hemoglobin and myoglobin, are helical in certain regions but not in others.

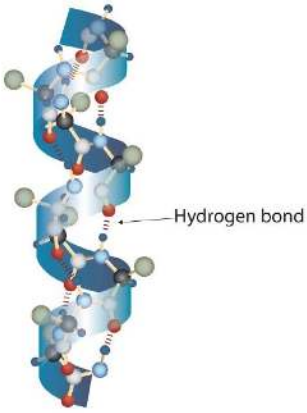


Figure 2.1.2: A Ball-and-Stick Model of an α-Helix. This ball-and-stick model shows the intrachain hydrogen bonding between carbonyl oxygen atoms and amide hydrogen atoms. Each turn of the helix spans 3.6 amino acids. Note that the side chains (represented as green spheres) point out from the helix.

Another common type of secondary structure, called the *β-pleated sheet conformation*, is a sheetlike arrangement in which two or more extended polypeptide chains (or separate regions on the same chain) are aligned side by side. The aligned segments can run either parallel or antiparallel—that is, the N-terminals can face in the same direction on adjacent chains or in different directions—and are connected by *interchain* hydrogen bonding (Figure 2.1.3). The β-pleated sheet is particularly important in structural proteins, such as silk fibroin. It is also seen in portions of many enzymes, such as carboxypeptidase A and lysozyme.

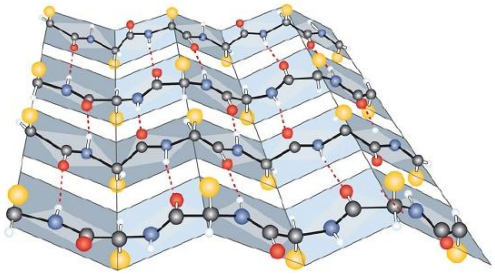


Figure 2.1.3: A Ball-and-Stick Model of the β-Pleated Sheet Structure in Proteins. The side chains extend above or below the sheet and alternate along the chain. The protein chains are held together by interchain hydrogen bonding.

The tertiary structure of proteins

Tertiary structure refers to the unique three-dimensional shape of the protein as a whole, which results from the **folding and bending** of the protein backbone. The tertiary structure is intimately tied to the proper biochemical functioning of the protein. Figure 2.1.4 shows a depiction of the three-dimensional structure of insulin.

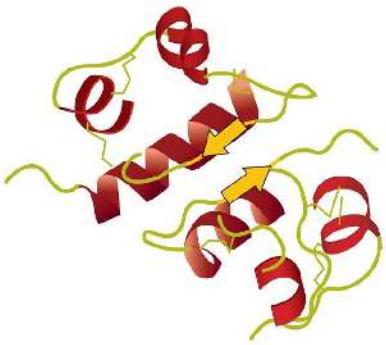


Figure 2.1.4: A Ribbon Model of the Three-Dimensional Structure of Insulin. The spiral regions represent sections of the polypeptide chain that have an α-helical structure, while the broad arrows represent β-pleated sheet structures.

The Quaternary Structure of Protein

When a protein contains more than one polypeptide chain, each chain is called a *subunit*. The arrangement of multiple subunits represents a fourth level of structure, the quaternary structure of a protein. Hemoglobin, with four polypeptide chains or subunits, is the most frequently cited example of a protein having quaternary structure (Figure 2.1.5). The quaternary structure of a protein is produced and stabilized by the same kinds of interactions that produce and maintain the tertiary structure.

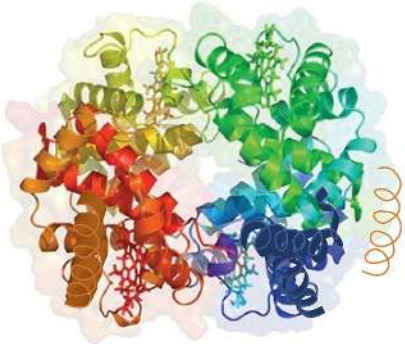


Figure 2.1.5: The Quaternary Structure of Hemoglobin. Hemoglobin is a protein that transports oxygen throughout the body.

Source: Image from the RCSB PDB (www.pdb.org) of PDB ID 1I3D (R.D. Kidd, H.M. Baker, A.J. Mathews, T. Brittain, E.N. Baker (2001) Oligomerization and ligand binding in a homotetrameric hemoglobin: two high-resolution crystal structures of hemoglobin Bart's (gamma(4)), a marker for alpha-thalassemia. Protein Sci. 1739–1749).

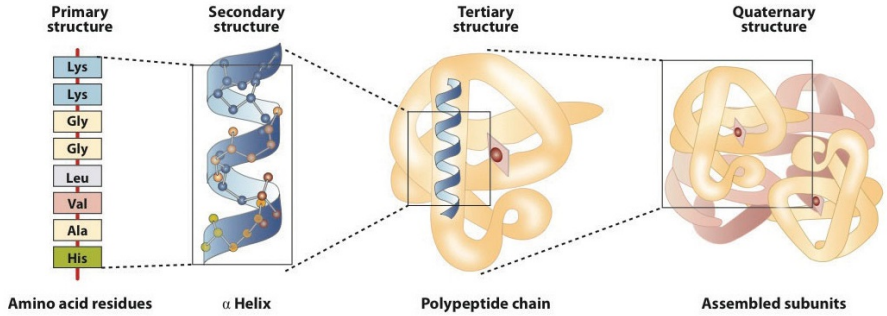


Figure 2.1.6: Levels of Structure in Proteins

The *primary structure* consists of the specific amino acid sequence. The resulting peptide chain can twist into an α-helix, which is one type of *secondary structure*. This helical segment is incorporated into the *tertiary structure* of the folded polypeptide chain. The single polypeptide chain is a subunit that constitutes the *quaternary structure* of a protein, such as hemoglobin that has four polypeptide chains.

What holds a protein into its tertiary structure?

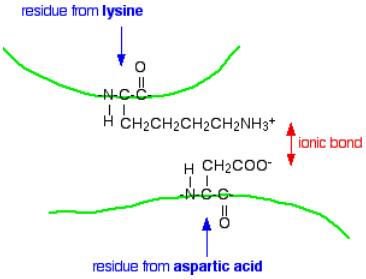
The tertiary structure of a protein is held together by interactions between the side chains - the "R" groups. There are several ways this can happen.

Ionic interactions

Some amino acids (such as aspartic acid and glutamic acid) contain an extra -COOH group. Some amino acids (such as lysine) contain an extra $-NH_2$ group.

You can get a transfer of a hydrogen ion from the -COOH to the $-NH_2$ group to form zwitterions just as in simple amino acids.

You could obviously get an ionic bond between the negative and the positive group if the chains folded in such a way that they were close to each other.

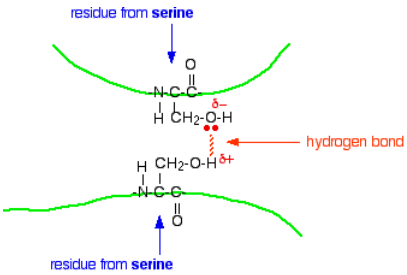


Hydrogen bonds

Notice that we are now talking about hydrogen bonds between side groups - not between groups actually in the backbone of the chain.

Lots of amino acids contain groups in the side chains which have a hydrogen atom attached to either an oxygen or a nitrogen atom. This is a classic situation where hydrogen bonding can occur.

For example, the amino acid serine contains an -OH group in the side chain. You could have a hydrogen bond set up between two serine residues in different parts of a folded chain.



You could easily imagine similar hydrogen bonding involving -OH groups, or -COOH groups, or $-CONH_2$ groups, or $-NH_2$ groups in various combinations - although you would have to be careful to remember that a -COOH group and an $-NH_2$ group would form a zwitterion and produce stronger ionic bonding instead of hydrogen bonds.

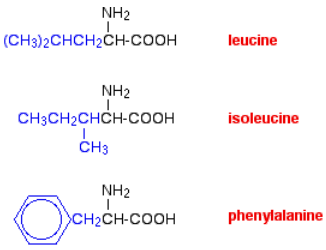
Hydrophobic interactions

This is considered a major driving force for protein folding. The hydrophobic amino acids, especially the hydrophobic R groups, interact with one another to form a "core" and push out water molecules.

van der Waals dispersion forces

Several amino acids have quite large hydrocarbon groups in their side chains. A few examples are shown below. Temporary fluctuating dipoles in one of these groups could induce opposite dipoles in another group on a nearby folded chain.

The dispersion forces set up would be enough to hold the folded structure together.



Sulfur bridges

Sulfur bridges which form between two cysteine residues have already been discussed under primary structures. Wherever you choose to place them doesn't affect how they are formed!

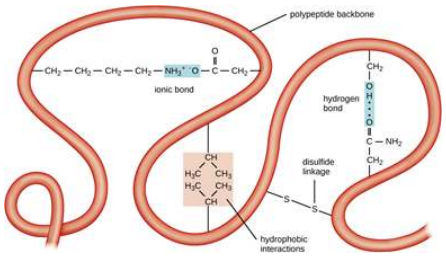
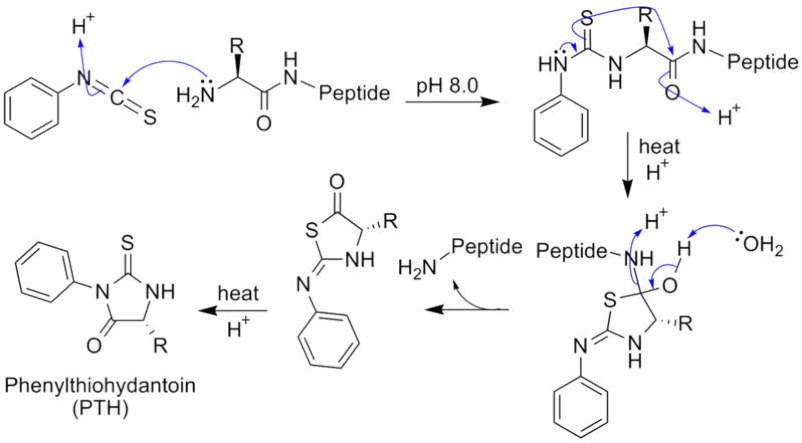


Figure 2.1.7: *Formation of sulfur linkages between two cysteine residues*

This page titled 2.1: The Structure of Proteins is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Jim Clark.

2.2: Protein Sequencing

Edman degradation is the method of sequencing amino acids in a peptide by sequentially removing one residue at a time from the amino end of a peptide. To solve the problem of damaging the protein by hydrolyzing conditions, Pehr Edman created a new way of labeling and cleaving the peptide. Edman thought of a way of removing only one residue at a time, which did not damage the overall sequencing. This was done by adding **Phenyl isothiocyanate (Edman's reagent)**, which creates a phenylthiocarbamoyl derivative with the N-terminal. The N-terminal is then cleaved under less harsh acidic conditions, creating a cyclic compound of phenylthiohydantoin PTH-amino acid. This does not damage the protein and leaves two constituents of the peptide. This method can be repeated for the rest of the residues, separating one residue at a time.



Advantage

Sensitive, simple, and inexpensive method of sequencing amino acids.

Edman degradation is very useful because it does not damage the protein.

The technique allows sequencing of the protein to be done in less time.

Limitations

The technique lacks high-throughput capabilities as sequencing proceeds on samples of single proteins only.

Additional Reading

Sequencing Larger Proteins

Larger proteins cannot be sequenced by the Edman sequencing because of the less than perfect efficiency of the method. A strategy called divide and conquer successfully cleaves the larger protein into smaller, practical amino acids. This is done by using a certain chemical or enzyme which can cleave the protein at specific amino acid residues. The separated peptides can be isolated by chromatography. Then they can be sequenced using the Edman method, because of their smaller size.

In order to put together all the sequences of the different peptides, a method of overlapping peptides is used. The strategy of divide and conquer followed by Edman sequencing is used again a second time, but using a different enzyme or chemical to cleave it into different residues. This allows two different sets of amino acid sequences of the same protein, but at different points. By comparing these two sequences and examining for any overlap between the two, the sequence can be known for the original protein.

For example, trypsin can be used on the initial peptide to cleave it at the carboxyl side of arginine and lysine residues. Using trypsin to cleave the protein and sequencing them individually with Edman degradation will yield many different individual results. Although the sequence of each individual cleaved amino acid segment is known, the order is scrambled. Chymotrypsin, which cleaves on the carboxyl side of aromatic and other bulky nonpolar residues, can be used. The sequence of these segments overlap with those of the trypsin. They can be overlapped to find the original sequence of the initial protein. However, this method is

limited in analyzing larger sized proteins (more than 100 amino acids) because of secondary hydrogen bond interference. Other weak intermolecular bonding such as hydrophobic interactions cannot be properly predicted. Only the linear sequence of a protein can be properly predicted assuming the sequence is small enough.

Mass spectrometry has been replacing traditional methods to determine the molecular mass and structure of a protein. Its power comes from its exquisite sensitivity and modern computational methods to determine structure through comparisons of ion fragment data with computer databases of known protein structures.

In mass spectrometry, a molecule is **first ionized** in an ion source. Sample introduction into the ion source occurs though simple diffusion of gases and volatile liquids from a reservoir, by injection of a liquid sample containing the analyte by spraying a fine mist, or for very large proteins by desorbing a protein from a matrix using a laser.The charged particles are then accelerated by an electric field into a **mass analyzer** where they are subjected to an external magnetic field. The external magnetic field interacts with the magnetic field arising from the movement of the charged particles, causing them to deflect. The deflection is proportional to the mass to charge ratio, m/z. Ions then enter the **detector** which is usually a photomultiplier. Analysis of complex mixtures is done by coupling HPLC with mass spectrometry in a LCMS. (http://www.chm.bris.ac.uk/ms/theory/...onisation.html)

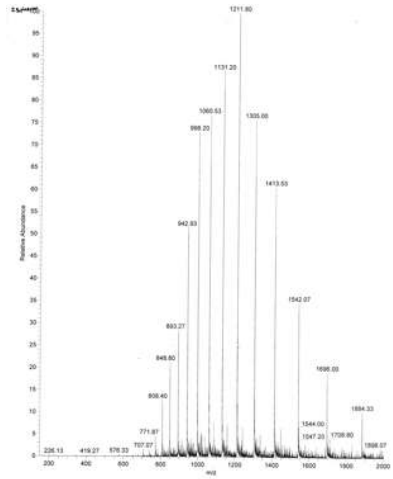
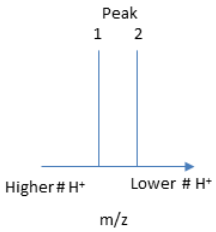


Figure: 2.3.1:
ESI Mass Spectrum of Apo-Myoglobin

The molecular mass of the protein can be determined by analyzing two adjacent peaks, as shown in the figure below.



If 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, then the following equations gives the molecular mass M of the protein for each peak:

Mpeak2 = n(m/z)peak2 - n(1.008) (1)

Mpeak1 = (n+1)(m/z)peak1 - (n+1)(1.008) (2)

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)peak2 - n(1.008) = (n+1)(m/z)peak1 - (n+1) (1.008) *(3)*

Solving for n gives:

n= [(m/z)peak1 - 1.008]/[(m/z)peak2 - (m/z)peak1] *(4)*

Knowing n, the molecular mass M 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.

Advantage

Protein sequencing by MS has become an invaluable tool in the field of **proteomics.** What we know today about protein structure, function, modification and global protein dynamics comes from the development of high-throughput MS proteomics workflows.

Contributors

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)
- Dr. Dietmar Kennepohl FCIC (Professor of Chemistry, Athabasca University)
- Wikibooks (Structural Biopchemistry). The content on this page is licensed under a CC-SA-BY 3.0 license.

2.2: Protein Sequencing is shared under a CC BY-SA license and was authored, remixed, and/or curated by LibreTexts.

2.3: Protein Structural Determination

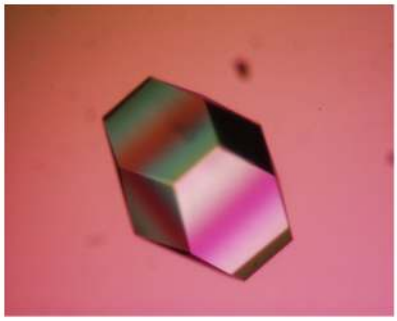
X-ray protein crystallography is a technique by which it is possible to determine the three dimensional positions of each atom in a protein. Now over 100 years old, x-ray crystallography was first used to determine the three dimensional structures of inorganic materials, then small organic molecules, and finally macromolecules like DNA and proteins. To date, about 100,000 protein structures have been published in the Protein Data Bank, with almost 10,000 added every year. To use this technique, the crystallographer obtains protein crystals, records the diffraction pattern formed by x-rays passed through the crystals, and then interprets the data using a computer. The result is a atomic-resolution model of a protein.

Technique

Obtaining crystals

The first and least certain step in crystallography of a protein is obtaining crystals of the protein of interest. Obtaining suitable amounts of the protein of interest is usually carried out in a straightforward manner using established molecular biology techniques such as molecular cloning and affinity chromatography. However, the crystallization step remains the bottleneck for this technique, with some proteins (particularly proteins that exist in the aliphatic environment of the plasma membrane) remaining intransigent to crystallization even in the face of the most diligent crystallographers. Thus, for each protein of interest, a large number of crystallization conditions must be tried, necessitating a relatively large amount (milligrams) of the pure protein.

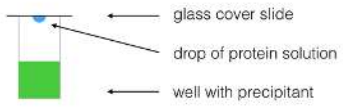
Protein production and purification



To produce suitable amounts of protein, contemporary crystallographers turn to molecular biology's old friend *Escherichia coli*. A gene which codes for the protein of interest is cloned into a small, circular piece of DNA known as an expression plasmid. The expression of the gene is typically under the control of an inducible promoter, and is regulated by the researcher rather than the bacteria. Cells are transformed with the expression plasmid, grown to high density, and induced to express the protein of interest. The cells are lysed chemically with detergents or physically with sonication, and the protein is purified, typically via affinity chromatography. High purity (greater than 95%) is desirable. Often, it takes multiple experiments before the method that obtains maximum protein is found.

Crystallization

The concentrated protein solution obtained is then subjected to a wide variety of crystallization conditions. Since we have no way of knowing *a priori* which set of conditions is right for obtaining crystals of a given protein, many different conditions are tried in parallel using a technique called drop diffusion.



In this technique, a small quantity (typically a microliter) of concentrated protein solution is mixed with an equal volume of precipitant. This drop is separated by air from a large volume of precipitant solution. The drop is hypotonic to the precipitant and slowly equilibrates to the concentration of the large volume of precipitant. Concomitantly, the concentration of protein increases. If this process occurs at just the right rate, the protein precipitates out of solution into an ordered lattice structure: a protein crystal.

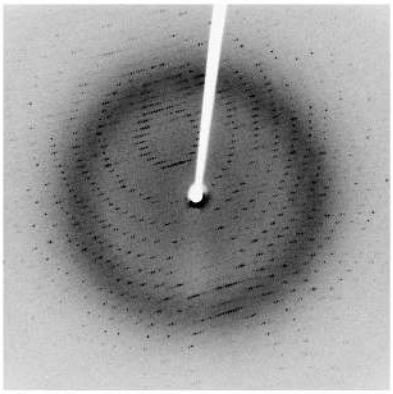
It is often said that this part of crystallography is more of an art than a science, and indeed there is little theoretical guidance available to the crystallographer who wishes to crystallize a new protein. Patience, and to some extent, luck, determine the sucess or failure of the crystallization of any particular protein.

Obtaining x-ray diffraction data

Once crystals of suitable size and composition are obtained, it is necessary to bombard the crystal with x-rays and observe the diffraction pattern. An x-ray diffractometer works in a similar manner to a light microscope. In a light microscope, the subject is irradiated with visible light (400 nm < λλ < 700 nm$), which is diffracted by a lens onto the retina, producing a macroscopic image of a microscopic object. Molecules such as proteins are much smaller than microscopic structures like cells, and, as such, require that a shorter wavelength of radiation be used during diffraction. X-rays, where $100 pm < λλ < 10,000 pm$, are the perfect size to diffract around atoms (32--225 pm), bonds (74--267 pm), and molecules (100 pm to hundreds of Angstroms). However, x-rays are difficult to focus in a manner analogous to the way a lens focuses visible light. Crystallographers employ computational methods to capture the x-ray scattering pattern (pictured at right) and infer the three-dimensional positions of atoms in a molecule.

X-ray sources

Traditionally, x-ray crystallographers filtered and directed the x-rays generated by radioactive cesium in their diffractometers, but today it is much more common to use synchrotron radiation to irradiate samples. Synchrotrons, huge hollow rings used to accelerate electrons for use in studies of subatomic particles, produce huge amounts of tunable (different wavelengths) x-ray radiation that is perfect for irradiating crystals.



Sample preparation

The crystal is suspended in aqueous solution containing a cryoprotectant in the eye of a small loop. The crystal and loop are cooled with a continuous stream of liquid nitrogen to prevent chemical damage by the x-rays. X-rays are directed through the crystal, and the diffraction pattern at any given moment is recorded by a detector. The crystal is rotated sightly and a new diffraction pattern is obtained. This process is repeated through 360 degrees along one axis (typically rotations through a smaller angle on another axis are also recorded to avoid blind spots) until the instrument has recorded a diffraction pattern for each position.

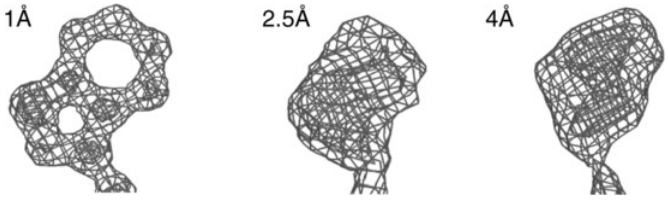
X-ray scattering

As an incident x-ray (electromagnetic wave) overlaps with an electron, it is elastically scattered, generating a secondary wave that has the same wavelength, but different direction, than the incident wave (thus the wave is "scattered" or "diffracted"). Due to the symmetry of the crystal and its many repeated units, these secondary waves interfere constructively at only one point along a circle drawn around the atom that scattered them. It is that point, described by Bragg's Law, that appears as a dark spot on the detector. An example diffraction pattern, from a SARS protease, is displayed at right.

Obtaining an electron density map

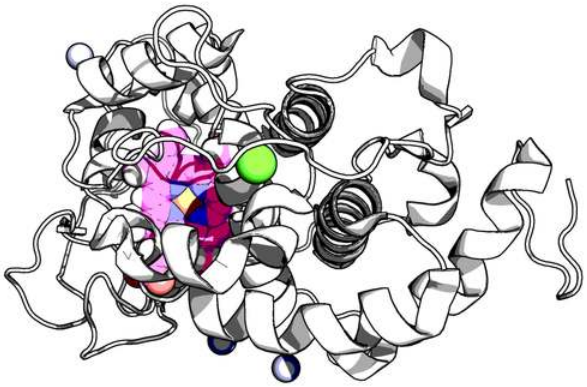
The data recorded by the detector during diffraction are now subjected to computational analysis. First, each spot in each diffraction image is indexed, integrated, merged, and scaled by a computer, producing a single text file from thousands of images. The position of each spot depends on the properties of the crystal, and as such is different for every protein. The process of converting the reciprocal space-representation of the crystal into an interpretable electron density map is known as phasing.

Shown below is the software PyMOL displaying the electron density map (white) for Protein Data Bank structure 4BLL, a peroxidase from the model organism *Pleurotus ostreatus*, overlayed with the model from the PDB structure (pink).



Obtaining a three-dimensional model

With sufficient resolution (less than 1.5 A), it is possible to automatically generate a model based on the electron density map and known bond angles and lengths, and known sizes of atoms. In practice, not all crystallography data is of such high quality. Often, the crystallographer uses molecular visualization software to manually fit a chemical model to the electron density data. The result is a model that can be viewed with molecular visualization software. An example, draw in PyMOL from PDB structure 4BLL, is below.



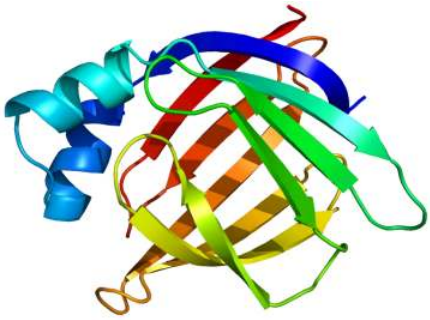
NMR of proteins and peptides

In this chapter you have learned enough about NMR to be able to understand how it is used to solve the structures of relatively small organic molecules. But what about really big organic molecules, like proteins?

X-ray crystallography, not NMR, is the most common way to determine the precise three-dimensional structure of a protein, and in a biochemistry class you will look at many images of protein structures derived from x-ray crystallography. While it is an immensely powerful tool for analyzing protein structure, crystallography has two major drawbacks. First, it relies on a researcher being able to get a protein to form regular, ordered crystals, which can be very challenging. Most proteins are globular, meaning they are (very roughly) spherical in shape. For a molecule to form crystals, it must pack together tightly in an ordered, repeating way: think of a neat stack of cube-shaped objects. Spheres, however, are inherently difficult to pack this way. Imagine trying to make a pile of tennis balls - they just roll apart, because so little of each ball's surface area comes into contact with its neighbor, thus there is very little friction (ie. noncovalent interactions!) holding them together. A large percentage of known proteins simply will not crystallize under any conditions that have been tried - therefore, we cannot determine their structure using x-ray crystallography.

Secondly, a lot of what is most interesting about proteins is how they move: flaps open and close when a substrate binds, or one part of the protein moves over to connect with another part. Protein action is dynamic. A crystal, on the other had, is static, or frozen. A protein structure determined by x-ray crystallography is like a still photograph of leaping dancer: we can infer from the picture what kind of movement might be taking place, but we can't get a full appreciation of the motion.

This leads to NMR, which of course is done in solution. It is easy to get most proteins into aqueous solution, so there are no worries about trying to make crystals. Also, a protein in solution is free to move, so NMR can potentially capture elements of protein dynamics. So why don't scientists always use NMR to look at proteins?



After working through a few NMR structure determination problems in this chapter, you have an appreciation for the brainwork required to figure out the structure of a small organic molecule based on its NMR structure: now imagine doing this with a protein, with its thousands of carbon and hydrogen atoms! Nevertheless, spectroscopists are gradually getting better and better at using NMR and computer-power to do just this. The advanced NMR techniques and methods of analysis are far beyond the scope of our discussion here, but you can see how useful it could be to protein scientists to be able to 'see' what a protein looks like using NMR, and if you are interested in this area of research you can learn about it in more advanced courses.

Note: The Spectral Database of Organic Compounds is a great resource for looking at NMR spectra (both proton and carbon) for a large number of compounds - the more examples you see, the better!

Contributors

- Organic Chemistry With a Biological Emphasis by Tim Soderberg (University of Minnesota, Morris)

2.3: Protein Structural Determination is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.4: Protein Folding and Prions

Protein Folding

Proteins are folded and held together by several forms of molecular interactions. The molecular interactions include the thermodynamic stability of the complex, the hydrophobic interactions and the disulfide bonds formed in the proteins.

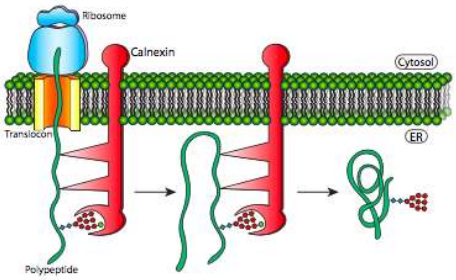


Figure 2.4.1: Protein folding is optimized in the ER. Proteins such as calnexin can temporarily bind to nascent polypeptides, preventing them from forming secondary structures from incomplete information, releasing the protein for folding once the entire polypeptide has been translated.

The biggest factor in a proteins ability to fold is the **thermodynamics** of the structure. The interaction scheme includes the short-range propensity to form extended conformations, residue-dependent long-range contact potentials, and orientation-dependent hydrogen bonds. The thermodynamics are a main stabilizing force within a protein because if it is not in the lowest energy conformation it will continue to move and adjust until it finds its most stable state. The use of energy diagrams and maps are key in finding out when the protein is in the most stable form possible.

The next type of interaction in protein folding is the **hydrophobic interactions** within the protein. The framework model and the hydrophobic collapse model represent two canonical descriptions of the protein folding process. The first places primary reliance on the short-range interactions of secondary structure and the second assigns greater importance to the long-range interactions of tertiary structure.[6] These hydrophobic interactions have an impact not just on the primary structure but then lead to changes seen in the secondary and tertiary structure as well. Globular proteins acquire distinct compact native conformations in water as a result of the hydrophobic effect.[7] When a protein has been folded in the correct way, it usually exists with the hydrophobic core as a result of being hydrated by waters in the system around it which is important because it creates a charged core to the protein and can lead to the creation of channels within the protein. The hydrophobic interactions are found to affect time correlation functions in the vicinity of the native state even though they have no impact on same time characteristics of the structure fluctuations around the native state.[7] The hydrophobic interactions are shown to have an impact on the protein even after it has found the most stable conformation in how the proteins can interact with each other as well as folding themselves.

Another type of interaction seen when the protein is folding is the **disulfide linkages** that form in the protein. (See figure 2.4.2). The disulfide bond, a sulfur- sulfur chemical bond that results from an oxidative process that links nonadjacent (in most cases) cysteine's of a protein.[9] These are a major way that proteins get into their folded form. The types of disulfide bonds are cysteine-cysteine linkage is a stable part of their final folded structure and those in which pairs of cysteines alternate between the reduced and oxidized states.[9] The more common is the linkages that cause the protein to fold together and link back on itself compared to the cysteines that are changing oxidation states because the bonds between cysteines once created are fairly stable. The exchange of disulfide bonding targets is catalyzed by protein disulfide isomerase (PDI).

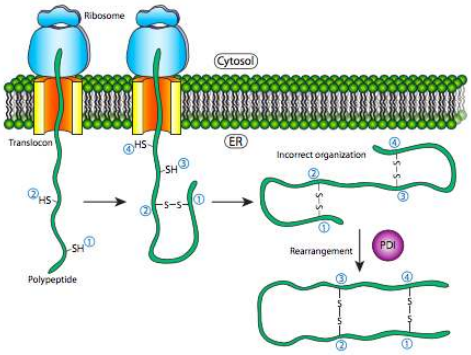


Figure 2.4.2: Disulfide linkages

Denaturation

For proteins, function is dependent on precise structure. Loss of the precise, folded structure of a protein is known as denaturation and is usually accompanied by loss of function. Anyone who has ever worked to purify an enzyme knows how easy it is for one to lose its activity. A few enzymes, such as ribonuclease, are remarkably stable under even very harsh conditions. For most others, a small temperature or pH change can drastically affect activity. The reasons for these differences vary, but relate to 1) the strength of the forces holding the structure together and 2) the ability of a protein to refold itself after being denatured. Let us consider these separately below.

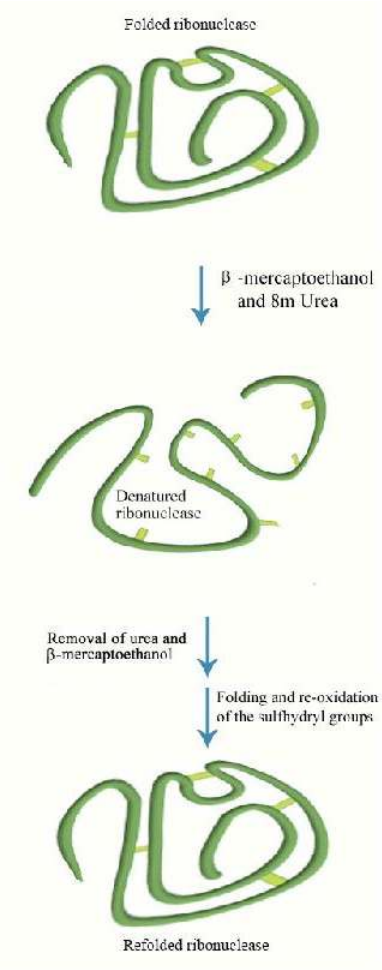


Figure 3.2.18: *Denaturation and renaturation of ribonuclease.*

Forces Stabilizing Structures

Amino acids are linked one to the other by peptide bonds. These covalent bonds are extraordinarily stable at neutral pHs, but can be broken by hydrolysis with heat under acidic conditions. Peptide bonds, however, only stabilize primary structure and, in fact, are the only relevant force responsible for it. Secondary structure, on the other hand, is generally stabilized by weaker forces, including hydrogen bonds. Hydrogen bonds are readily disrupted by heat, urea, or guanidinium chloride.

Forces stabilizing tertiary structure include ionic interactions, disulfide bonds, hydrophobic interactions, metallic bonds, and hydrogen
bonds. Of these, the ionic interactions are most sensitive to pH changes. Hydrophobic bonds are most sensitive to detergents. Thus, washing one's hands helps to kill bacteria by denaturing critical proteins they need to survive. Metallic bonds are sensitive to oxidation/reduction. Breaking disulfide bonds requires either a strong oxidizing agent, such as performic acid or a strong reducing agent
on another disulfide, such as mercaptoethanol or dithiothreitol.

Quaternary structures are stabilized by the same forces as tertiary structure and have the same sensitivities.

Refolding Denatured Proteins

All of the information for protein folding is contained in the primary structure of the protein. It may seem curious then that most proteins do not refold into their proper, fully active form after they have been denatured and the denaturant is removed. A few do, in fact, refold correctly under these circumstances. A good example is bovine ribonuclease (also called RNase). Its catalytic activity is very resistant to heat and urea. However, if one treats the enzyme with mercaptoethanol (which breaks disulfide bonds) prior to urea treatment and
heating, activity is lost, indicating that the covalent disulfide bonds help stabilize the overall enzyme structure. If one allows the enzyme
mixture to cool back down to room temperature, over time some enzyme activity reappears, indicating that ribonuclease can re-fold under the proper conditions.

Irreversible Denaturation

Most enzymes, however, do not behave like ribonuclease. Once denatured, their activity cannot be recovered to any significant extent. This may seem to contradict the idea that folding information is inherent to the sequence of amino acids in the protein. It does not. The reason most enzymes can't refold properly is due to two phenomena. 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 protein synthesis starts. In many cases, the proper folding of newly made polypeptides is also assisted by special proteins called chaperones. Chaperones bind to newly made proteins, preventing interactions that might result in misfolding. Thus, early folding and the assistance of chaperones eliminate some potential "wrong-folding" interactions that can occur if the entire sequence was present when folding started.

Denatured full-length polypeptides have many more potential wrong folds that can occur. A second reason most proteins don't refold properly after denaturation is probably that folding, like any other natural phenomenon, is driven by energy minimization. Though the folded structure may have a low energy, the path leading to it may not be all downhill. Like a chemical reaction that has energies of activation that must be overcome for the reaction to occur, folding likely has peaks and valleys of energy that do not automatically lead directly to the proper fold. Again, folding during synthesis leads the protein along a better-defined path through the energy maze of folding that denatured full-length proteins can't navigate.

Misfunctions

Proteins can malfunction for several reasons. When a protein is miss folded it can lead to denaturation of the protein. **Denaturation** is the loss of protein structure and function.[1] The miss folding does not always lead to complete lack of function but only partial loss of functionality. The miss functioning of proteins can sometimes lead to diseases in the human body.

Alzheimer's Disease

Alzheimer's Disease (AD) is a neurological degenerative disease that affects around 5 million Americans, including nearly half of those who are age 85 or older.[10] The predominant risk factors of AD are age, family history, and heredity. Alzheimer's disease

typically results in memory loss, confusion of time and place, misplacing places, and changes in mood and behavior.[11] AD results in dense plaques in the brain that are comprised of fibrillar β-amyloid proteins with a well-orders β-sheet secondary structure.[12] These plaques visually look like voids in the brain matter (see figure 2.4.3) and are directly connected to the deterioration of thought processes. It has been determined that AD is a protein misfolding disease, where the misfolded protein is directly related to the formation of these plaques in the brain.[13] It is yet to be fully understood what exactly causes this protein misfolding to begin, but several theories point to oxidative stress in the brain to be the initiating factor. This oxidation results in damage to the phospholipids in the brain, which has been found to result in a faster accumulation of amyloid β-proteins.

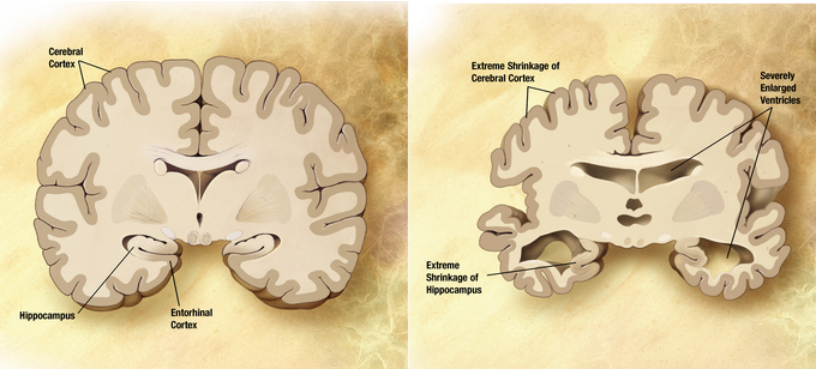


Figure 2.4.3: The figure showing combination of two brain diagrams in one for comparison. In the left normal brain, in the right brain of a person with Alzheimer's disease.

Cystic Fibrosis

Cystic Fibrosis (CF) is a chronic disease that affects 30,000 Americans. The typical affects of CF is a production of thick, sticky mucus that clogs the lungs and leads to life-threatening lung infection, and obstructs the pancreas preventing proper food processing.[15] CF is caused by protein misfolding. This misfolding then results in some change in the protein known as cystic fibrosis transmembrane conductance regulator (CFTR), which can result in this potentially fatal disease.[16] In approximately 70% of CF cases, a deletion of phenylalanine at position 508 in the CFTR is deleted. This deletion of Phe508 seems to be directly connected to the formation of CF.[17] The protein misfolding that results in CF occurs prior to birth, but it is not entirely clear as to why.

Resources

1. Garrett, R.H., and Grisham, C.M. *Biochemistry fourth edition*; Brooks/Cole. Australia, 2010. p. 93-95, 135, 143, 160.
2. Chang, D.T., Syu Y., Lin, P., Predicting the protein-protein interactions using primary structures with predicted protein surface. BCM Bioinformatics. Jan 2010
3. Bodis, P., Schwartz, E., Koepf, M. Cornelissen, J. Rowan A.E., Roeland J. Nolte, M. and Woutersen S. Vibrational self-trapping in beta-sheet structures observed with femtosecond nonlinear infrared spectroscopy. The Journal Of Chemical Physics **131**, 2009
4. Yasuda, S., Yoshidome, T. Oshima, H. Kodama R., Harano, Y. and Kinoshita, M. Effects of side-chain packing on the formation of secondary structures in protein folding. The Journal of Chemical Physics **132,** 2010
5. Papandreoul, N., Berezovsky I.N., Lopes, A., Eliopoulos, E., and Chomilier, J. Universal positions in globular proteins from observation to simulation. Eur. J. Biochem. 271, 4762–4768 (2004)
6. Hausratha, A.C. A Kinetic theory of tertiary contact formation coupled to the helix-coil transition in polypeptides. The Journal of Chemical Physics. 2006
7. Cieplaka M. and Niewieczerza S. Hydrodynamic interactions in protein folding. Journal of Chemical Physics. **130** 2009
8. Gront, D. Kolinskia, A. and Skolnick, J. A new combination of replica exchange Monte Carlo and histogram analysis for protein folding and thermodynamics. Journal of Chemical Physics. Vol. 115 2001
9. Kadokura, H., Katzen F. and Beckwith, J. Protein disulfide bond formation in prokatyotes. Annual Review Biochem 2003.
10. Alzheimer's Disease. Centers for Disease Control and Prevention. 2010, http://www.cdc.gov/aging/aginginfo/alzheimers.htm.
11. Alzheimer's Association. 2011, http://www.alz.org.

12. Riordan, J.M. et al. Identification of the cysticfibrosisgene: cloning and characterization of complementary DNA. *Science* **1989**, 245, 1066-1073.

Prions

A prion is an infectious agent composed of protein in a misfolded form. This is the central idea of the Prion Hypothesis, which remains debated. This is in contrast to all other known infectious agents (virus /bacteria/fungus/parasite) which must contain nucleic acids (either DNA, RNA, or both). The word *prion*, coined in 1982 by Stanley B. Prusiner, is derived from the words protein and infection. Prions are responsible for the transmissible spongiform encephalopathies in a variety of mammals, including bovine spongiform encephalopathy (BSE, also known as “mad cow disease”) in cattle and Creutzfeldt–Jakob disease (CJD) in humans. All known prion diseases affect the structure of the brain or other neural tissue, are currently untreatable and universally fatal.

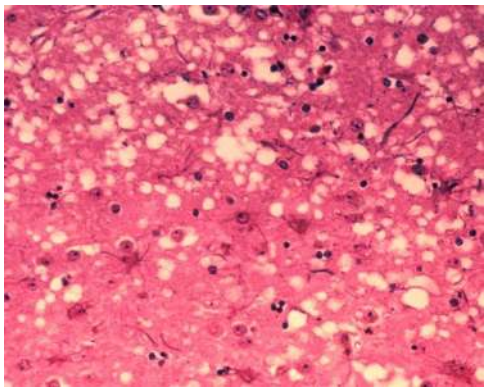


Figure: ***Prion-affected tissue****: This micrograph of brain tissue reveals the cytoarchitectural histopathologic changes found in bovine spongiform encephalopathy. The presence of vacuoles, i.e. microscopic “holes” in the gray matter, gives the brain of BSE-affected cows a sponge-like appearance when tissue sections are examined in the lab.*

Prions propagate by transmitting a misfolded protein state. When a prion enters a healthy organism, it induces existing, properly folded proteins to convert into the disease-associated prion form; it acts as a template to guide the misfolding of more proteins into prion form. These newly-formed prions can then go on to convert more proteins themselves; triggering a chain reaction. All known prions induce the formation of an amyloid fold, in which the protein polymerises into an aggregate consisting of tightly-packed beta sheets. Amyloid aggregates are fibrils, growing at their ends, and replicating when breakage causes two growing ends to become four growing ends.

The incubation period of prion diseases is determined by the exponential growth rate associated with prion replication, which is a balance between the linear growth and the breakage of aggregates. Propagation of the prion depends on the presence of normally-folded protein in which the prion can induce misfolding; animals which do not express the normal form of the prion protein cannot develop nor transmit the disease.

All known mammalian prion diseases are caused by the so-called prion protein, PrP. The endogenous, properly-folded form is denoted PrPC (for Common or Cellular) while the disease-linked, misfolded form is denoted PrPSc (for Scrapie, after one of the diseases first linked to prions and neurodegeneration.) The precise structure of the prion is not known, though they can be formed by combining PrPC, polyadenylic acid, and lipids in a Protein Misfolding Cyclic Amplification (PMCA) reaction.

Proteins showing prion-type behavior are also found in some fungi, which has been useful in helping to understand mammalian prions. Fungal prions do not appear to cause disease in their hosts.

The first hypothesis that tried to explain how prions replicate in a protein-only manner was the heterodimer model. This model assumed that a single PrPSc molecule binds to a single PrPC molecule and catalyzes its conversion into PrPSc. The two PrPSc molecules then come apart and can go on to convert more PrPC. However, a model of prion replication must explain both how prions propagate, and why their spontaneous appearance is so rare. Manfred Eigen showed that the heterodimer model requires PrPSc to be an extraordinarily effective catalyst, increasing the rate of the conversion reaction by a factor of around 10^{15}. This problem does not arise if PrPSc exists only in aggregated forms such as amyloid, where cooperativity may act as a barrier to spontaneous conversion. What is more, despite considerable effort, infectious monomeric PrPSc has never been isolated.

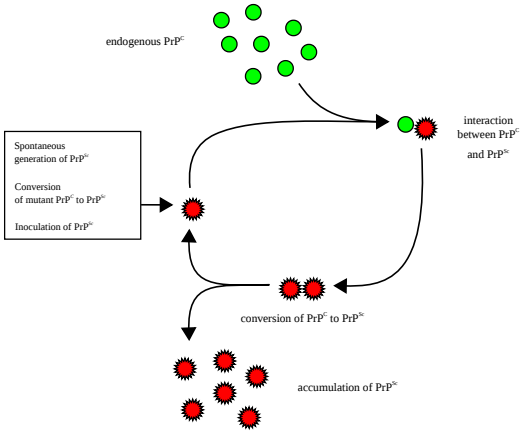


Figure: ***Diagram demonstrating how prion numbers increase****: Heterodimer model of prion propagation.*

An alternative model assumes that PrPSc exists only as fibrils, and that fibril ends bind PrPC and convert it into PrPSc. If this were all, then the quantity of prions would increase linearly, forming even longer fibrils. But exponential growth of both PrPSc and of the quantity of infectious particles is observed during prion disease. This can be explained by taking into account fibril breakage. A mathematical solution for the exponential growth rate resulting from the combination of fibril growth and fibril breakage has been found.

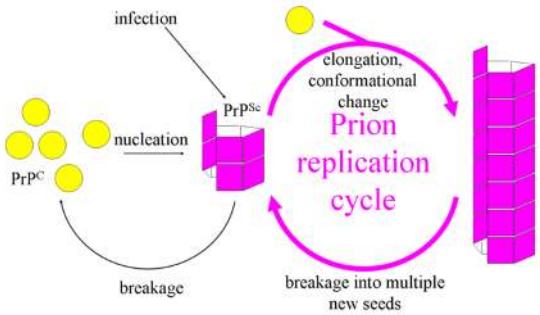


Figure: **The Fibril Model**: Illustrating prion propagation.

The protein-only hypothesis has been criticised by those who feel that the simplest explanation of the evidence to date is viral. For more than a decade, Yale University neuropathologist Laura Manuelidis has been proposing that prion diseases are caused instead by an unidentified slow virus. In January 2007, she and her colleagues published an article reporting to have found a virus in 10%, or less, of their scrapie-infected cells in culture.

The virion hypothesis states that TSEs are caused by a replicable informational molecule (likely to be a nucleic acid) bound to PrP. Many TSEs, including scrapie and BSE, show strains with specific and distinct biological properties, a feature which supporters of the virion hypothesis feel is not explained by prions.

Recent studies propagating TSE infectivity in cell-free reactions and in purified component chemical reactions strongly suggest against TSE's viral nature. Using a similar defined recipe of multiple components (PrP, POPG lipid, RNA), Jiyan Ma and colleagues generated infectious prions from recombinant PrP expressed from E. coli, casting further doubt on this hypothesis.

Key Points

- Prions are responsible for the transmissible spongiform encephalopathies in a variety of mammals, including bovine spongiform encephalopathy and Creutzfeldt–Jakob disease in humans. All known prion diseases affect the structure of neural tissue, are currently untreatable and universally fatal.
- Prions propagate by transmitting a misfolded protein state. When a prion enters a healthy organism, it induces existing, properly-folded proteins to convert into the disease-associated, prion form; it then acts as a template to guide the misfolding of

more proteins into prion form.

- All known mammalian prion diseases are caused by the so-called prion protein, PrP. The endogenous, properly folded, form is denoted PrPC while the disease-linked, misfolded form is denoted PrPSc.The precise structure of the prion is not known.
- The first hypothesis to explain how prions replicate in a protein-only manner was the heterodimer model, which assumed that a single PrPSc molecule binds to a single PrPC molecule and catalyzes its conversion into PrPSc. The two PrPSc molecules then come apart and can go on to convert more PrPC.
- An alternative model assumes that PrPSc exists only as fibrils and that fibril ends bind PrPC and convert it into PrPSc. The exponential growth of both PrPSc and of the quantity of infectious particles observed during prion disease can be explained by taking fibril breakage into account.
- The protein-only hypothesis has been criticised by those who feel that the simplest explanation of the evidence to date is viral. However, recent studies propagating TSE infectivity in cell-free reactions and in purified component chemical reactions strongly suggest against TSE's viral nature.

Key Terms

- **Amyloid**: Insoluble fibrous protein aggregates sharing specific structural traits. They arise from at least 18 inappropriately folded versions of proteins and polypeptides present naturally in the body. These misfolded structures alter their proper configuration such that they erroneously interact with one another or other cell components forming insoluble fibrils. They have been associated with the pathology of more than 20 serious human diseases in that, abnormal accumulation of amyloid fibrils in organs may lead to amyloidosis, and may play a role in various neurodegenerative disorders.
- **Creutzfeldt–Jakob disease**: A degenerative neurological disorder (brain disease) that is incurable and invariably fatal. In CJD, the brain tissue develops holes and takes on a sponge-like texture, due to a type of infectious protein called a prion.
- **prion**: A self-propagating misfolded conformer of a protein that is responsible for a number of diseases that affect the brain and other neural tissue.

LICENSES AND ATTRIBUTIONS

CC LICENSED CONTENT, SHARED PREVIOUSLY

- Curation and Revision. **Provided by**: Boundless.com. **License**: *CC BY-SA: Attribution-ShareAlike*

CC LICENSED CONTENT, SPECIFIC ATTRIBUTION

- Subviral agent. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Subviral_agent%23Subviral_agents. **License**: *CC BY-SA: Attribution-ShareAlike*
- Virology. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Virology. **License**: *CC BY-SA: Attribution-ShareAlike*
- Hepatitis delta virus. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Hepatitis_delta_virus%23Virology. **License**: *CC BY-SA: Attribution-ShareAlike*
- Virus. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Virus. **License**: *CC BY-SA: Attribution-ShareAlike*
- Satellite (biology). **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Satellite_(biology). **License**: *CC BY-SA: Attribution-ShareAlike*
- Defective interfering particle. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Defective_interfering_particle. **License**: *CC BY-SA: Attribution-ShareAlike*
- General Biology/Classification of Living Things/Viruses. **Provided by**: Wikibooks. **Located at**: http://en.wikibooks.org/wiki/General_Biology/Classification_of_Living_Things/Viruses. **License**: *CC BY-SA: Attribution-ShareAlike*
- Helper virus. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Helper%20virus. **License**: *CC BY-SA: Attribution-ShareAlike*
- Satellite. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Satellite. **License**: *CC BY-SA: Attribution-ShareAlike*
- Hepatocellular carcinoma 1. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/File:Hepatocellular_carcinoma_1.jpg. **License**: *Public Domain: No Known Copyright*

- Avsunviroidae. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Avsunviroidae. **License**: *CC BY-SA: Attribution-ShareAlike*
- Pospiviroidae. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Pospiviroidae. **License**: *CC BY-SA: Attribution-ShareAlike*
- Avocado sunblotch viroid. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Avocado_sunblotch_viroid. **License**: *CC BY-SA: Attribution-ShareAlike*
- Viroid. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Viroid. **License**: *CC BY-SA: Attribution-ShareAlike*
- Virusoid. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Virusoid. **License**: *CC BY-SA: Attribution-ShareAlike*
- viroid. **Provided by**: Wiktionary. **Located at**: http://en.wiktionary.org/wiki/viroid. **License**: *CC BY-SA: Attribution-ShareAlike*
- Virusoid. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Virusoid. **License**: *CC BY-SA: Attribution-ShareAlike*
- Hepatocellular carcinoma 1. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/File:Hepatocellular_carcinoma_1.jpg. **License**: *Public Domain: No Known Copyright*
- PSTviroid. **Provided by**: Wikimedia. **Located at**: http://commons.wikimedia.org/wiki/File:PSTviroid.png?uselang=he. **License**: *Public Domain: No Known Copyright*
- Prion. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Prion. **License**: *CC BY-SA: Attribution-ShareAlike*
- prion. **Provided by**: Wiktionary. **Located at**: http://en.wiktionary.org/wiki/prion. **License**: *CC BY-SA: Attribution-ShareAlike*
- Amyloid. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Amyloid. **License**: *CC BY-SA: Attribution-ShareAlike*
- Creutzfeldt. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/Creutzfeldt?Jakob+disease. **License**: *CC BY-SA: Attribution-ShareAlike*
- Hepatocellular carcinoma 1. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/File:Hepatocellular_carcinoma_1.jpg. **License**: *Public Domain: No Known Copyright*
- PSTviroid. **Provided by**: Wikimedia. **Located at**: http://commons.wikimedia.org/wiki/File:PSTviroid.png?uselang=he. **License**: *Public Domain: No Known Copyright*
- Histology bse. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/File:Histology_bse.jpg. **License**: *Public Domain: No Known Copyright*
- Prion propagation. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/File:Prion_propagation.svg. **License**: *CC BY: Attribution*
- Prion replication. **Provided by**: Wikipedia. **Located at**: http://en.wikipedia.org/wiki/File:Pr...eplication.png. **License**: *CC BY-SA: Attribution-ShareAlike*

2.4: Protein Folding and Prions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.5: Denaturation of proteins

The highly organized structures of proteins are truly masterworks of chemical architecture. But highly organized structures tend to have a certain delicacy, and this is true of proteins. Denaturation is the term used for any change in the three-dimensional structure of a protein that renders it incapable of performing its assigned function. **A denatured protein cannot do its job**. (Sometimes denaturation is equated with the precipitation or coagulation of a protein; our definition is a bit broader.) A wide variety of reagents and conditions, such as heat, organic compounds, pH changes, and heavy metal ions can cause protein denaturation.

Table 2.5: Protein Denaturation Methods

Method	Effect on Protein Structure
Heat above 50°C or ultraviolet (UV) radiation	Heat or UV radiation supplies kinetic energy to protein molecules, causing their atoms to vibrate more rapidly and disrupting relatively weak hydrogen bonding and dispersion forces.
Use of organic compounds, such as ethyl alcohol	These compounds are capable of engaging in intermolecular hydrogen bonding with protein molecules, disrupting intramolecular hydrogen bonding within the protein.
Salts of heavy metal ions, such as mercury, silver, and lead	These ions form strong bonds with the carboxylate anions of the acidic amino acids or SH groups of cysteine, disrupting ionic bonds and disulfide linkages.
Alkaloid reagents, such as tannic acid (used in tanning leather)	These reagents combine with positively charged amino groups in proteins to disrupt ionic bonds.

Anyone who has fried an egg has observed denaturation. The clear egg white turns opaque as the albumin denatures and coagulates. No one has yet reversed that process. However, given the proper circumstances and enough time, a protein that has unfolded under sufficiently gentle conditions can refold and may again exhibit biological activity (Figure 2.5.1). Such evidence suggests that, at least for these proteins, the primary structure determines the secondary and tertiary structure. A given sequence of amino acids seems to adopt its particular three-dimensional (3D) arrangement naturally if conditions are right.

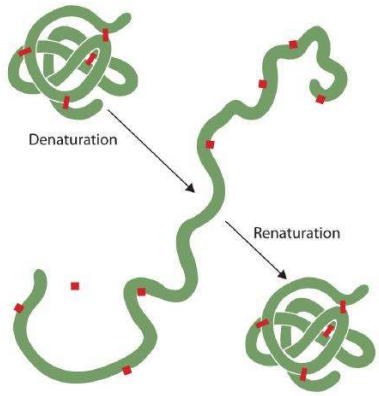


Figure 2.5.1: Denaturation and Renaturation of a Protein. The denaturation (unfolding) and renaturation (refolding) of a protein is depicted. The red boxes represent stabilizing interactions, such as disulfide linkages, hydrogen bonding, and/or ionic bonds.

The primary structures of proteins are quite sturdy. In general, fairly vigorous conditions are needed to hydrolyze peptide bonds. At the secondary through quaternary levels, however, proteins are quite vulnerable to attack, though they vary in their vulnerability to denaturation. The delicately folded globular proteins are much easier to denature than are the tough, fibrous proteins of hair and skin.

Summary

Proteins can be divided into two categories: fibrous, which tend to be insoluble in water, and globular, which are more soluble in water. A protein may have up to four levels of structure. The primary structure consists of the specific amino acid sequence. The

resulting peptide chain can form an α-helix or β-pleated sheet (or local structures not as easily categorized), which is known as secondary structure. These segments of secondary structure are incorporated into the tertiary structure of the folded polypeptide chain. The quaternary structure describes the arrangements of subunits in a protein that contains more than one subunit. Four major types of attractive interactions determine the shape and stability of the folded protein: *ionic bonding, hydrogen bonding, disulfide linkages, and dispersion forces*. A wide variety of reagents and conditions can cause a protein to unfold or denature.

2.5: Denaturation of proteins is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.6: Amino Acids and Proteins (Exercises)

Questions

1.1: Amino Acids

Q1.Read the material at:

https://chem.libretexts.org/Courses/University_of_Arkansas_Little_Rock/CHEM_4320%2F%2F5320%3A_Biochemistry_1/02%3A__Protein_Structure/2.6%3A_Amino_Acids_and_Proteins_(Exercises) and answer the following questions:

a. What are essential amino acids?
b. What are nonessential amino acids?
c. What happens if you are deficient in an amino acid?

Q2. Draw the functional groups present in all amino acids.

Q3. Complete the following for threonine, lysine, and tyrosine.

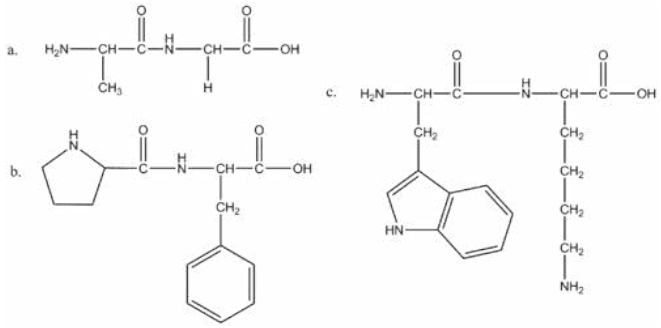
a. Draw the amino acid.
b. Circle the side chain.
c. Identify whether it is polar, nonpolar, acidic, or basic.
d. At what pH will it exist as a zwitterion?
e. What is the range of pH values when it will be positively charged?
f. What is the range of pH values when it will be negatively charged?

1.5: Peptides

Q4. Draw the two dipeptides formed from each pair of amino acids.

a. tyrosine and lysine
b. threonine and gluatmine
c. alanine and histidine

Q5. Draw and give the full names of the amino acids in the following dipeptides.

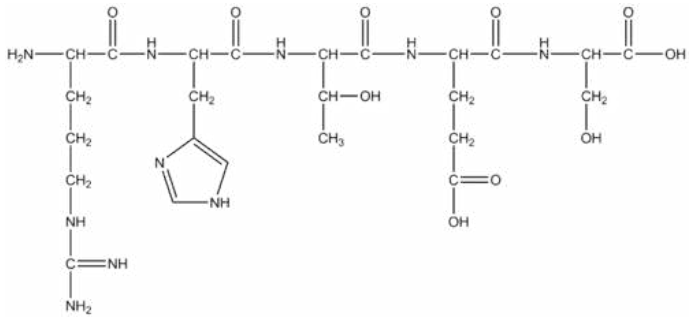


Q6. List of all of the possible polypeptides that can be formed from threonine, alanine, and phenylalanine (use three character abbreviations for each amino acid).

Q7. Draw the following polypeptides.

a. Ser-Tyr-Gln
b. Lys-Met-Gly

Q8. Identify each of the amino acids in the polypeptide and then name it using the three character abbreviations.



2.1: Protein Structure

Q9. Describe the four levels of protein structure.

Q10. What levels of structure involve hydrogen bonding?

Q11. What types of structure is the result of interactions between amino acids that are far apart in the primary structure?

Q12. What types of interactions hold the secondary structure together?

Q13. What types of interactions hold the tertiary structure together?

Q14. What levels of structure are affected by denaturation?

Q15. A protein has one subunit. Would it have a quaternary structure?

Answers

1.1: Amino Acids

Q1

a. Essential amino acids are those you get from your diet.
b. Nonessential amino acids are produced in the body.
c. Illness and/or degradation of body's proteins.

Q2

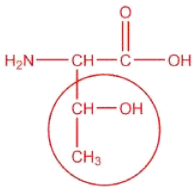
amine and carboxylic acid

Q3

Complete the following for threonine, lysine, and tyrosine.

threonine

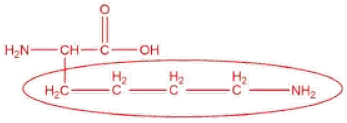
a.



b.
c. polar
d. 5.60
e. < 5.60
f. > 5.60

lysine

a.



b.

c. basic

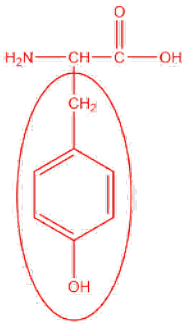
d. 9.47

e. < 9.47

f. > 9.47

tyrosine

a.



b.

c. polar

d. 5.63

e. < 5.63

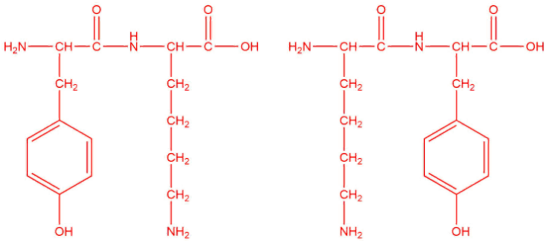
f. > 5.63

1.5: Peptides

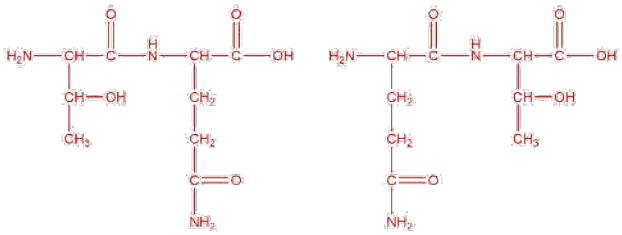
Q4

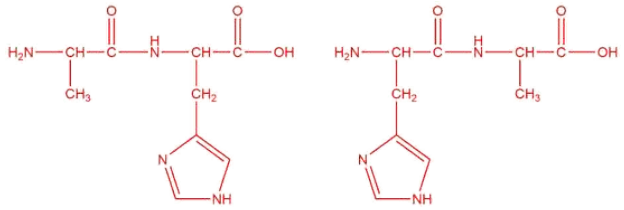
Draw the two dipeptides formed from each pair of amino acids.

a.

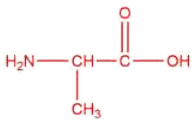
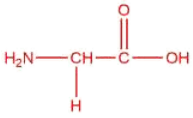
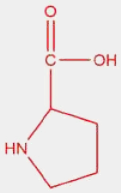
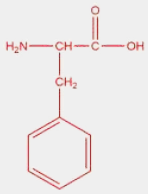
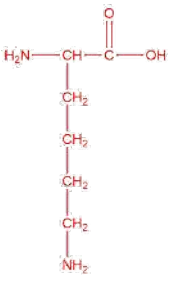


b.



c. 

Q5

a.	alanine 	glycine 
b.	proline 	phenylalanine 
c.	tryptophan	lysine 

Q6

Thr-Ala-Phe

Thr-Phe-Ala

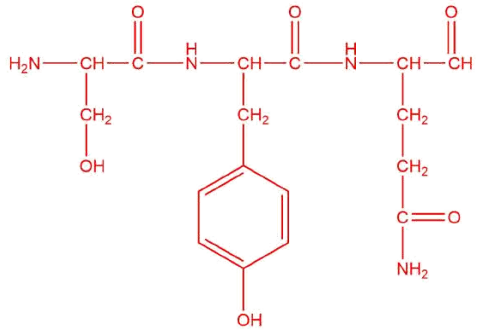
Ala-Thr-Phe

Ala-Phe-Thr

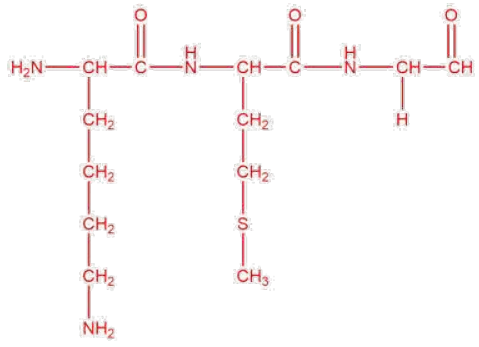
Phe-Ala-Thr

Phe-Thr-Ala

Q7

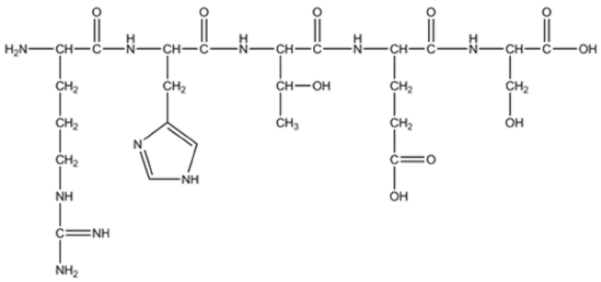


a.



b.

Q8



arginine histidine threonine glutamic acid serine

Arg-His-Thr-Glu-Ser

2.2: Protein Structure

Q9

Primary - sequence of amino acids

Secondary - alpha helix and Beta-pleated sheets held together by hydrogen bonds

Tertiary - third level of structure of protein often forming globular or fibrous structure, held together by variety of attractive forces

Quaternary - complex of multiple proteins held together to function as one, held together by variety of attractive forces (same as tertiary)

Q10

secondary, tertiary, and quaternary structures

Q11

tertiary structures

Q12

hydrogen bonds

Q13

London dispersion forces, hydrogen bonds, dipole-dipole forces, ion-dipole interactions, salt bridges, and disulfide bonds

Q14

secondary, tertiary, and quaternary

Q15

No, a quaternary structure must have multiple subunits.

Concept Review Exercises

1. What is the predominant attractive force that stabilizes the formation of secondary structure in proteins?
2. Distinguish between the tertiary and quaternary levels of protein structure.
3. Briefly describe four ways in which a protein could be denatured.

Answers

1. hydrogen bonding
2. Tertiary structure refers to the unique three-dimensional shape of a single polypeptide chain, while quaternary structure describes the interaction between multiple polypeptide chains for proteins that have more than one polypeptide chain.
3. (1) heat a protein above 50°C or expose it to UV radiation; (2) add organic solvents, such as ethyl alcohol, to a protein solution; (3) add salts of heavy metal ions, such as mercury, silver, or lead; and (4) add alkaloid reagents such as tannic acid

1. Classify each protein as fibrous or globular.
    1. albumin
    2. myosin
    3. fibroin
2. Classify each protein as fibrous or globular.
    1. hemoglobin
    2. keratin
    3. myoglobin
3. A protein has a tertiary structure formed by interactions between the side chains of the following pairs of amino acids. For each pair, identify the strongest type of interaction between these amino acids.
    1. aspartic acid and lysine
    2. phenylalanine and alanine
    3. serine and lysine
    4. two cysteines
4. A protein has a tertiary structure formed by interactions between the side chains of the following pairs of amino acids. For each pair, identify the strongest type of interaction between these amino acids.
    1. valine and isoleucine
    2. asparagine and serine
    3. glutamic acid and arginine
    4. tryptophan and methionine

5. What level(s) of protein structure is(are) ordinarily disrupted in denaturation? What level(s) is(are) not?

6. Which class of proteins is more easily denatured—fibrous or globular?

This page titled 2.6: Amino Acids and Proteins (Exercises) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Allison Soult.

CHAPTER OVERVIEW

3: Methods of Protein Purification and Characterization

3.1: Protein Purification

3.2: Cell Disruption

3.3:Cell Fractionation and Centrifugation

3.4: Chromatography

3.4.1. Affinity Chromatography

3.4.2. Gel Exclusion Chromatography

3.4.3. Ion Exchange Chromatography

3.5: Electrophoresis

3.5.1: Blotting

3.5.2: Exercises

3: Methods of Protein Purification and Characterization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.1: Protein Purification

Assays, Specific Activity, Initial Fractionation

A successful protein purification procedure can be nothing short of amazing. Whether you are starting off with a recombinant protein which is produced in *E. coli*, or trying to isolate a protein from some mammalian tissue, you are typically starting with gram quantities of a complex mixture of protein, nucleic acids, polysaccharide, etc. from which you may have to extract milligram (or microgram!) quantities of desired protein at high purity, and hopefully with high yield.

The first step in any purification is the development of a **specific assay** for the protein of interest. The specific assay can be based upon some *unique characteristic* of the protein of interest

- Enzymatic activity
- Immunological activity
- Physical characteristics (e.g. molecular mass, spectroscopic properties, etc.)
- Biological activity
- Ideally, an assay should be
  - **Specific** (you don't want a false positive)
  - **rapid** (you don't want to wait a week for the results)
  - **sensitive** (you don't want to consume all your sample in order to assay it)
  - **quantitative** (you need an accurate way to measure the quantity of your protein at each step in the purification)

Protein purification can be thought of as a series of fractionation steps designed so that:

- The protein of interest is found almost exclusively in one fraction (and with good yield)
- A significant amount of the contaminants can be found in a different fraction

During purification you will need to monitor several parameters, including:

1. **Total sample volume**
2. **Total sample protein** (can be estimated by A_{280}; 1.4 ~ 1.0 mg/ml)
3. **Units of activity of desired protein** (based on specific assay)

This basic information will allow you to keep track of the following information during each step of purification:

1. **% yield** for each purification step
2. **Specific activity** of the desired protein (units/mg total protein)
3. **Purification enhancement** of each step (e.g. "3.5x purification)

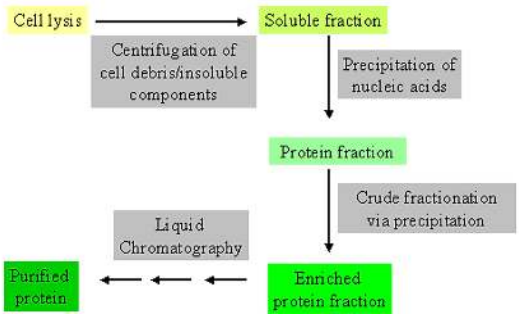
In designing a purification scheme you typically have to balance **purification** with **yield**.

- For example, it may be relatively straightforward to obtain 90% pure material with good yield.
- However, it may be difficult to improve that purity an additional few percentile with good yield.
- The planned application of the purified protein determines the target purity.
- If the protein is to be used to determine amino acid sequence information, maybe 90% is acceptable. However, if the material is to be used in clinical trials, 99.99+% may be the target purity.

Initial steps in purification

Figure 3.1.1: Purification steps

It is extremely helpful to have some information not only on the general physical and chemical characteristics of the protein you are trying to purify, but also on the *contaminating* components. For example, many *E. coli* proteins are generally *low molecular weight* (<50,000 Da) and somewhat *acidic* in isoelectric point. Usually the initial steps in purification make use of general physical and/or chemical differences between soluble proteins and other cell components. Soluble proteins can be separated from general cellular debris, and intact cells, by *centrifugation*. Thus, cells are physically disrupted (via homogenization or a cell press) to allow release of cell contents. This is then followed by centrifugation to separate generally soluble components from those which are insoluble.



Monitoring the Purification Process

- There are several criteria. One criteria is that we cannot improve upon the *specific activity* of our sample. This value refers to the *functional activity of our sample in relationship to the total protein concentration of the sample*.
- In the initial stages of purification this value will be low (not much activity in relationship to the total amount of protein).
- This value will increase after each purification step as we *remove* other proteins from the sample.
- At some point the specific activity will *plateau*, and by definition, *if it is pure we cannot increase the specific activity*.
- There may be a published value for the specific activity which we can compare ours to.

Also, each step of the purification should be monitored by gel electrophoresis.

The following chart represents the typical data one would monitor during a purification:

Table 3.1: *Monitoring the Purification Process*

Step	Total protein (mg)	Total activity (units)	Specific activity (units/mg)	Purification	% Yield
Crude cell lysate	5500	6600	1.2		
30-70% Ammonium sulfate cut	1020	5910	5.8	4.8	89.5
DEAE Sephadex pool	187	5070	27.1	4.7	85.8
CM Sephadex pool	102	4420	43.3	1.6	87.2
Phenyl Sepharose pool	56	3930	70.2	1.6	88.9
Gel Filtration pool	32	2970	92.8	1.3	75.6
Affinity resin type #1 pool	5.8	2520	434.5	4.7	84.8
Affinity resin type #2 pool	5.3	2390	450.9	1.0	94.8
Total purification	**376**				
Total yield (%)	**36**				

This page titled 3.1: Protein Purification is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Michael Blaber.

3.2: 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.

This page titled 3.2: Cell Disruption is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

- **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.

3.3:Cell Fractionation and Centrifugation

Fractionation of samples typically starts with centrifugation. Using a centrifuge, one can remove cell debris, and fractionate organelles, and cytoplasm. For example, nuclei, being relatively large, can be spun down at fairly low speeds. Once nuclei have been sedimented, the remaining solution, or supernatant, can be centrifuged at higher speeds to obtain the smaller organelles, like mitochondria. Each of these fractions will contain a subset of the molecules in the cell.

Cells are disrupted in a homogenizer and the resulting mixture, called the homogenate, is centrifuged in a step-by-step fashion of increasing centrifugal force. The denser material will form a pellet at lower centrifugal force than will the less-dense material. The isolated fractions can be used for further purification.

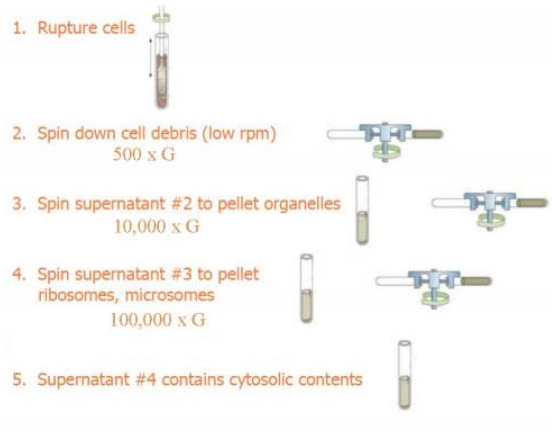


Figure 3.3.1: *Fractionation of samples typically starts with centrifugation. Using a centrifuge, one can remove cell debris, and fractionate organelles, and cytoplasm.*

This page titled 3.3:Cell Fractionation and Centrifugation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

3.4: Chromatography

"Chroma" refers to color and "graphy" refers to writing.

Chromatography is a method by which a mixture is separated by distributing its components between two phases. The stationary phase remains fixed in place while the mobile phase carries the components of the mixture through the medium being used. The stationary phase acts as a constraint on many of the components in a mixture, slowing them down to move slower than the mobile phase. The movement of the components in the mobile phase is controlled by the significance of their interactions with the mobile and/or stationary phases. Because of the differences in factors such as the solubility of certain components in the mobile phase and the strength of their affinities for the stationary phase, some components will move faster than others, thus facilitating the separation of the components within that mixture.

Theory

The distribution of a solute between the mobile and stationary phases in chromatography is described by κ, the partition coefficient, defined by:

$$\kappa = \frac{C_s}{C_m}$$

where Cs is the concentration of solute in the stationary phase and Cm is the concentration of the solute in the mobile phase. The mobile phase serves to carry the sample molecules through the chromatographic column. During the sample molecules transportation through the column, each analyte is retained according to that compound's characteristic affinity for the stationary phase. The time that passes between the sample injection and peak maximum is called the **retention time**. The area underneath each peak is proportional to the amount of co responding analyte in solution.

Retention Time

The retention time, tR, is given in seconds by:

$$t_R = t_S + t_M$$

where tS is the time the analyte spends in the stationary phase and tM is the time spent in the mobile phase. tM is often referred to as the dead, or void time, as all components spend tM in the mobile phase.

An example you can do at home involves a coffee filter (stationary phase) and water (mobile phase) in the separation and analysis of the dye in a water soluble marker pen (the mixture of compounds). If you draw a line with a purple marker on a strip of the filter paper, and place the bottom in a dish of water, the water will wick up the paper, and the differential affinity of the dye inks for the solid support (cellulose) results in differential migration. In the example below, the red ink has some affinity for the paper, but the blue ink does not (and migrates with the front edge of the mobile phase)

Figure 3.4.1: *Chromatography of purple pen*

Chromatography in biochemistry typically utilizes not paper, but beads of polysaccharide (often chemically derivatized) packed into a column, as the solid support. The solid support is often called the chromatography "resin". Some common types of chromatographic resins include:

- Ion exchange

- Affinity
- Hydrophobic
- Gel filtration

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

3.4.1. Affinity Chromatography

Affinity chromatography is another powerful technique of purifying proteins. This technique takes advantage of the high affinity of many proteins for specific chemical groups. Biological macromolecules, such as enzymes and other proteins, interact with other molecules with high specificity through several different types of bonds and interaction. Such interactions include hydrogen bonding, ionic interaction, disulfide bridges, hydrophobic interaction, and more. The high selectivity of affinity chromatography is caused by allowing the desired molecule to interact with the stationary phase and be bound within the column in order to be separated from the undesired material which will not interact and elute first. The molecules no longer needed are first washed away with a buffer while the desired proteins are let go in the presence of the eluting solvent (of higher salt concentration). This process creates a competitive interaction between the desired protein and the immobilized stationary molecules, which eventually lets the now highly purified proteins be released.

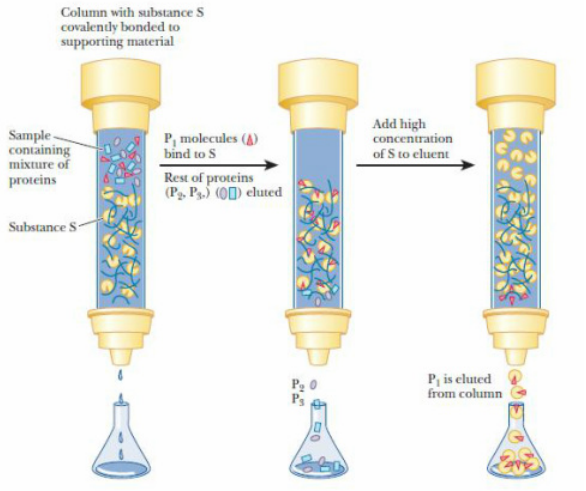
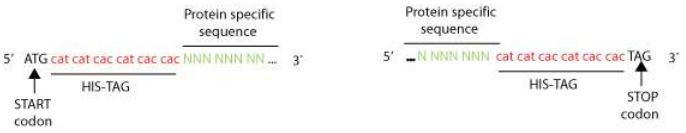


Figure 3.4.1.1: *Affinity chromatography is a powerful technique of purifying proteins.*

Purifying Proteins by Affinity Tag

Protein tags are peptide sequences genetically grafted onto a recombinant protein. Often these tags are removable by chemical agents or by enzymatic means, such as proteolysis or intein splicing. Tags are attached to proteins for various purposes.

Affinity tags are appended to proteins so that they can be purified from their crude biological source using an affinity technique. These include chitin binding protein (CBP), maltose binding protein (MBP), and glutathione-S-transferase (GST). The poly (His) tag is a widely-used protein tag; it binds to metal matrices.



Adding Polyhistidine Tags

This is an example of a primer designed to add a 6xHis-tag using PCR. Eighteen bases coding six histidines are inserted right after the START codon or right before the STOP codon. At least 16 bases specific to the gene of interest are needed next to the His-tag. With 6 His, the protein will have an added 1 kDa of molecular weight. Oftentimes, a linker (such as gly-gly-gly or gly-ser-gly) is placed between the protein of interest and the 6 His tag. This is to prevent the polyhistidine tag from affecting the activity of the protein being tagged.

Solubilization tags are used, especially for recombinant proteins expressed in chaperone-deficient species such as E. coli, so as to assist in the proper folding in proteins and keep them from precipitating. These include thioredoxin (TRX) and poly (NANP). Some affinity tags have a dual role as a solubilization agent, such as MBP and GST.

Chromatography tags are used to alter chromatographic properties of the protein to afford different resolution across a particular separation technique. These often consist of polyanionic amino acids, such as FLAG-tag.

Epitope tags are short peptide sequences which are chosen because high-affinity antibodies can be reliably produced in many different species. These are usually derived from viral genes, which explain their high immunoreactivity. Epitope tags include V5-tag, c-myc-tag, and HA-tag. These tags are particularly useful for western blotting, immunofluorescence and immunoprecipitation experiments, although they also find use in antibody purification.

Fluorescence tags are used to give visual readout on a protein. GFP and its variants are the most commonly used fluorescence tags. More advanced applications of GFP include using it as a folding reporter (fluorescent if folded, colorless if not).

Protein tags are also useful for specific enzymatic modification (such as biotin ligase tags) and chemical modification (FlAsH) tag. Often tags are combined to produce multifunctional modifications of the protein. However, with the addition of each tag comes the risk that the native function of the protein may be abolished or compromised by interactions with the tag.

Examples of peptide tags include

- AviTag, a peptide allowing biotinylation by the enzyme BirA and so the protein can be isolated by streptavidin (GLNDIFEAQKIEWHE)
- Calmodulin-tag, a peptide bound by the protein calmodulin (KRRWKKNFIAVSAANRFKKISSSGAL)
- FLAG-tag, a peptide recognized by an antibody (DYKDDDDK)
- HA-tag, a peptide recognized by an antibody (YPYDVPDYA)
- His-tag, 5-10 histidines bound by a nickel or cobalt chelate (HHHHHH)
- Myc-tag, a short peptide recognized by an antibody (EQKLISEEDL)
- S-tag (KETAAAKFERQHMDS)
- SBP-tag, a peptide which binds to streptavidin (MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP)
- Softag 1, for mammalian expression (SLAELLNAGLGGS)
- Softag 3, for prokaryotic expression (TQDPSRVG)
- V5 tag, a peptide recognized by an antibody (GKPIPNPLLGLDST)
- Xpress tag (DLYDDDDK)

Examples of protein tags include

- BCCP (Biotin Carboxyl Carrier Protein), a protein domain recognized by streptavidin
- Glutathione-S-transferase-tag, a protein which binds to immobilized glutathione
- Green fluorescent protein-tag, a protein which is spontaneously fluorescent and can be bound by nanobodies
- Maltose binding protein-tag, a protein which binds to amylose agarose
- Nus-tag
- Strep-tag, a peptide which binds to streptavidin or the modified streptavidin called streptactin (Strep-tag II: WSHPQFEK)
- Thioredoxin-tag

This page titled 3.4.1. Affinity Chromatography is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

3.4.2. 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.

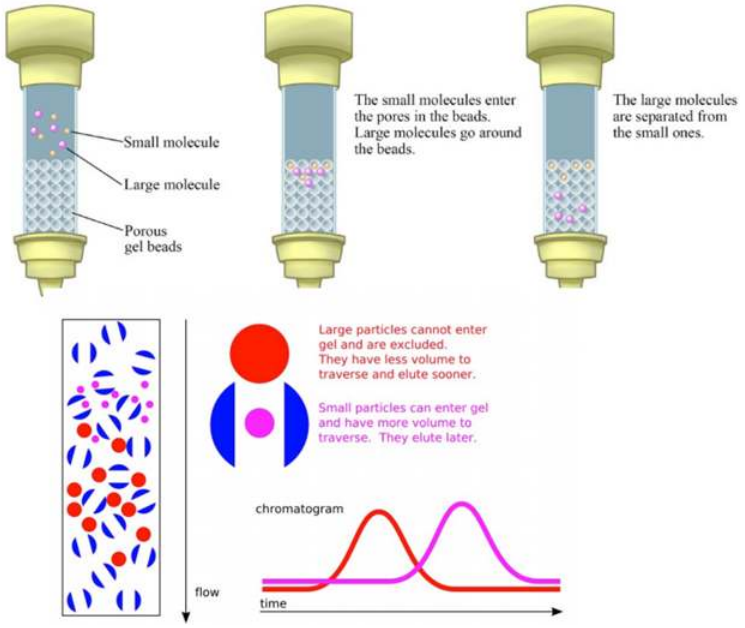


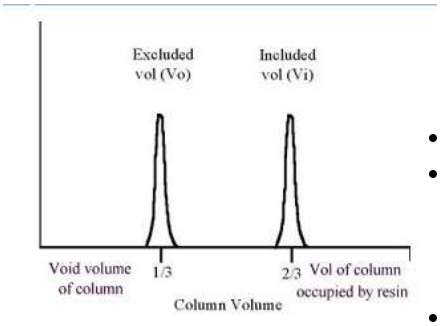
Figure 3.4.2.1: *Gel Exclusion Chromatography*

- Gel filtration does not rely on any chemical interaction with the protein, rather it is based on a **physical** property of the protein - that being the *effective molecular radius* (which relates to mass for most typical globular proteins).
- Gel filtration resin can be thought of as beads which contain pores of a defined size range.
- Large proteins which cannot enter these pores pass around the *outside* of the beads.
- Smaller proteins which can enter the pores of the beads have a longer, tortuous path before they exit the bead.
- Thus, a sample of proteins passing through a gel filtration column will *separate based on molecular size*: The big ones will elute first and the smallest ones will elute last (and "middle" sized proteins will elute in the middle).
- If your protein is unusually "small" or "large" *in comparison to contaminating proteins* then gel filtration may work quite well.

Where will a protein elute in a gel filtration experiment?

- There are two extremes in the separation profile of a gel filtration column.
- There is a critical molecular mass (large mass) which will be ***completely excluded*** from the gel filtration beads. All solutes in the sample which are equal to, or larger, than this critical size will behave identically: they will all eluted in the excluded volume of the column
- There is a critical molecular mass (small mass) which will be ***completely included*** within the pores of the gel filtration beads. All solutes in the sample which are equal to, or smaller, than this critical size will behave identically: they will all eluted in the included volume of the column
- Solutes between these two ranges of molecular mass will elute between the excluded and included volumes

Figure 3.4.2.2: Excluded vs. included volume



As a general rule of thumb, the excluded volume (Vo) is approximately equal to one third of the column volume, the included volume is approximately equal to two thirds of the column volume

- In gel filtration the resolution is a function of column length (the longer the better)
- However, one drawback is related to the maximum sample volume which can be loaded. The larger the volume of sample loaded, the more the overlap between separated peaks. Generally speaking, the sample size one can load is limited to about *3-5% of the total column volume*.
- Thus, gel filtration is best saved for *the end stages of a purification* ,when the sample can be readily concentrated to a small volume.
- Gel filtration can also be used to remove salts from the sample, due to its ability to separate "small" from "large" components.
- Finally, gel filtration can be among the most "gentle" purification methods due to the lack of chemical interaction with the resin.

This page titled 3.4.2. Gel Exclusion Chromatography is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

3.4.3. 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, in **cation exchange column**, the chemical groups attached to the beads are negatively charged groups and 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.

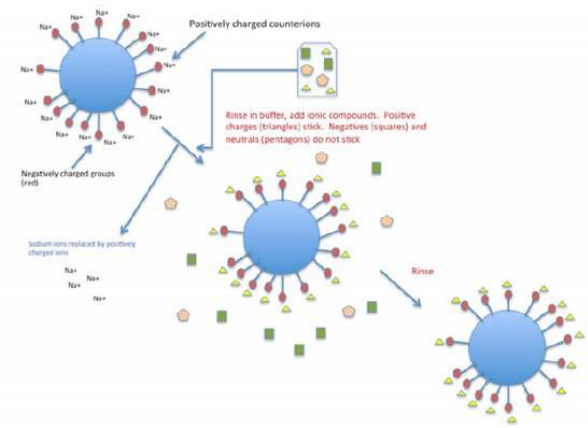
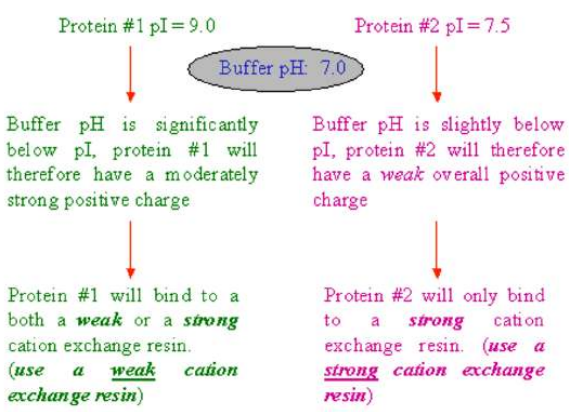


Figure 3.4.3.1: Cation exchange chromatography

Usually, samples are loaded under *low ionic strength* conditions and bound material is eluted using either a step or gradient elution of buffer with *higher ionic strength*.

Generally speaking, a protein will bind to a cation exchange resin if the buffer pH is *lower than the isoelectric point (pI)* of the protein, and will bind to an anion exchange resin if the pH is *higher than the pI*.



Knowledge of the pI of the protein is therefore helpful in designing a purification protocol using ion exchange resins (however, you can always simply try different resins to see which works best).

Elution of proteins from ion exchange resins

- Proteins bound to ion exchange resins are bound via non-covalent ionic (salt-bridge) interactions. We can compete for these ionic binding sites on the resin with other ionic groups, namely, ***salts***
- There are two general types of methods when eluting with a salt solution: 1. **Gradient elution** and 2. **Step elution**
- A gradient elution refers to a smooth transition of salt concentration (from low to high) in the elution buffer. Weakly binding proteins elute first, and stronger binding proteins elute last (i.e. they require higher salt concentrations in the buffer to compete them off the column)
- A gradient salt concentration can be made using a **gradient maker**. In its simplest form, this consists of two containers (***must be the same shape***) connected by a siphon (or tube at the bottom). One container contains the low salt buffer, and the other

contains high salt buffer. The buffer is withdrawn from the low salt container:

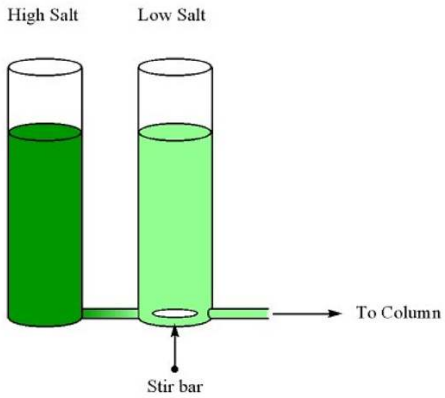


Figure 3.4.3.2: Gradient maker

This will produce a linear gradient from low to high salt concentrations over the total volume of the gradient

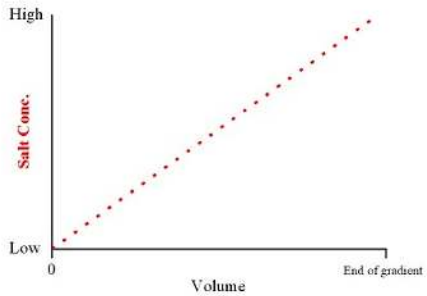


Figure 3.4.3.3: Salt concentration and volume

If we know the concentration range of salt over which a protein of interest will elute we can simply elute with a buffer containing that concentration of salt. This is known as a **step elution**.

Step elutions are generally faster to run, and elute the protein in a smaller overall volume than with gradient elutions. They generally work best when contaminants elute at a significantly different salt concentration than the protein of interest.

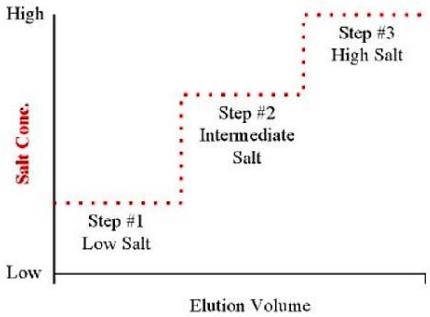


Figure 3.4.3.4: Step elution

Note that after ion exchange chromatography the protein of interest will be in a buffer with a potentially high salt concentration. This must be taken into account before proceeding with the next step in the purification scheme.

Dialysis

- After an ammonium sulfate precipitation step, or an ion exchange chromatography step, the protein of interest may be in a high salt buffer. This may be undesirable for several reasons. How do we get rid of salt in our sample?
- One of the most common methods is that of **dialysis**
- The method of dialysis makes use of **semi-permeable membranes**. In the simplest example, this membrane is manufactured in the form of tubing (looking much like a sausage casing)
- The main feature of this membrane is that it is porous. However, the pore size is such that while small salt ions can freely pass through the membrane, larger protein molecules cannot (i.e. they are retained). Thus, dialysis membranes are characterized by the molecular mass of the smallest typical globular protein which it will retain.
- This is commonly referred to as the **cutoff** of the tubing (e.g. Spectrapore #6 dialysis tubing has a cutoff of 1,000 Daltons, meaning that a 1,000 Dalton protein will be retained by the tubing but that smaller molecular mass solutes will pass through the tubing)
- Dialysis proceeds by placing a high salt sample in dialysis tubing (i.e. the dialysis "bag") and putting it into the desired low salt buffer.

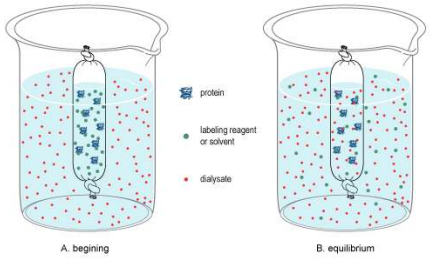


Figure 3.4.3.5: Dialysis (https://en.wikipedia.org/wiki/Dialysis_(biochemistry)

Over time the concentration of low molecular mass solutes within the bag, and in the low salt buffer, will come achieve equilibrium. In practical terms (for the above case) salt molecules will diffuse out of the bag into the low salt buffer.

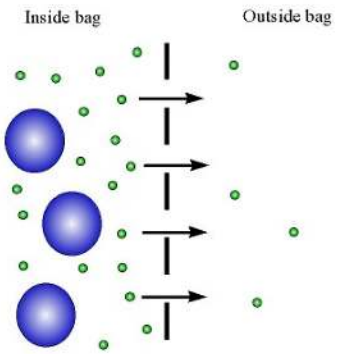


Figure 3.4.3.6: Salt diffusion

Concentration

- What if our protein sample is too dilute for our needs? How can we concentrate our samples?
- One common method is, again, to use a semi-permeable membrane for this purpose.
- A very simple method is to place our sample in a dialysis bag and coat it with a high molecular weight solute which can readily be dissolved by the buffer.
- In another variation, the semi-permeable membrane is manufactured into a flat disk and placed at the bottom of a container which holds our sample. In one method the container is pressurized and forces buffer out of the container (protein is retained and is concentrated). In another method, the vessel is centrifuged and the centripetal force achieves the same goal as pressure in the previous example.

For both dialysis and concentration, it is essential that the membrane does not interact with the protein (i.e. has no affinity for, and will not bind, the protein)

Contributors

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

This page titled 3.4.3. Ion Exchange Chromatography is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

3.5: Electrophoresis

SDS-PAGE

Like DNA and RNA, proteins are large macromolecules. Proteins, however, vary 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).

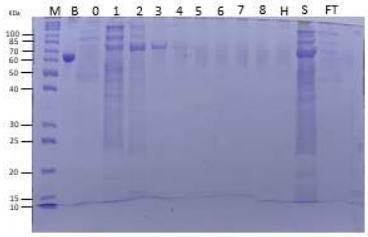


Figure 3.5.3: *SDS-PAGE separation of proteins.*

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.

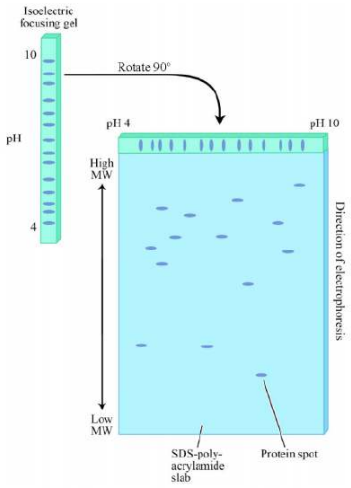


Figure 3.5.4: *2D Gel Electrophoresis*

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 left 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.

Contributors

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

This page titled 3.5: Electrophoresis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

3.5.1: Blotting

Blotting provides a means of identifying specific molecules out of a mixture. It employs three main steps. First, the mixture of molecules is separated by gel electrophoresis. The mixture could be DNA (Southern Blot), RNA (Northern Blot), or protein (Western Blot) and the gel could be agarose (for DNA/RNA) or polyacrylamide (for protein). Second, after the gel run is complete, the proteins or nucleic acids in the gel are transferred out of the gel onto a membrane/paper that physically binds to the molecules. This “blot", as it is called, has an imprint of the bands of nucleic acid or protein that were in the gel (see figure at left). The transfer can be accomplished by diffusion or by using an electrical current to move the molecules from the gel onto the membrane. The membrane may be treated to covalently link the bands to the surface of the blot. Last, a visualizing agent specific for the molecule of interest in the mixture is added to the membrane. For DNA/RNA, that might be a complementary nucleic acid sequence that is labeled in some fashion (radioactivity or dye). For a protein, it would typically involve an antibody that specifically binds to the protein of interest. The bound antibody can then be targeted by another antibody specific for the first antibody. The secondary antibody is usually linked to an enzyme which, in the presence of the right reagent, catalyzes a reaction that produces a signal (color or light) indicating where the antibody is bound. If the molecule of interest is in the original mixture, it will “light" up and reveal itself.

Western Blots

In a western blot procedure, proteins are first separated on an SDS-PAGE gel and then transferred to a membrane. This membrane replica is treated with antibodies that specifically recognize a protein or epitope of interest. Additional processing steps generate a signal at the position of the bound antibody. Between the steps, various washes are done to increase the signal-to-noise ratio on the final, developed blot. The major steps involved in a typical western blot are as follows:

- Electrophoretic transfer of proteins from an SDS-PAGE gel to a membrane
- Blocking of nonspecific protein binding sites on transfer membranes
- Incubation of the membrane with a primary antibody specific for the epitope of interest
- Incubation with a secondary antibody that recognizes primary antibodies
- Visualization of bound antibodies

Electrophoretic transfer of proteins from an SDS-PAGE gel to a membrane

The first step in a western blot is to generate a replica of the SDS-PAGE gel by transferring proteins electrophoretically to a synthetic membrane with a high protein binding capacity. The membranes made of polyvinylidine fluoride (PVDF), a kind of plastic, are hydrophobic and the dry membranes do not wet properly with water. Therefore, PVDF membranes are first wet with methanol, then rinsed with deionized water, and finally rinsed with transfer buffer. They must not be allowed to dry out during the transfer and immunoblot procedures. If they do dry out, they must be re-wet with methanol and rinsed with water before proceeding.

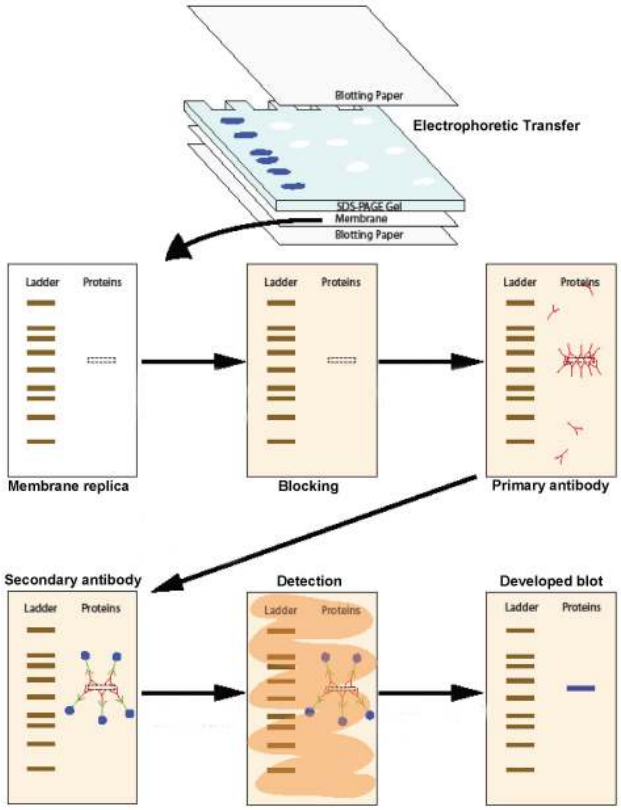


Figure 3.1.1.3: *The major steps in a typical western blot.*

During the transfer process, the gel and membrane are placed directly against each other within a “sandwich” of pre-wet filter papers and foam pads. During the electrophoretic transfer, current should flow evenly across the entire surface area of the gel. It is important, therefore, that air bubbles are not trapped between the gel and membrane. After the electrophoretic transfer, which can be done in a few hours or overnight with reduced voltage, the membrane replica with the transferred proteins can be allowed to dry out and stored for later visualization with antibodies.

Blocking of non-specific protein binding sites on membranes

The transfer membranes used in western blots bind proteins nonspecifically. Before the membranes are incubated with specific (and expensive) antibodies, they must be pretreated with blocking solutions that contain high concentrations of abundant (and cheap) proteins to saturate non-specific binding sites. If the transfer membranes are not adequately blocked before the antibody is applied, the nonspecific sites on the membranes will absorb some of the antibodies, reducing the amount of antibody available to bind the

Primary antibody binding

Either polyclonal or monoclonal antibodies can be used as the primary antibody on western blots. Antibodies can be directed toward a naturally-occurring protein or toward an epitope attached to an overexpressed protein (as we are doing). Increasingly, researchers are using epitope-tagged proteins in their experiments, because antibodies against naturally- occurring proteins are expensive and time-consuming to prepare.

Secondary antibody binding

The secondary antibodies used in western blots are designed to bind the FC fragments of primary antibodies, taking advantage of cross-species differences in antibody sequences. Secondary antisera are generally prepared by injecting an animal with FC fragments of IgGs from a second species. The first animal recognizes the FC fragments as foreign antigens and produces antibodies that bind the FC fragments.

Visualization of bound antibody

In this final step, the western blot is incubated with substrates for the enzyme that has been conjugated to the secondary antibody.

Contributors

Template:ContribOOG

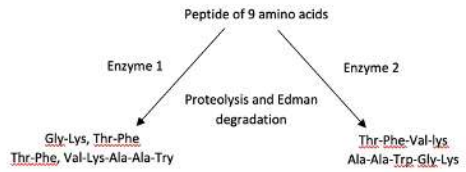
Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

Contributed by Clare M. O'Connor, Associate Professor Emeritus (Biology) at Boston College

This page titled 3.5.1: Blotting is shared under a CC BY-SA license and was authored, remixed, and/or curated by Todd Nickle and Isabelle Barrette-Ng.

3.5.2: Exercises

___ 1. What is the possible identity of enzyme 2?



a. papain

b. cyanogen bromide

c. pepsin

d. trypsin

e. chymotrypsin

___ 2. Which of the following techniques may be used to determine the tertiary structure of a protein?

a. X-ray crystallography

b. HPLC

c. ELISA

d. 1D proton NMR

e. mass spectrometry

___3. One liter of an aqueous solution contains 0.10 M aspartate at pH= pI, and 0.05 M NaOH is subsequently added. What is the pH after adding NaOH?

a. 3.1

b. 3.6

c. 4.1

d. 6.2

e. 8.0

___ 4. Which of the following amino acids would most likely be buried in the interior of a water-soluble protein?

a. Arg

b. Thr

c. Trp

d. Glu

e. His

___ 5. Which of the following described the quaternary structure of a protein?

a. helix or sheet stabilized by hydrogen bonds

b. formation of all four kinds of noncovalent bonds

c. a multi-peptide structure

d. requires disulfide bonds

e. none of the above are correct

___6. Somatostatin, a peptide hormone which inhibits the release of pituitary growth hormone, has the sequence:

___ 7. Which of the following agents is expected to alter the covalent structure of somatostatin?

a. SDS

b. heat to 100 ^{o}C

c. urea

d. cyanogen bromide

e. none of the above

___8. Which technique uses the attraction of the protein for a particular chemical group during protein purification?

a. affinity chromatography

b. gel-filtration chromatography

c. HPLC

d. SDS PAGE

e. salting out

___9. The AIGFRLKT was treated with the enzyme chymotrypsin. Which of the following could best separate the resulted peptides?

a. gel-filtration chromatography

b. affinity chromatography

c. ion-exchange chromatography

d. all of the above

e. none of the above

___10. You have purified a protein. Determination by gel-filtration chromatography yields 50 kDa, but mass spectrometry (MS) yielded a 25-kDa species. When the protein was per-treated with beta-mercaptoethanol followed MS determination, a single molecular species of 12.5 kDa showed up. Which of the following properly describes the structure of the protein? ***Hints: a dimer contains two subunits and a tetramer contains four subunits; each subunit is a separate polypeptide chain***

a. a dimer of 25 kDa subunits connected by disulfide bonds

b. a tetramer of 12.5 kDa subunits with all subunits linked to each other by disulfide bonds

c. a tetramer of 12.5 kDa subunits with two subunits forming 25 kDa dimers using disulfide bonds

d. a tetramer of 25 kDa subunits linked by disulfide bonds

___11. What are some of the modifications that proteins acquire?

a. cleavage of the protein

b. carboxylation

c. phosphorylation

d. all of the above

___12. Which of the following amino acids would most likely be found in the triple helix of collagen?

a. Asp

b. His

c. Phe

d. Gln

e. Gly

___13. Which atom(s) in proteins are regular hydrogen-bond acceptors?

a. carbon

b. oxygen

c. nitrogen

d. sulfur.

e. oxygen and nitrogen

14. You plan to purify and subsequently characterize a native protein.

a. List 4 techniques for purification of your protein (not denatured).

b. After purification, you would like to know the purity of your protein. List one technique you may use.

c. Now you would like to know the 3-D structure of the protein. List 2 techniques you may use.

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

CHAPTER OVERVIEW

4: Overview of Hemoglobin and Myoglobin

4.1: Myoglobin, Hemoglobin, and their Ligands
4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin
4.3: Exercises

4: Overview of Hemoglobin and Myoglobin is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.1: Myoglobin, Hemoglobin, and their Ligands

Hemoglobin:

- 1st. protein whose molecular weight was determined
- 1st protein to be assigned a specific function - dioxygen transport
- has a prosthetic group (non-amino acid) heme group (protoporphyrin IX with a ferrous ion) covalently attached to the protein. The heme group binds dioxygen.
- 1st protein in which a point mutation (single base pair charge) causes a single amino acid change in the protein, marking the start of molecular medicine
- 1st protein with high resolution x-ray structure
- theory for dioxygen binding explain control of enzyme activity
- the binding of dioxygen is regulated by binding of H^+, CO_2, and bisphosphoglcyerate which bind to sites (allosteric) distant from oxygen binding site.
- crystals of deoxy-Hb shatter on binding dioxygen, indicating significant conformational changes on binding.

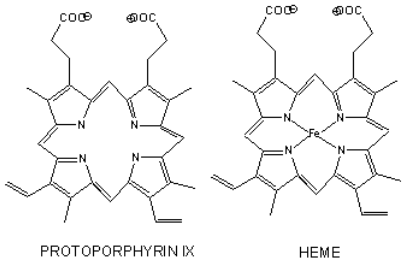


Figure 4.1.1: *The heme group contains protoporphyrin IX, with four tetrapyrrole rings linked by methene bridges. Attached to the tetrapyrrole structure are four methyl, two vinyl, and two proprionate groups. These can be arranged in 15 ways, only one (IX) occurs in biological systems.*

The heme fits into a hydrophobic crevice in the proteins with the proprionate groups exposed to solvent. The Fe^{2+} ion is coordinated to 4 N's on the 4 pyrrole rings, The 5th ligand is a supplied by proximal His (the 8th amino acid on helix F) of the protein. In the absence of dioxygen, the 6 ligand is missing. and the geometry of the complex is square pyramidal with the Fe above the plane of the heme ring. A distal His (E7) is on the opposite side of the heme ring, but too far to coordinate with the Fe. When dioxgen binds, it occupies the 6th coordination site and pulls the Fe into the plane of the ring, leading to octahedral geometry. CO, NO, and H_2S also bind to the 6th site, but with higher affinity than dioxygen, which can lead to CO poisoning. The distal His keeps these ligands (including dioxygen) bound in a bent, non-optimal geometry. This minimizes the chances of CO poisoning.

Myoglobin

- Myoglobin is a relatively small protein that contains 150 amino acids.
- Mb is extremely compact, and consists of 75% alpha helical structure.
- The interior amino acids are almost entirely nonpolar. The only polar amino acids found completely buried are the two His (proximal and distal) found at the active site of dioxygen binding.

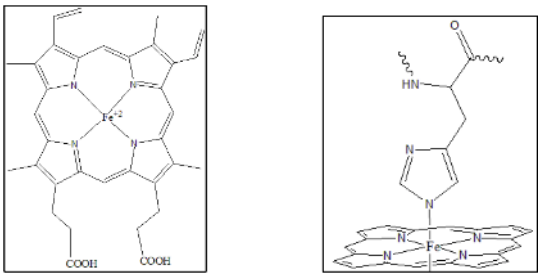


Figure 4.1.2: *The skeletal structure of the heme prosthetic group found within the structure of myoglobin. The porphyrin ring contains four pyrrole nitrogens bound to a ferrous (Fe(II)) ion center. There are six coordination sites in the Fe(II) ion; four are occupied by the pyrrole nitrogens, one is occupied by a proximal histidine, one site can be occupied by a dioxygen molecule (not shown).*

Difference between Hb and Mb

Hb is a tetramer of two α and two β subunits held together by IMF's (an example of quarternary protein structure), and 4 bound hemes, each of which can bind a dioxygen. In a fetus, two other subunits make up Hb (two zeta - ζand two epsilion - ε subunits - analogous to the two αand two β subunits, respectively). This changes in fetuses to two α and two γ subunits. Fetal Hb has higher affinity for dioxygen than adult Hb. Mb is a single polypeptide chains which has higher affinity for dioxygen than Hb.

The α and β chains are similar to Mb, which is unexpected since only 24 of 141 residues in the α and β chains of Hb are identical to amino acids in Mb. This suggests that different sequences can fold to similar structures. The globin fold of Mb and each chain of Hb is common to vertebrates and must be nature's design for dioxygen carriers. A comparison of the sequence of Hb from 60 species show much variability of amino acids, with only 9 identical amino founds. These must be important for structure/function. All internal changes are conservative (e.g. changing a nonpolar for a nonpolar amino acid). Not even Pro's are conserved, suggesting there are different ways to break helices. The two active site His are conserved, as is Gly B6 (required for a reverse turn).

This page titled 4.1: Myoglobin, Hemoglobin, and their Ligands is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin

Oxygen Transport

Many microorganisms and most animals obtain energy by respiration, the oxidation of organic or inorganic molecules by O_2. At 25°C, however, the concentration of dissolved oxygen in water in contact with air is only about 0.25 mM. Because of their high surface area-to-volume ratio, aerobic microorganisms can obtain enough oxygen for respiration by passive diffusion of O_2 through the cell membrane. As the size of an organism increases, however, its volume increases much more rapidly than its surface area, and the need for oxygen depends on its volume. Consequently, as a multicellular organism grows larger, its need for O_2 rapidly outstrips the supply available through diffusion. Unless a transport system is available to provide an adequate supply of oxygen for the interior cells, organisms that contain more than a few cells cannot exist. In addition, O_2 is such a powerful oxidant that the oxidation reactions used to obtain metabolic energy must be carefully controlled to avoid releasing so much heat that the water in the cell boils. Consequently, in higher-level organisms, the respiratory apparatus is located in internal compartments called mitochondria, which are the power plants of a cell. Oxygen must therefore be transported not only to a cell but also to the proper compartment within a cell.

Myoglobin and Hemoglobin

Myoglobin is a relatively small protein that contains 150 amino acids. The functional unit of myoglobin is an **iron–porphyrin complex** that is embedded in the protein (Figure 4.2.1). In myoglobin, the heme iron is five-coordinate, with only a single histidine imidazole ligand from the protein (called the proximal histidine because it is near the iron) in addition to the four nitrogen atoms of the porphyrin. A second histidine imidazole (the distal histidine because it is more distant from the iron) is located on the other side of the heme group, too far from the iron to be bonded to it. Consequently, the iron atom has a vacant coordination site, which is where O_2 binds.

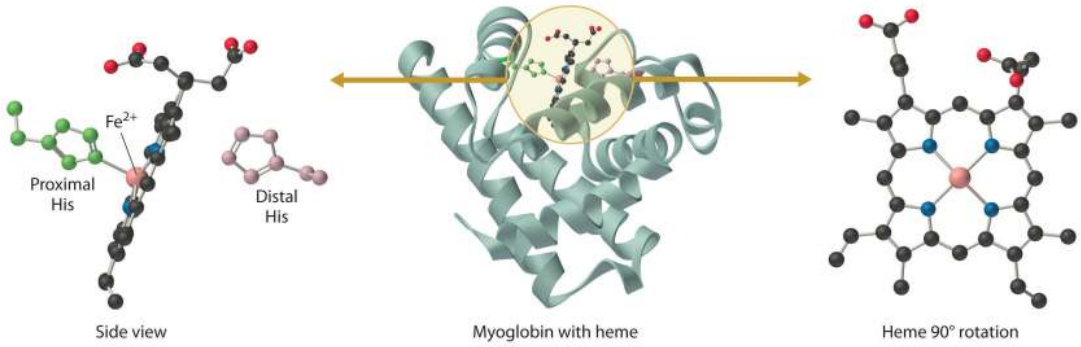


Figure 4.2.1: The Structure of Deoxymyoglobin, Showing the Heme Group. The iron in deoxymyoglobin is five-coordinate, with one histidine imidazole ligand from the protein. Oxygen binds at the vacant site on iron.

In the ferrous form (deoxymyoglobin), the iron is five-coordinate and high spin. Because high-spin Fe^{2+} is too large to fit into the "hole" in the center of the porphyrin, it is about 60 pm above the plane of the porphyrin. When O_2 binds to deoxymyoglobin to form oxymyoglobin, the iron is converted from five-coordinate (high spin) to six-coordinate (low spin; Figure 4.2.2). Because low-spin Fe^{2+} and Fe^{3+} are smaller than high-spin Fe^{2+}, the iron atom moves into the plane of the porphyrin ring to form an octahedral complex. The O_2 pressure at which half of the molecules in a solution of myoglobin are bound to O_2 ($P_{1/2}$) is about 1 mm Hg (1.3×10^{-3} atm).

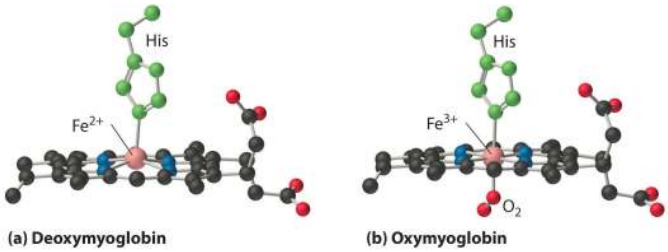


Figure 4.2.2: Oxygen Binding to Myoglobin and Hemoglobin. (a) The Fe^{2+} ion in deoxymyoglobin is high spin, which makes it too large to fit into the "hole" in the center of the porphyrin. (b) When O_2 binds to deoxymyoglobin, the iron is converted to low-spin Fe^{3+}, which is smaller, allowing the iron to move into the plane of the four nitrogen atoms of the porphyrin to form an octahedral complex.

Hemoglobin consists of two subunits of 141 amino acids and two subunits of 146 amino acids, both similar to myoglobin; it is called a tetramer because of its four subunits. Because hemoglobin has very different O_2-binding properties, however, it is not simply a "super myoglobin" that can carry four O_2 molecules simultaneously (one per heme group). The shape of the O_2-binding curve of myoglobin can be described mathematically by the following equilibrium:

$$MbO_2 \rightleftharpoons Mb + O_2 \tag{4.2.1}$$

$$K_{diss} = \frac{[Mb]\,[O_2]}{[MbO_2]} \tag{4.2.2}$$

The O_2-binding curve of hemoglobin is S shaped (Figure 4.2.3). As shown in the curves, at low oxygen pressures, the affinity of deoxyhemoglobin for O_2 is substantially lower than that of myoglobin, whereas at high O_2 pressures the two proteins have comparable O_2 affinities. The physiological consequences of unusual S-shaped O_2-binding curve of hemoglobin are enormous. In the lungs, where O_2 pressure is highest, the high oxygen affinity of deoxyhemoglobin allows it to be completely loaded with O_2, giving four O_2 molecules per hemoglobin. In the tissues, however, where the oxygen pressure is much lower, the decreased oxygen affinity of hemoglobin allows it to release O_2, resulting in a net transfer of oxygen to myoglobin.

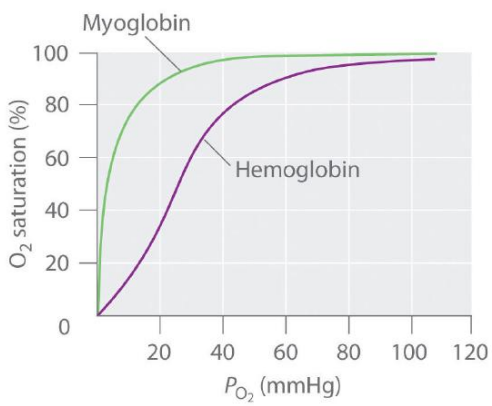


Figure 4.2.3: The O_2-Binding Curves of Myoglobin and Hemoglobin. Plots of Y (fractional saturation) vs L (pO2) are hyperbolic for Mb, but sigmoidal for Hb, suggesting cooperative binding of oxygen to Hb (binding of the first oxygen facilitates binding of second, etc).

The S-shaped O_2-binding curve of hemoglobin is due to a phenomenon called **cooperativity**, in which the affinity of one heme for O_2 depends on whether the other hemes are already bound to O_2. Cooperativity in hemoglobin requires an interaction between the four heme groups in the hemoglobin tetramer, even though they are more than 3000 pm apart, and depends on the change in structure of the heme group that occurs with oxygen binding. The structures of deoxyhemoglobin and oxyhemoglobin are slightly different, and as a result, deoxyhemoglobin has a much lower O_2 affinity than myoglobin, whereas the O_2 affinity of oxyhemoglobin is essentially identical to that of oxymyoglobin. Binding of the first two O_2 molecules to deoxyhemoglobin causes the overall structure of the protein to change to that of oxyhemoglobin; consequently, the last two heme groups have a much higher affinity for O_2 than the first two.

The affinity of Hb, but not of Mb, for dioxygen depends on pH. This is called the **Bohr effect**, after the father of Neils Bohr, who discovered it.

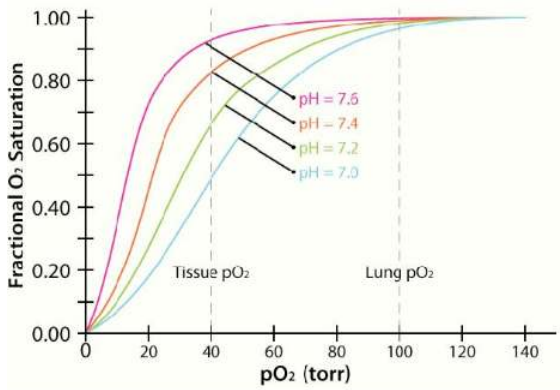


Figure 4.2.4: The Bohr Effect

Decreasing pH shifts the oxygen binding curves to the right (to decreased oxygen affinity). Increased $[H^+]$ will cause protonation of basic side chains. In the pH range for the Bohr effect, the mostly likely side chain to get protonated is His (pKa around 6), which then becomes charged. The mostly likely candidate for protonation is His 146 (on the β chain - CH_3) which can then form a salt bridge with Asp 94 of the β(FG1) chain. This salt bridge stabilizes the positive charge on the His and raises its pKa compared to the oxyHb state. Carbon dioxide binds covalently to the N-terminus to form a negatively charge carbamate which forms a salt bridge with Arg 141 on the alpha chain. BPG, a strongly negatively charged ligand, binds in a pocket lined with Lys 82, His 2, and His 143 (all on the beta chain). It fits into a cavity present between the β subunits of the Hb tetramer in the T state. Notice all these allosteric effectors lead to the formation of more salt bridges which stabilize the T or deoxy state. The central cavity where BPG binds between the β subunits become much smaller on oxygen binding and the shift to the oxy or R state. Hence BPG is extruded from the cavity.

The binding of H^+ and CO_2 helps shift the equilibrium to deoxyHb which faciliates dumping of oxygen to the tissue. It is in respiring tissues that CO_2 and H^+ levels are high. CO_2 is produced from the oxidation of glucose through glycolysis and the Krebs cycle. In addition, high levels of CO_2 increase H^+ levels through the following equilibrium:

$$\mathbf{H_2O + CO_2 \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-} \tag{4.2.3}$$

In addition, H^+ increases due to production of weak acids such as pyruvic acid in glycolysis .

Hb, by binding CO_2 and H^+, in addition to O_2, serves an additional function: it removes excess CO_2 and H^+ from the tissues where they build up. When deoxyHb with bound H^+ and CO_2 reaches the lungs, they leave as O_2 builds and deoxyHb is converted to oxyHb.

2,3-BPG

Another molecule favoring the release of oxygen by hemoglobin is 2,3- bisphosphoglycerate (also called 2,3-BPG or just BPG - Figure 4.2.5). Like protons and carbon dioxide, 2,3-BPG is produced by actively respiring tissues, as a byproduct of glucose metabolism. The 2,3-BPG molecule fits into the ‘hole of the donut’ of adult hemoglobin. 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.

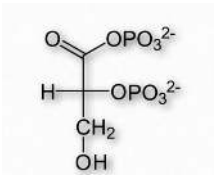


Figure 4.2.5: 2,3- bisphosphoglycerate

Carbon Monoxide

CO is a highly toxic gas without color and odor. It is commonly produced the partial combustion of carbon-containing compounds. It competes with oxygen for hemoglobin binding. Its binding affinity is ~ 200 fold tighter.

Summary

Hemoglobin is a protein found in red blood cells (RBCs) that is comprised of two alpha and two beta subunits that surround an iron-containing heme group. Oxygen readily binds this heme group. The ability of oxygen to bind increases as more oxygen molecules are bound to heme. Disease states and altered conditions in the body can affect the binding ability of oxygen, and increase or decrease its ability to dissociate from hemoglobin.

Carbon dioxide can be transported through the blood via three methods. It is dissolved directly in the blood, bound to plasma proteins or hemoglobin, or converted into bicarbonate. The majority of carbon dioxide is transported as part of the bicarbonate system. Carbon dioxide diffuses into red blood cells. Inside, carbonic anhydrase converts carbon dioxide into carbonic acid (H_2CO_3), which is subsequently hydrolyzed into bicarbonate (HCO_3^-) and H^+. The H^+ ion binds to hemoglobin in red blood cells, and bicarbonate is transported out of the red blood cells in exchange for a chloride ion. This is called the **chloride shift**. Bicarbonate leaves the red blood cells and enters the blood plasma. In the lungs, bicarbonate is transported back into the red blood cells in exchange for chloride. The H^+ dissociates from hemoglobin and combines with bicarbonate to form carbonic acid with the help of carbonic anhydrase, which further catalyzes the reaction to convert carbonic acid back into carbon dioxide and water. The carbon dioxide is then expelled from the lungs.

Q1: what are the differences between concerted model and sequential model?

https://bio.libretexts.org/Bookshelv...ng_-_Hill_Plot

Contributors

Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

Contributed by OpenStax, General Biology at OpenStax CNX

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.3: Exercises

Multiple Choice

1. Which of the following will NOT facilitate the transfer of oxygen to tissues?

1. decreased body temperature
2. decreased pH of the blood
3. increased carbon dioxide
4. increased exercise

2. The majority of carbon dioxide in the blood is transported by _____.

1. binding to hemoglobin
2. dissolution in the blood
3. conversion to bicarbonate
4. binding to plasma proteins

3. The majority of oxygen in the blood is transported by _____.

1. dissolution in the blood
2. being carried as bicarbonate ions
3. binding to blood plasma
4. binding to hemoglobin

Review Questions

1. **Hemoglobin** is able to use 90% of its potential oxygen-carrying capacity effectively. Under similar conditions, **myoglobin** would be able to use only 7% of its potential capacity. What accounts for this dramatic difference?

2. How does **2,3-BPG** affect oxygen affinity so significantly?

3. Explain the heterotropic regulation of **hemoglobin** by hydrogen ions and carbon dioxide increases the oxygen-transporting efficiency of this magnificent allosteric **protein**.

4. What would happen if no carbonic anhydrase were present in red blood cells?

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

CHAPTER OVERVIEW

5: Michaelis-Menten Enzyme Kinetics, Inhibitors, pH optima; Bi-Substrate Reactions

5.1: Catalytic Efficiency of Enzymes
5.2: Enzyme Parameters
5.3: Michaelis-Menten Kinetics
5.4: Enzyme Inhibition
5.5: Temperature, pH, and enzyme concentration on the rate of a reaction
5.6: Multi-Substrate Sequential Mechanisms
5.7: Double displacement reaction

5: Michaelis-Menten Enzyme Kinetics, Inhibitors, pH optima; Bi-Substrate Reactions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.1: Catalytic Efficiency of Enzymes

Introduction

An enzyme's active sites are usually composed of amino acid residues; depending on which amino acid residues are present, the specificity of the substrate can vary greatly. Depending on the pH level, the physical properties (mainly the electric charge) of an enzyme can change. A change in the electric charge can alter the interaction between the active site amino acid residues and the incoming substrate. With that said, the substrate can bind to the active site via hydrogen bonding or van der Waals forces. Once the substrate binds to the active site it forms an enzyme-substrate complex that is then involved in further chemical reactions.

In order for an enzyme to be active and be energetically favorable to allow a chemical reaction to proceed forward, a substrate must bind to an enzyme's "active site". An active site can be thought of as a lock and the substrate as a key; this is known as the **lock and key model**. A key (substrate) must be inserted and turned (chemical reaction), then the lock (enzyme) opens (production of products). Note that an enzyme might have more than one active site. Another theory on the active site-substrate relationship is the **induced fit theory**, which is quite opposite of the lock and key theory (where the active site is seemingly inflexible). In the induced fit theory, the active site of the enzyme is very flexible, and only changes its conformation when the substrate binds to it.

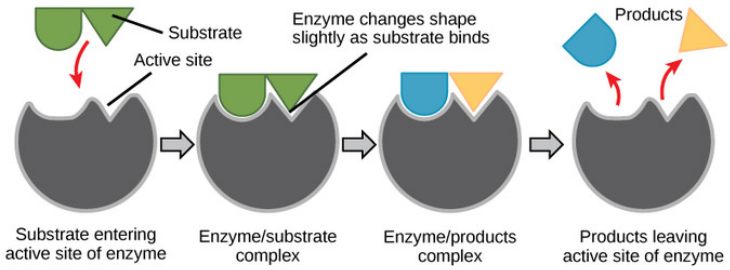


Figure 5.1.1: According to the induced fit model, both enzyme and substrate undergo dynamic conformational changes upon binding. The enzyme contorts the substrate into its transition state, thereby increasing the rate of the reaction.

Enzymes work as a catalyst by lowering the Gibbs free energy of activation of the enzyme-substrate complex. Below are two figures showing a basic enzymatic reaction with and without a catalyst.

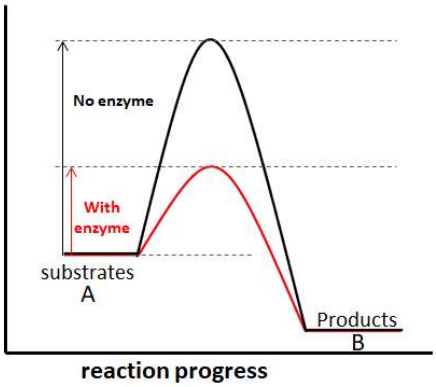


Figure 5.1.2: The energies of the stages of a chemical reaction. Without enzyme (Black line), substrates need a lot of activation energy to reach a transition state, which then decays into lower-energy products. When enzyme catalyzed (Red line), the enzyme binds the substrates (ES), then stabilizes the transition state ($ES^{\ddagger}$) to reduce the activation energy required to produce products (EP) which are finally released.

Notice the difference of the Gibbs energy with the presence of a catalyst; it is higher in value when a catalyst is absent. The Gibbs energy can tell us whether or not a reaction is spontaneous judging from its value. In this case $\Delta G^{\ddagger}$, the Gibbs free energy, can be conceptualized as the probability of a reaction occurring in nature without any interference. With a lower (more negative) Gibbs energy value, the probability of reaction is higher, and vice versa. Enzymes are here to increase the probability of the chemical reaction via many mechanisms, with one of the more prominent explanations being that it adopts a more favorable conformation.

Enzyme Active Site and Substrate Specificity

Enzymes bind with chemical reactants called substrates. There may be one or more substrates for each type of enzyme, depending on the particular chemical reaction. In some reactions, a single-reactant substrate is broken down into multiple products. In others, two substrates may come together to create one larger molecule. Two reactants might also enter a reaction, both become modified, and leave the reaction as two products.

The enzyme's active site binds to the substrate. Since enzymes are proteins, this site is composed of a unique combination of amino acid residues (side chains or R groups). Each amino acid residue can be large or small; weakly acidic or basic; hydrophilic or hydrophobic; and positively-charged, negatively-charged, or neutral. The positions, sequences, structures, and properties of these residues create a very specific chemical environment within the active site. A specific chemical substrate matches this site like a jigsaw puzzle piece and makes the enzyme specific to its substrate.

Active Sites and Environmental Conditions

Environmental conditions can affect an enzyme's active site and, therefore, the rate at which a chemical reaction can proceed. Increasing the environmental temperature generally increases reaction rates because the molecules are moving more quickly and are more likely to come into contact with each other.

However, increasing or decreasing the temperature outside of an optimal range can affect chemical bonds within the enzyme and change its shape. If the enzyme changes shape, the active site may no longer bind to the appropriate substrate and the rate of reaction will decrease. Dramatic changes to the temperature and pH will eventually cause enzymes to denature.

Enzyme-Substrate Complex

When an enzyme binds its substrate, it forms an enzyme-substrate complex. This complex lowers the activation energy of the reaction and promotes its rapid progression by providing certain ions or chemical groups that actually form covalent bonds with molecules as a necessary step of the reaction process. Enzymes also promote chemical reactions by bringing substrates together in an optimal orientation, lining up the atoms and bonds of one molecule with the atoms and bonds of the other molecule. This can contort the substrate molecules and facilitate bond-breaking. The active site of an enzyme also creates an ideal environment, such as a slightly acidic or non-polar environment, for the reaction to occur. The enzyme will always return to its original state at the completion of the reaction. One of the important properties of enzymes is that they remain ultimately unchanged by the reactions they catalyze. After an enzyme is done catalyzing a reaction, it releases its products (substrates).

5.1: Catalytic Efficiency of Enzymes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **Catalytic Efficiency of Enzymes** by LibreTexts is licensed CC BY-NC-SA 4.0.

5.2: Enzyme Parameters

Scientists spend a considerable amount of time characterizing enzymes. To understand how they do this and what the characterizations tell us, we must first understand a few parameters. Imagine I wished to study the reaction catalyzed by an enzyme I have just isolated. I would be interested to understand how fast the enzyme works and how much affinity the enzyme has for its substrate(s).

To perform this analysis, I would perform the following experiment. Into 20 different tubes, I would put 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. I would let the reaction proceed for a fixed, short amount of time and then I would measure the amount of product contained in each tube. For each reaction, I would determine the velocity of the reaction as the concentration of product found in each tube divided by the time. I would then plot the data on a graph using velocity on the Y-axis and the concentration of substrate on the X-axis.

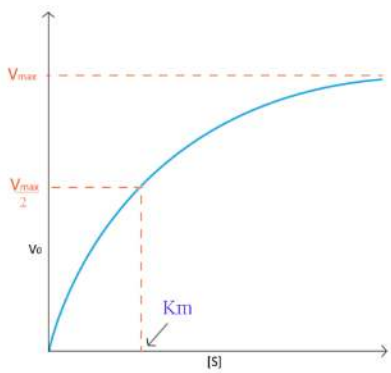


Figure 5.2.1: Velocity vs Substrate concentration plot

Typically, I would generate a curve like that shown on Figure 5.2.1. Notice how the velocity increase is almost linear in the tubes with the lowest amounts of substrate. This indicates that substrate is limiting and the enzyme converts it into product as soon as it can bind it. 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. What is happening is the enzyme is becoming saturated with substrate at higher concentrations of the latter. 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.

Turnover Number

On a plot of Velocity versus Substrate Concentration (V vs. $[S]$), the maximum velocity (known as $\mathrm{V_{max}}$) is the value on the Y axis that the curve asymptotically approaches. It should be noted that the value of $\mathrm{V_{max}}$ depends on the amount of enzyme used in a reaction. Double the amount of enzyme, double the $\mathrm{V_{max}}$. If one wanted to compare the velocities of two different enzymes, it would be necessary to use the same amounts of enzyme in the different reactions they catalyze. It is desirable to have a measure of velocity that is independent of enzyme concentration. For this, we define the value $\mathrm{K_{cat}}$, also known as the **turnover number**. Mathematically,

$$\mathrm{K_{cat}} = \frac{\mathrm{V_{max}}}{[\mathrm{Enzyme}]} \tag{4.7.1}$$

To determine $\mathrm{K_{cat}}$, one must obviously know the $\mathrm{V_{max}}$ at a particular concentration of enzyme, but the beauty of the term is that it is a measure of velocity independent of enzyme concentration, thanks to the term in the denominator. Kcat is thus a constant for an enzyme under given conditions. The units of $\mathrm{K_{cat}}$ are $\mathrm{time^{-1}}$. An example would be 35/second. This would mean that each molecule of enzyme is catalyzing the formation of 35 molecules of product every second. While that might seem like a high value, there are enzymes known (carbonic anhydrase, for example) that have $\mathrm{K_{cat}}$ values of 10^6/second. This astonishing number illustrates clearly why enzymes seem almost magical in their action.

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 5.2.2: Common turnover numbers

Michaelis constant

Another parameter of an enzyme that is useful is known as $\mathrm{K_M}$, the **Michaelis constant**. What it measures, in simple terms, is the affinity an enzyme has for its substrate. Affinities of enzymes for substrates vary considerably, so knowing $\mathrm{K_M}$ helps us to understand how well an enzyme is suited to the substrate being used. Measurement of KM depends on the measurement of V_{max}. On a V vs. [S] plot, $\mathrm{K_M}$ is determined as the x value that give $\mathrm{V_{max}/2}$. A common mistake students make in describing V_{max} is saying that $\mathrm{K_M = V_{max}/2}$. This is, of course not true. $\mathrm{K_M}$ is a substrate concentration and is the amount of substrate it takes for an enzyme to reach $V_{max}/2$. On the other hand $V_{max}/2$ is a velocity and is nothing more than that. The value of $\mathrm{K_M}$ is inversely related to the affinity of the enzyme for its substrate. High values of $\mathrm{K_M}$ correspond to low enzyme affinity for substrate (it takes more substrate to get to V_{max}). Low $\mathrm{K_M}$ values for an enzyme correspond to high affinity for substrate.

5.2: Enzyme Parameters is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.8: Enzyme Parameters** by Kevin Ahern & Indira Rajagopal is licensed CC BY-NC-SA 4.0. Original source: https://biochem.oregonstate.edu/undergraduate/educational-resources.

5.3: Michaelis-Menten Kinetics

Introduction

The general reaction scheme of an enzyme-catalyzed reaction is as follows:

$$E + S \xrightarrow{k_1} [ES] \xrightarrow{k_2} E + P \tag{5.3.1}$$

The enzyme interacts with the substrate by binding to its active site to form the enzyme-substrate complex, ES. That reaction is followed by the decomposition of ES to regenerate the free enzyme, E, and the new product, P.

To begin our discussion of enzyme kinetics, let's define the number of moles of product (P) formed per time as *V*. The variable, *V*, is also referred to as the rate of catalysis of an enzyme. For different enzymes, *V* varies with the concentration of the substrate, S. At low S, *V* is linearly proportional to S, but when S is high relative to the amount of total enzyme, *V* is independent of S. **Concentrations is important in determining the initial rate of an enzyme-catalyzed reaction.**

To understand Michaelis-Menten Kinetics, we will use the general enzyme reaction scheme shown below, which includes the back reactions in addition the the forward reactions:

$$E + S \xrightarrow{k_1} [ES] \xrightarrow{k_2} E + P \tag{5.3.2}$$

$$E + S \xleftarrow{k_1} [ES] \xleftarrow{k_2} E + P \tag{5.3.3}$$

The table below defines each of the rate constants in the above scheme.

***Table 1**: Model parameters*

Rate Constant	Reaction
k_1	The binding of the enzyme to the substrate forming the enzyme substrate complex.
k_2	Catalytic rate; the catalysis reaction producing the final reaction product and regenerating the free enzyme. This is the rate limiting step.
k_1	The dissociation of the enzyme-substrate complex to free enzyme and substrate .
k_2	The reverse reaction of catalysis.

Substrate Complex

$$E + S \xrightarrow{k_1} ES \; v = k_1[E][S]$$

$$ES \xrightarrow{k_1} E + S \; v = k_1[ES]$$

$$ES \xrightarrow{k_2} E + P \; v = k_2[ES]$$

$$E + P \xrightarrow{k_2} ES \; v = k_2[E][P] = 0$$

The ES complex is formed by combining enzyme E with substrate S at rate constant k_1. The ES complex can either dissociate to form E_F (free enzyme) and S, or form product P at rate constant k_2 and k-1, respectively. The velocity equation can be derived in either of the 2 methods that follow.

Method 1: The Rapid Equilibrium Approximation

E, S, and the ES complex can equilibrate very rapidly. The instantaneous velocity is the catalytic rate that is equal to the product of ES concentration and k_2 the catalytic rate constant.

$$v = k_2[ES] \tag{5.3.4}$$

The total enzyme concentration (E_T) is equal to the concentration of free enzyme E (E_F) plus the concentration of the bound enzyme in ES complex:

$$[E]_T = [E_F] + [ES] \tag{5.3.5}$$

$$K_s = \frac{k_1}{k_1} = \frac{[E][S]}{[ES]} \tag{5.3.6}$$

$$K_s \frac{([E_o] - [ES])[S]}{[ES]} \tag{5.3.7}$$

$$[ES] = \frac{[E_o][S]}{K_s + [S]} \tag{5.3.8}$$

$$v_o = \left(\frac{dP}{dt}\right)_o = k_2[ES] \tag{5.3.9}$$

$$v = \left(\frac{dP}{dt}\right)_o = k_2[E_o] = V_{max}[E_o][S]K_s + [S]$$

At high substrate concentrations, $[S] >> K_s$ we get:

$$v = \left(\frac{dP}{dt}\right)_o = k_2[E_o] = V_{max}$$

Method 2: The Steady-State Approximation

The figure below shows the relatively low and constant concentration of the enzyme-substrate complex due to the complex's slow formation and rapid consumption. Note the falling substrate concentration and the rising product concentration.

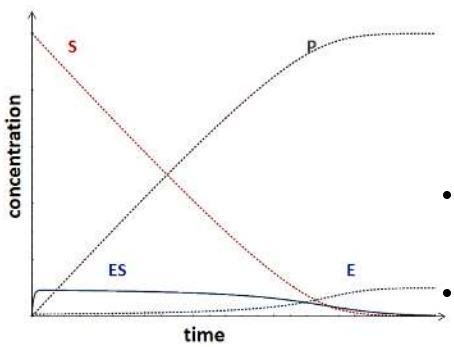


Figure 5.3.1: Change in concentrations over time for enzyme E, substrate S, complex ES and product P.

The rates of formation and breakdown of the E - S complex are given in terms of known quantities:

- The rate of formation of ES = $k_1[E][S]$

  (with the assumption that [P] =0)
- The rate of breakdown of ES = $k_2[ES] + k_1[ES] = (K_2 + K_1)[ES]$

  At steady state,

$$\frac{d[ES]}{dt} = k_1[E][S] + k_2[ES] + k_1[ES] = 0 \tag{5.3.10}$$

Therefore, rate of formation of [ES] is equal to the rate of breakdown of [ES]

So,

$$k_1[E][S] = (k_2 + k_1)[ES] \tag{5.3.11}$$

Dividing through by k_1:

$$[E][S] = \frac{(k_2 + k_1)}{k_1}[ES] \tag{5.3.12}$$

Substituting $\frac{(k_2 + k_1)}{k_1}$ with k_M:

$$[E][S] = K_M[ES]$$

$$k_M = \frac{breakdown[ES]}{formation[ES]} \tag{5.3.13}$$

K_m implies that half of the active sites on the enzymes are filled. Different enzymes have different K_m values. They typically range from 10^{-1} to 10^{-7} M. The factors that affect K_m are:

- pH
- temperature
- ionic strengths
- the nature of the substrate

Substituting $[E_F]$ with $[E_T]$-[ES]: $E_T = [ES] + [E_F]$

$$([E_T] - [ES])\,[S] = k_M\,[ES]$$

$$[E_T]\,[S] - [ES][S] = k_M\,[ES]$$

$$[E_T]\,[S] = [ES]\,[S] + k_M\,[ES]$$

$$[E_T]\,[S] = [ES]\,([S] + k_M)$$

Solving for [ES]:

$$[ES] = \frac{([E_T][S])}{([S] + k_M)}$$

The rate equation from the rate limiting step is:

$$v_o = \frac{dP}{dt} = k_2[ES]$$

Multiplying both sides of the equation by k_2:

$$k_2[ES] = k_2(\frac{([E_T][S])}{(K_M + [S])} \tag{5.3.14}$$

$$v = k_2(\frac{([E_T][S])}{(K_M + [S])} \tag{5.3.15}$$

When $S >> K_M$, v_o is approximately equal to $k_2[E_T]$. When the [S] great, most of the enzyme is found in the bound state ([ES]) and $V_o = V_{max}$

We can then substitue $k_2[E_T]$ with $V_{max}$to get the **Michaelis Menten Kinetic Equation:**

$$v = \frac{(v_{max}[S])}{(k_M + [S])}$$

Reaction Order Note

When $[S] << K_m$**,**

$$v = \frac{V_{max}[S]}{K_m}$$

This means that the rate and the substrate concentration are directly proportional to each other. The reaction is first-order kinetics.

When $[S] >> K_m$**,**

$$v = V_{max} \tag{5.3.16}$$

This means that the rate is equal to the maximum velocity and is independent of the substrate concentration. The reaction is zero-order kinetics.

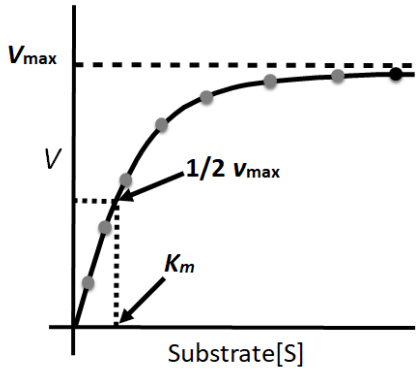


Figure 5.3.2: Diagram of reaction velocity and Michaelis-Menten kinetics.

$$v = \frac{V_{max}}{2}, K_m = [S]\text{n}$$

$$v = \frac{V_{max}}{2} = \frac{V_{max}[S]}{K_m + [S]} \tag{5.3.17}$$

Therefore, K_m is equal to the concentration of the substrate when the rate is half of the maximum velocity. From the Michaelis Menten Kinetic equation, we have many different ways to find K_m and V_{max} such as the Lineweaver-Burk plot, Hanes-Woolf plot, and Eadie-Hofstee plot, etc.

Lineweaver-Burk Plot

For example, by taking the reciprocal of the Michaelis Menten Kinetics Equation, we can obtain the Lineweaver-Burk double reciprocal plot:

$$v = \frac{(V_{max}[S])}{(K_M + [S])} \tag{5.3.18}$$

$$\frac{1}{v} = \frac{(k_M + [S])}{v_{max}[S]} \tag{5.3.19}$$

$$\frac{1}{v} = \left(\frac{K_m}{V_{max}}\right)\left(\frac{1}{[S]}\right) + \frac{1}{V_{max}} \tag{5.3.20}$$

Apply this to equation for a straight line $y = mx + b$ and we have:

$$y = \frac{1}{v} \tag{5.3.21}$$

$$x = \frac{1}{[S]} \tag{5.3.22}$$

$$m = slope = \frac{K_m}{V_{max}} \tag{5.3.23}$$

$$b = y - intercept = \frac{1}{V_{max}} \tag{5.3.24}$$

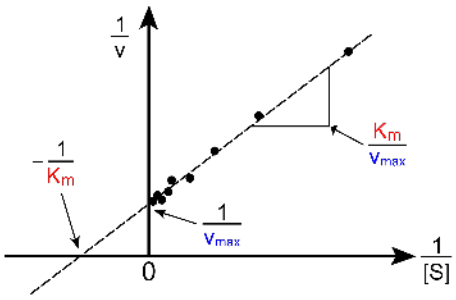
When we plot $y = \frac{1}{v}$ versus $x = \frac{1}{[S]}$, we obtain a straight line.

$$x - intercept = \frac{-1}{K_m} \tag{5.3.25}$$

$$y - intercept = \frac{1}{V_{max}} \tag{5.3.26}$$

$$slope = \frac{K_m}{V_{max}} \tag{5.3.27}$$

Figure 5.3.3. *An example of a Lineweaver-Burke plot.*



References

1. Chang, Raymond. *Physical Chemistry for the Biosciences*. Sansalito, CA: University Science, 2005. Page 363-371.
2. Atkins, Peter and de Paula, Julio. *Physical Chemistry for the Life Sciences*. New York, NY: W. H. Freeman and Company, 2006. Page 309-313.
3. Stryer, Lubert. *Biochemistry (Third Edition)*. New York, NY: W.H. Freeman and Company, 1988. Page 187-191.

Contributors

- Han Le, Sandy Algaze, Eva Tan

5.3: Michaelis-Menten Kinetics is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.4: Enzyme Inhibition

Inhibition caused by drugs may be either reversible or irreversible. A reversible situation occurs when an equilibrium can be established between the enzyme and the inhibitory drug. A competitive inhibition occurs when the drug, as "mimic" of the normal substrate competes with the normal substrate for the active site on the enzyme. Concentration effects are important for competitive inhibition.

Reversible inhibitors

A reversible inhibitor inactivates an enzyme through noncovalent, more easily reversed, interactions. Unlike an irreversible inhibitor, a reversible inhibitor can dissociate from the enzyme. Reversible inhibitors include competitive inhibitors and noncompetitive inhibitors. (There are additional types of reversible inhibitors.)

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. When the drug methotrexate is present, some of the 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.

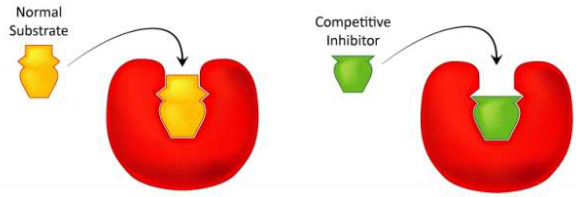


Figure 5.4.1: *Competitive Inhibition*

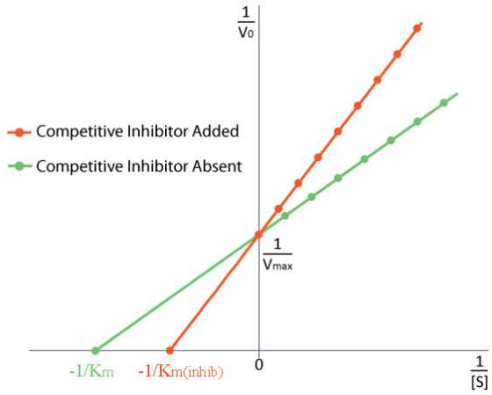


Figure 5.4.2: *Line-Weaver Burk Plot of competitive inhibition*

No Effect On *Vmax*

How do we study competitive inhibition. It is typically done as follows. First one performs a set of *V* vs. [S] reactions without inhibitor (20 or so tubes, with buffer and constant amounts of enzyme, varying amounts of substrate, equal reaction times). *V* vs. [S] is plotted, as well as 1/*v* vs. 1/[S], if desired. 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 inhibitor 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). Graphically, the results of

these experiments are shown above. Notice that at high substrate concentrations, the competitive inhibitor has essentially no effect, causing the *Vmax* for the enzyme to remain unchanged. This is due to the fact that at high substrate concentrations, the inhibitor doesn't compete well. However, at lower substrate concentrations it does.

Increased *Km*

Note that the apparent *Km* of the enzyme for the substrate increases (-1/*Km* gets closer to zero - red line above) when the inhibitor is present, 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 reducing the amount of active enzyme at lower 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.

Studies of competitive inhibition have provided helpful information about certain enzyme-substrate complexes and the interactions of specific groups at the active sites. As a result, pharmaceutical companies have synthesized drugs that competitively inhibit metabolic processes in bacteria and certain cancer cells. Many drugs are competitive inhibitors of specific enzymes.

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 (rectangular site below). 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].

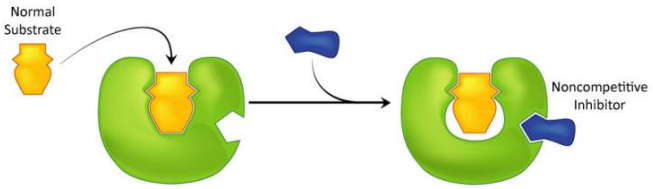


Figure 5.4.3: *Non-competitive Inhibition*

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 above. As you can see, Vmax is reduced in non-competitive inhibition compared to uninhibited reactions. This makes sense if we remember that Vmax is dependent on the amount of enzyme present. Reducing the amount of enzyme present reduces Vmax. In competitive inhibition, this doesn't occur detectably, because at high substrate concentrations, there is essentially 100% of the enzyme active and the Vmax 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 KM of active enzymes and KM is a constant for a given enzyme.

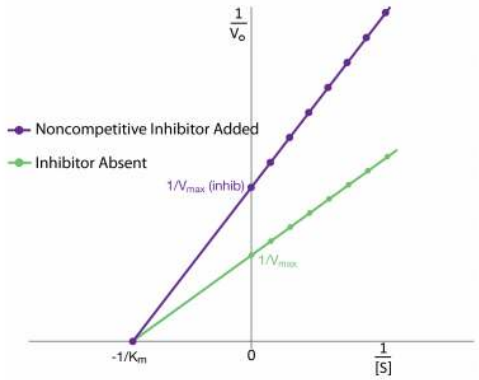


Figure 5.4.4: *Line-Weaver Burk Plot of noncompetitive inhibition*

Feedback inhibition is a normal biochemical process that makes use of noncompetitive inhibitors to control some enzymatic activity. In this process, the final product inhibits the enzyme that catalyzes the first step in a series of reactions. Feedback inhibition is used to regulate the synthesis of many amino acids. For example, bacteria synthesize isoleucine from threonine in a series of five enzyme-catalyzed steps. As the concentration of isoleucine increases, some of it binds as a noncompetitive inhibitor to the first enzyme of the series (threonine deaminase), thus bringing about a decrease in the amount of isoleucine being formed.

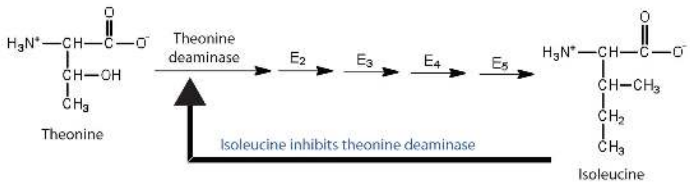


Figure 5.4.5: *Feedback Inhibition of Threonine Deaminase by Isoleucine. Threonine deaminase is the first enzyme in the conversion of threonine to isoleucine. Isoleucine inhibits threonine deaminase through feedback inhibition.*

Uncompetitive Inhibition

A third type of enzymatic inhibition is that of uncompetitive inhibition, which has the odd property of a reduced V*max* as well as a reduced Km. The explanation for these seemingly odd results is due to the fact that the uncompetitive inhibitor binds only to the enzyme-substrate (ES) complex. 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 result in reduced V*max*.

The reduced Km is a bit harder to conceptualize. The answer lies in the fact 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 Vmax and increases an enzyme's affinity for its substrate.

Summary

In **competitive inhibition** the substrate and the inhibitor compete for the same active site on the enzyme. Because the substrate cannot bind to an enzyme–inhibitor complex, EI, the enzyme's catalytic efficiency for the substrate decreases. With **noncompetitive inhibition** the substrate and the inhibitor bind to different active sites on the enzyme, forming an enzyme–substrate–inhibitor, or ESI complex. The formation of an ESI complex decreases catalytic efficiency because only the enzyme–substrate complex reacts to form the product. Finally, in **uncompetitive inhibition** the inhibitor binds to the enzyme–substrate complex, forming an inactive ESI complex.

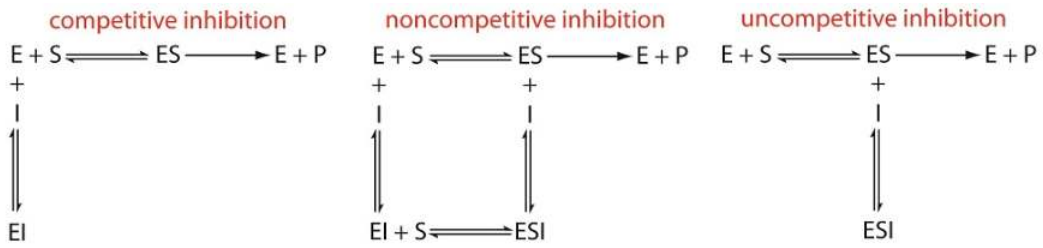


Figure 5.4.6*: Mechanisms for the reversible inhibition of enzyme catalysis. E: enzyme, S: substrate, P: product, I: inhibitor, ES: enzyme–substrate complex, EI: enzyme–inhibitor complex, ESI: enzyme–substrate–inhibitor complex.*

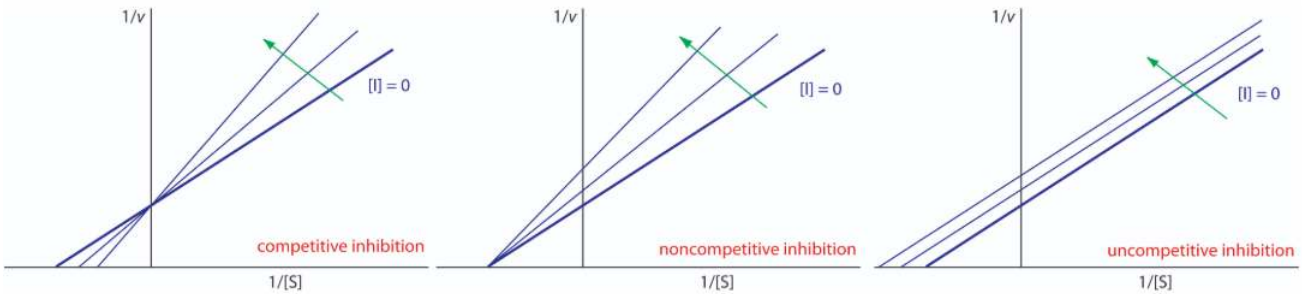
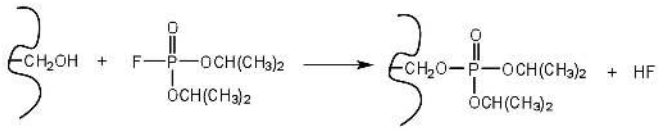


Figure 5.4.6: *Linweaver–Burk plots for competitive inhibition, noncompetitive inhibition, and uncompetitive inhibition. The thick blue line in each plot shows the kinetic behavior in the absence of inhibitor, and the thin blue lines in each plot show the change in behavior for increasing concentrations of the inhibitor. In each plot, the inhibitor's concentration increases in the direction of the green arrow.*

Irreversible inhibitors

An irreversible inhibitor inactivates an enzyme by bonding **covalently** to a particular group at the active site. The inhibitor-enzyme bond is so strong that the inhibition cannot be reversed by the addition of excess substrate. The nerve gases, especially Diisopropyl fluorophosphate (DIFP), irreversibly inhibit biological systems by forming an enzyme-inhibitor complex with a specific OH group of serine situated at the active sites of certain enzymes. The peptidases trypsin and chymotrypsin contain serine groups at the active site and are inhibited by DIFP.



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 and thus cannot be removed. 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 5.4.7), which covalently links to the bacterial enzyme, **D-D 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. If one were to measure the kinetics of suicide inhibitors under conditions where there was more enzyme than inhibitor, they would resemble non-competitive inhibition's kinetics because both involve reducing the amount of active enzyme by a fixed amount in a set of reactions.

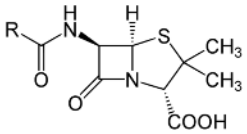


Figure 5.4.7: *Penicillin*

Contributors

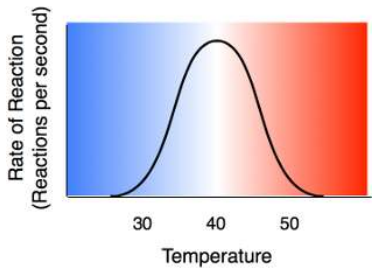
Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

This page titled 5.4: Enzyme Inhibition is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

5.5: Temperature, pH, and enzyme concentration on the rate of a reaction

Temperature

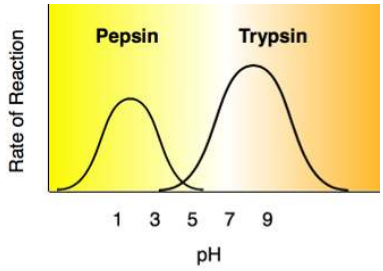
Higher temperature generally causes more collisions among the molecules and therefore increases the rate of a reaction. More collisions increase the likelihood that substrate will collide with the active site of the enzyme, thus increasing the rate of an enzyme-catalyzed reaction. Above a certain temperature, activity begins to decline because the enzyme begins to denature. The rate of chemical reactions therefore increases with temperature but then decreases as enzymes denature.



pH

Each enzyme has an optimal pH. A change in pH can alter the ionization of the R groups of the amino acids. When the charges on the amino acids change, hydrogen bonding within the protein molecule change and the molecule changes shape. The new shape may not be effective.

The diagram below shows that pepsin functions best in an acid environment. This makes sense because pepsin is an enzyme that is normally found in the stomach where the pH is low due to the presence of hydrochloric acid. Trypsin is found in the duodenum, and therefore, its optimum pH is in the neutral range to match the pH of the duodenum.



Most cells form hydrogen peroxide (H_2O_2) as a waste product of aerobic respiration. Hydrogen peroxide is toxic and must be converted to water and oxygen by the enzyme catalase.

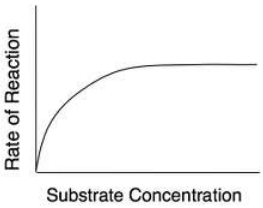
$$2\,H_2O \xrightarrow{\text{catalase}} 2\,H_2O + O_2 \tag{5.5.1}$$

Hydrogen peroxide is also commonly used as a household disinfectant. It bubbles when it is applied to cuts and scrapes because **catalase** is present in the fluids of the broken cells. As the equation above shows, the bubbles are oxygen gas (O_2).

Substrate Concentration

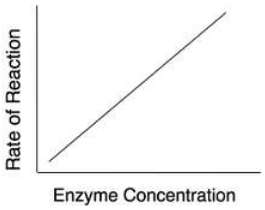
At lower concentrations, the active sites on most of the enzyme molecules are not filled because there is not much substrate. Higher concentrations cause more collisions between the molecules. With more molecules and more collisions, enzymes are more likely to encounter molecules of reactant.

The maximum velocity of a reaction is reached when the active sites are almost continuously filled. Increased substrate concentration after this point will not increase the rate. *Reaction rate therefore increases as substrate concentration is increased but it levels off.*



Enzyme Concentration

If there is insufficient enzyme present, the reaction will not proceed as fast as it otherwise would because all of the active sites are occupied with the reaction. Additional active sites could speed up the reaction. *As the amount of enzyme is increased, the rate of reaction increases.*



Contributors

Template:ContribGregory

Curation and Revision. **Provided by**: Boundless.com. **License**: ***CC BY-SA: Attribution-ShareAlike***

5.5: Temperature, pH, and enzyme concentration on the rate of a reaction is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

5.6: Multi-Substrate Sequential Mechanisms

The Michaelis –Menten model of enzyme kinetics was derived for **single substrate reactions**. Enzymatic reactions requiring multiple substrates and yielding multiple products are more common than single-substrate reaction. In these types of reactions, the all the substrates involved are bound to the enzyme before catalysis of the reaction takes place to release the products. Sequential reactions can be either ordered or random. In contrast to the Michealis-Menton kinetics where a binary Enzyme-Substrate complex is generated in the mechanism ([ES], in bi-substrate enzyme reactions, a ternary complex of the enzyme and two substrates is generated:

$$\mathrm{A} + \mathrm{B} \overset{\mathrm{E}}{\rightleftharpoons} \mathrm{P} + \mathrm{Q} \tag{5.6.1}$$

Bisubstrate reactions account for ~ 60% of the known enzymatic reactions. Multi-substrate reactions follow complex rate equations that describe how the substrates bind and in what sequence. The analysis of these reactions is much simpler if the concentration of substrate ***A*** is kept constant and substrate ***B*** varied. Under these conditions, the enzyme behaves just like a single-substrate enzyme and a plot of *v* by *[S]* gives apparent *KM* and *Vmax* constants for substrate ***B***. If a set of these measurements is performed at different fixed concentrations of ***A***, these data can be used to work out what the mechanism of the reaction is. For an enzyme that takes two substrates ***A*** and ***B*** and turns them into two products ***P*** and ***Q,*** there are two types of mechanism: ternary complex and ping–pong.

How do you resolve the enzymes kinetics of these more complicated systems? The answer is fairly straightforward. You keep one of the substrates (***B***, for example) fixed, and vary the other substrate (***A***) and obtain a **series** of hyperbolic plots of v_o vs ***A*** at different fixed ***B*** concentrations. This would give a series of linear 1/*v* vs. 1/***A*** double-reciprocal plots (Lineweaver-Burk plots) as well. The pattern of Lineweaver-Burk plots depends on how the reactants and products interact with the enzyme.

Sequential Mechanism

In this mechanism, both substrates must bind to the enzyme before any products are made and released. The substrates might bind to the enzyme in a **random fashion** (***A*** first then ***B*** or vice-versa) or in an **ordered fashion** (***A*** first followed by ***B***). For both mechanisms, Lineweaver-Burk plots at varying ***A*** and different fixed values of ***B*** give a series of intersecting lines. Derivative curves can be solved to obtain appropriate kinetic constants. In ordered sequential reactions, all the substrates are first bound to the enzyme in a **defined order or sequence**. The products, too, are released after catalysis in a defined order or sequence.

An example is the lactate dehydrogenase enzyme, which is a protein that catalyzes glucose metabolism. In this ordered mechanism, the coenzyme, NADH, always binds first, with pyruvate binding afterward. During the reaction, the pyruvate is reduced to lactate while NADH is oxidized to NAD^+ by the enzyme. Lactate is then released first, followed by the release of NAD^+.

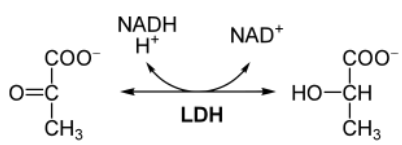


Figure 5.6.1*: Reaction of the lactate dehydrogenase: pyruvate (left) is oxidized to lactate (right) by expense of NADH*

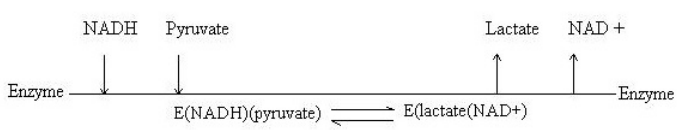


Figure 5.6.2*: Ordered Sequential Mechanism for the lactate dehydrogenase enzyme*

This is a characteristic of a ternary complex, which consists of three molecules that are bound together. Before catalysis, the substrates and coenzyme are bound to the enzyme. After catalysis, the complex consists of the enzyme and products, NAD^+ and lactate.

In random sequential reactions, the substrates and products are bound and then released in no preferred order, or "random" order (Figure 3.4.1). An example is the creatine kinase enzyme, which catalyzes creatine and ATP (the two substrates) to form phosphocreatine and ADP (teh Products) in Figure 3.4.4. In this case, **either** substrates may bind first and either products may be

released first. A ternary complex is still observed for random sequential reactions. Before catalysis, the complex is generated that includes the enzyme, ATP, and creatine. After catalysis, the complex consists of the enzyme, ADP, and phosphocreatine.

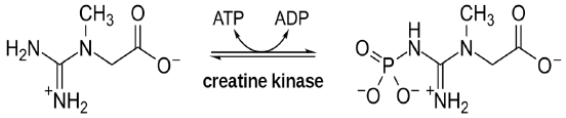


Figure 5.6.3: The metabolism of creatine kinase follows a random sequential mechanism.

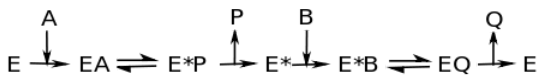
This page titled 5.6: Multi-Substrate Sequential Mechanisms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

5.7: Double displacement reaction

The simplest of enzymes will involve one substrate binding to the enzyme and producing a product plus the enzyme. However, the majority of enzymes are more complex and catalyze reactions involving multiple substrates. Binding of two substrates can occur through two mechanisms: sequential mechanism and non-sequential mechanism. In sequential mechanisms both substrates bind the enzyme and the reaction proceeds to form products which are then released from the enzyme. This mechanism can be further subdivided into random and ordered reactions. For random reactions the order in which the substrates bind does not matter. In ordered reactions one substrate must bind the enzyme before the second substrate is able to bind. Non-Sequential mechanism does not require both substrates to bind before releasing the first product. This page will focus on the non-sequential mechanism, which is also known as the "ping-pong" mechanism. It is called this because the enzyme bounces back and forth from an intermediate state to its standard state.The enzyme acts like a ping-pong ball, bouncing from one state to another.

The Mechanism

Ping-pong mechanism, also called a double-displacement reaction, is characterized by the change of the enzyme into an intermediate form when the first substrate to product reaction occurs. It is important to note the term intermediate indicating that this form is only temporary. At the end of the reaction the enzyme MUST be found in its original form. An enzyme is defined by the fact that it is involved in the reaction and is not consumed. Another key characteristic of the ping-pong mechanism is that one product is formed and released before the second substrate binds. The figure below explains the Ping Pong mechanism through an enzymatic reaction.



This image shows that as substrate A binds to the enzyme, enzyme-substrate complex EA forms. At this point, the intermediate state, E* forms. P is released from E* , then B binds to E*. B is converted to Q, which is released as the second product. E* becomes E, and the process can be repeated. Often times, E* contains a fragment of the original substrate A.This fragment can alter the function of the enzyme, gets attached to substrate B, or both.

One example of a ping-pong enzyme is low molecular weight protein tyrosine phosphatase. It reacts with the small substrate p-initrophenylphosphate (A) which binds to the enzyme covalently with the expulsion of the product P, the p-nitrophenol leaving group. Water (B) then comes in and covalently attacks the enzyme, forming an adduct with the covalently bound phosphate releasing it as inorganic phosphate. In this particular example, however, you can't vary the water concentration and it would be impossible to generate the parallel Lineweaver-Burk plots characteristic of ping-pong kinetics.

Example 1: Chymotrypsin

An example of the ping-pong mechanism would be the action of chymotrypsin. When reacted with p-nitrophenyl acetate (A), the reaction of chymotrypsin is seen to occur in two steps. In the first step, the substrate reacts extremely fast with the enzyme, leading to the formation of a small amount of p-nitrophenolate (P). In the second step, the substrate-enzyme interaction results in the formation of acetate ion (Q). The action of chymotrypsin is a ping-pong reaction because the binding of the two substrates causes the enzyme to switch back and forth between two states. Please refer to the section Chymotrypsin and pre-steady-state enzyme kinetics for more details on the action of chymotrypsin.

Example 2: Pyruvate Carboxylase

Another example of an enzyme that exhibits a ping-pong mechanism is pyruvate carboxylase. This enzyme catalyzes the addition of carbon dioxide to pyruvate in order to form oxaloacetate. (leads to gluconeogenesis) This biotin-containing enzyme works by binding CO_2 (A) to form carboxybiotin (EA). The biotin swings over towards pyruvate (E*P) and releases CO_2 (P, due to the fact that it had been moved from its original binding site) Pyruvate (B), in close proximity to CO_2, attacks the partial positive of Carbon in CO_2 (E*B). Oxaloacetate is formed within the enzyme (EQ) and gets released (Q). While this attack is occurring, biotin swings back to its initial position, (E* --> E) and is ready to bind another CO_2.

Contributors

Melissa Hill, Laura Lan, Andrew Jilwan

References

1. Chang, Raymond. 2005. Physical Chemistry for the Biosciences. Sausalito (CA): University Science Books. p. 372-377.
2. Cleland, Wallace. "Derivation of Rate Equations for Multisite Ping-Pong Mechanisms with Ping-Pong Reactions at One or More Sites." *Journal of Biological Chemistry* 248.24 (1973): 8353-355. Print.
3. Garrett, R., and Charles M. Grisham. "Gluconeogenesis, Glycogen Metabolism, and the Pentose Phosphate Pathway." *Biochemistry*. Belmont, CA: Brooks/Cole, Cengage Learning, 2010. 664-66. Print.
4. Garrett, Reginald and Charles M. Grisham. "Enzymes-Kinetics and Specificity". *Biochemistry, Fourth Edition.* Belmont, CA: Brooks/Cole, Cengage Learning, 2010. p. 406-407.

5.7: Double displacement reaction is shared under a CC BY license and was authored, remixed, and/or curated by LibreTexts.

6: Classification and Catalytic Strategies of Enzymes

Classification

http://nptel.ac.in/courses/104102016/36

Enzymes are generally named according to the reaction they catalyze or by suffixing **“ase”** after the name of substrate. The International Union of Biochemistry and Molecular Biology developed a nomenclature for enzymes. Each enzyme is described by a sequence of four numbers preceded by **"EC"**. EC denotes **Enzyme Commission** and the number of enzyme is called **EC numbers**.

When classified, each enzyme is assigned the EC number, in the form of digits separated by periods. The first number categorizes the enzyme based on its reaction.

- The first three numbers represent the **class, subclass and sub-subclass** to which an enzyme belongs, and the fourth digit is a serial number to identify the particular enzyme within a sub-subclass.
- The class, subclass and sub-subclass provide additional information about the reaction classified. For example, in the case of EC 1.2.3.4, the digits indicate that the enzyme is an oxidoreductase (class 1), that it acts on the aldehyde or oxo group of donors (subclass 2), that oxygen is an acceptor (sub-subclass 3) and that it was the fourth enzyme classified in this sub-subclass (serial number 4).
- The last printed list of enzymes appeared in the year 1992. Since then it has been updated and maintained online.

Accordingly the enzymes are classified into **seven main family classes** and many sub-family classes.

Oxidoreductases, Transferases, Hydrolases, Lyases, Isomerases, Ligases, **Translocases**

https://en.wikibooks.org/wiki/Struct...Classification

https://www.creative-enzymes.com/resource/enzyme-definition-and-classification_18.html

Catalytic Strategies

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 of reactions.

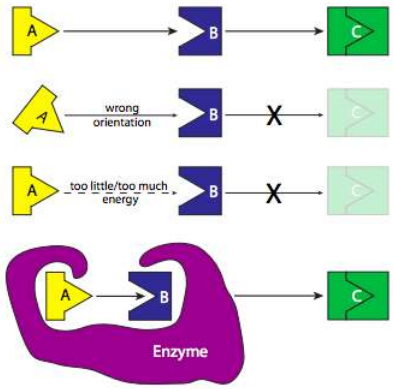


Figure 6.1: Enzymes can lower activation energy by binding substrates individually and bringing them together under optimal conditions to react.

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.

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.

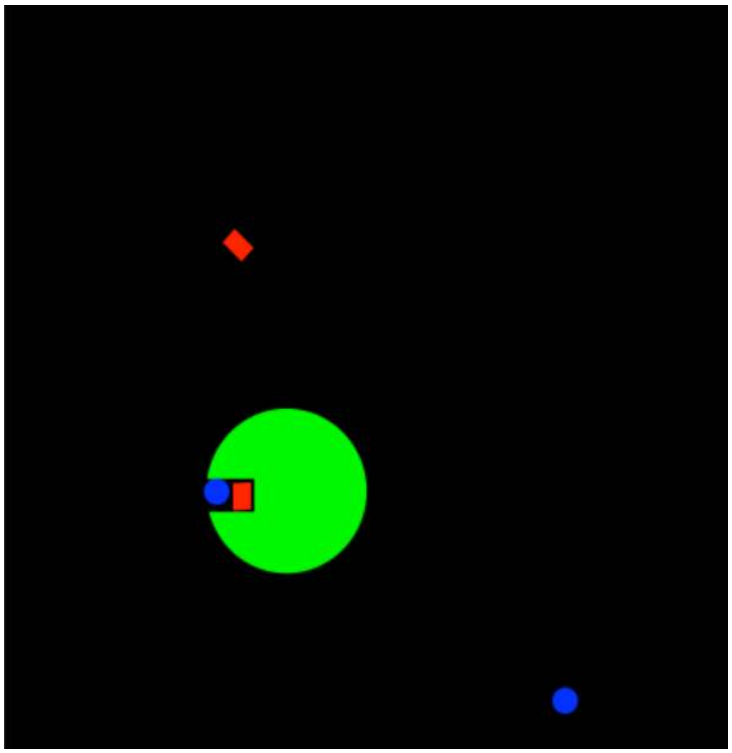
Once a substrate has been bound, it is the enzyme's job to quickly transform the substrate into product. The enzyme does so by carrying the substrate over a catalytic pathway. In a catalytic pathway, the reaction takes a different course than it would on its own. Sometimes the catalytic pathway is longer, involving additional steps, but because the energetic terrain is easier to traverse than in the un-catalyzed process, the catalytic reaction actually takes much less time.

Enzymes have a range of structures and reaction properties, so there are a wide number of different reactions they can catalyze. Nevertheless, there are a few common strategies displayed in catalytic reactions that are useful to know.

Approximation

When we make an approximation, we are getting close to the answer. If someone asks us what time it is, and it is 3:02 pm, we probably tell them it's about three o'clock. That's close enough.

In enzyme catalysis, approximation means getting things close to each other. If we have a substrate that is going to react with something, the enzyme can bind the substrate in such a way as to get the substrate in close proximity to the reactant. Sometimes, but not always, that might mean binding two substrates, so that they can more easily react with each other. If you want to meet your friend, it's a much better idea to say, let's meet at Cafe Santropol at 7:00, rather than wander the streets of Montreal, along with four million other people, hoping to bump into your friend. The reactant and substrate are much more likely to encounter each other in the protected confines of the enzyme than they are floating around in the wilderness of the cell.



- In approximation, two substrates are held close together within the enzyme
- This proximity makes them react together more easily

Sometimes, this idea can be a little more subtle. Imagine that the reactant is the enzyme itself, so that just by binding the single substrate, the reaction is already much more likely to occur. The substrate may be held in such a way that it is already in close proximity to amino acid side chains that will work on it and transform it into a new molecule.

- In approximation, the substrate is held in position so that subsequent reaction is much more likely

Approximation is really an entropy factor. By binding the substrate, we can limit its degrees of freedom, restricting its location or even its orientation so that there is no way the reactant can miss its intended target.

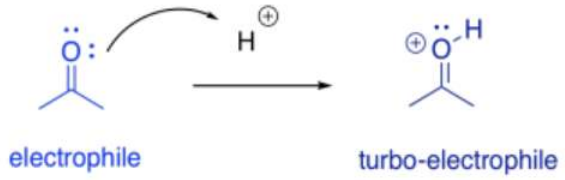
Of course, the substrates still have to find the active site of the enzyme. Sometimes, that task is aided by sticky surfaces on the enzyme; there can be groups on the surface of the enzyme that can interact with substrates, so that the substrates are less likely to drift away after a random collision with the enzyme. With its movement thus restricted, the substrate is more likely to move into the active site than drift off through the cell once more.

Although we are discussing enzymes in the cell, other kinds of catalysis make use of approximation as well. For example, transition metal catalysis often makes use of solid chunks of metal to catalyze the reactions of gaseous vapors. The surface of the metal gives the gas-phase molecules a place to bind, giving them a place to gather, rather than wandering around in three dimensions in the gas phase. Once they are together in one place, they are more likely to react with each other or with additional species on the surface of the metal.

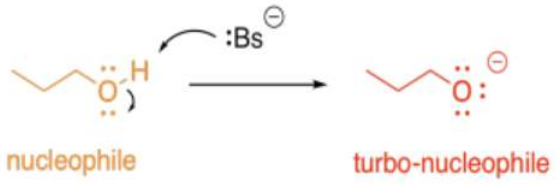
Acid-Base Catalysis

Acid-base catalysis is a very common phenomenon. So many reactions involve the addition or removal of protons, especially the carbonyl reactions that are so prevalent in biochemical pathways, that proton donors and acceptors become key players. Acidic and basic side chains of the amino acids in the protein naturally fill these roles.

Acid-base catalysis can provide mechanistic advantages by rapidly enhancing the electrophilicity of a molecule. Any carbonyl compound is electrophilic, but if it gets protonated, the overall positive charge makes the carbonyl even more electrophilic.



Alternatively, acid-base catalysis might increase the nucleophilicity of a molecule. Any alcohol is nucleophilic, because it has lone pairs. However, if its proton is removed, it becomes even more nucleophilic, because of the overall negative charge.



So, in the most straightforward case, adding a proton might accelerate a reaction involving an electrophile. Removing a proton might accelerate a reaction involving a nucleophile.

- Acid/base catalysis involves rapid proton shuttling
- Acidic side chains can activate electrophiles
- Basic side chains can activate nucleophiles

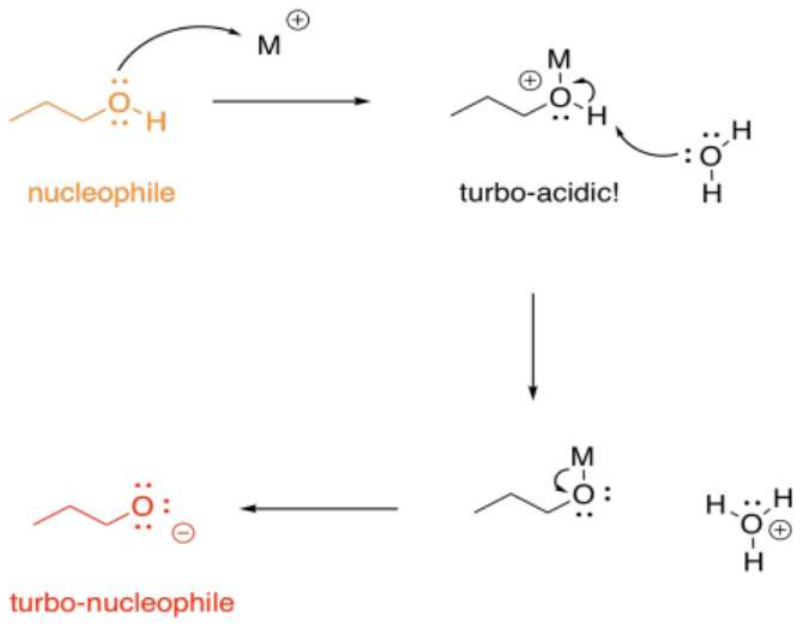
There are other variations on this approach. For example, consider a keto-enol tautomerism. We think of ketones classically as electrophiles, but their enol isomer is easily accessible in general, and the enol form is an excellent nucleophile. As part of a series of reaction, rapid conversion of a ketone into an enol might be a key step.

Metal Ion Catalysis

Metal ion catalysis can often be thought of as a special case of acid-base catalysis. With metal ions, we get Lewis acid catalysis. Lewis acid catalysis can accelerate reactions in a couple of different ways, in close analogy with general acid/base catalysis.

It sometimes helps to think of a metal as a great, big proton. That's an oversimplification, as we'll see in a moment. Nevertheless, it can be useful to keep the analogy in mind. When a compound binds to a metal ion, the effect can be similar to binding a proton. The compound has just donated a pair of electrons elsewhere (to the metal ion or to the proton), and so the compound suddenly looks electron-deficitient. It has enhanced electrophilicity.

On the other hand, we don't usually think of protonation as causing increased basicity; that would be completely backwards. With metals, though, that can happen, indirectly. If an alcohol, for example, donates a lone pair to a metal, the oxygen becomes positively charged. It becomes much more acidic. Suddenly, the alcohol can be deprotonated by a very, very weak base, such as water. That leads to formation of an alkoxide ion, which is much more nucleophilic than the original alcohol. Water would really never take a proton from an alcohol, but it can do it once the oxygen has a positive charge.



Metal ions can play a number of other roles in catalysis, but that's enough to get an idea of just some of the ways in which they might be useful. To learn more, we would have to explore more transition metal reactivity, including the ability of metals to donate and accept individual electrons.

Covalent catalysis

See 6.1 serine protease.

Suggested homework: lock-and-key model vs induced fit model.

Contributors

Contributed by E. V. Wong

Chris P Schaller, Ph.D., (College of Saint Benedict / Saint John's University)

Axolotl Academica Publishing (Biology) at Axolotl Academica Publishing

This page titled 6: Classification and Catalytic Strategies of Enzymes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

6.1: Serine proteases

Serine proteases are just one type of endoproteases. However, they are extremely abundant in both prokaryotes and eukaryotes. Protease A, a chymotrypsin-like protease from Stremptomyces griseus, has a very different primary sequence than chymotrypsin, but its overall tertiary structure is quite similar to chymotrypsin, The positions of the catalytic triad amino acids in the primary sequences of the protein are very similar, indicating that the genes for the proteins diverged from a common precursor gene. In contrast, subtilisin, a serine protease from B. Subtilis, has both limited sequence and tertiary structure homology to chymotrypsin. However, when folded it also has a catalytic triad (Ser 221 - His 64 - Asp 32) similar to that of chymotrypsin (Ser 195 - His 57 - Asp 102).

Proteases have multiple functions, other than in digestion, including degrading old or misfolded proteins and activating precursor proteins (such as clotting proteases and proteases involved in programmed cell death). In general, the catalytic strategies used by four classes of enzymes: the serine proteases, carbonic anhydrases, restriction endonucleases, and nucleoside monophosphate (NMP) kinases. The first three classes of enzymes catalyze reactions that require the addition of water to a substrate. Proteases can also be integral membrane proteins, and carry out their activities in the hydrophobic environment of the membrane. For example, aberrant cleavage of the amyloid precursor protein by the membrane protease presenillin can lead to the development of Alzheimers.

Chymotrypsin

Consider the mechanism of catalysis of the enzyme known as chymotrypsin. Found in our digestive system, chymotrypsin's catalytic action 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. It acts fairly specifically, cutting not all peptide bonds, but only those that are adjacent to specific 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.

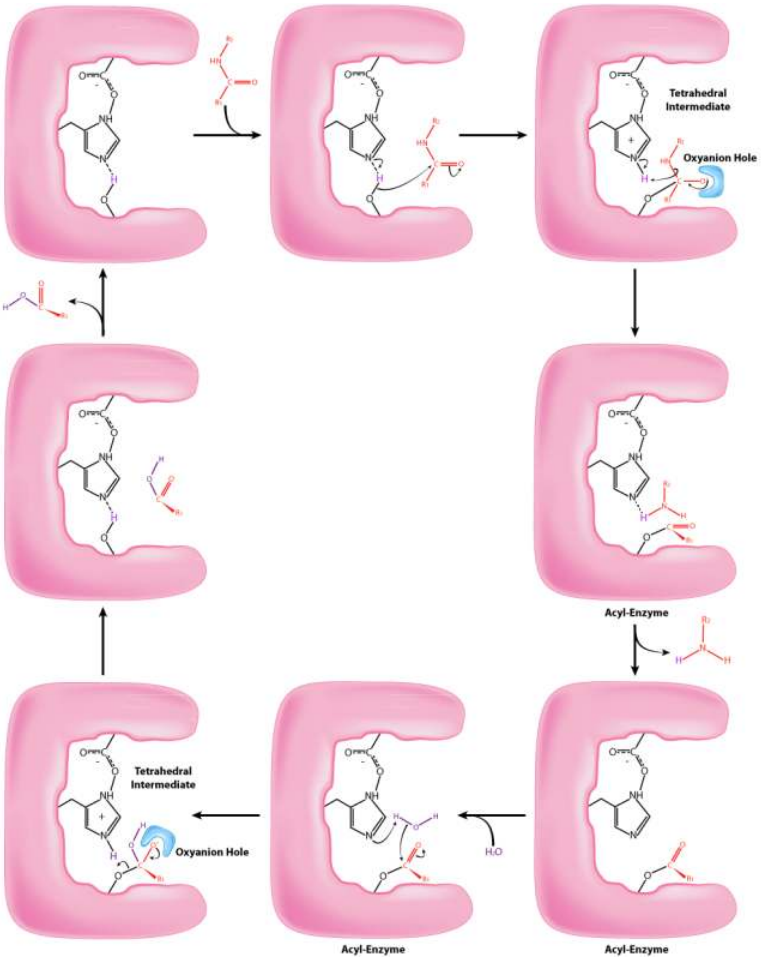


Figure 6.1.1: *mechanism of catalysis of the enzyme chymotrypsin.*

The process starts with the binding of the substrate in the S1 pocket. 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 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. When the proper substrate binds, it stays and its presence induces an ever so slight shift in the shape of the enzyme. *This subtle shape change on the binding of the proper substrate starts the steps of the catalysis and is the reason that the enzyme shows specificity for cutting at specific enzyme positions 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 of the enzyme upon binding of the proper substrate cause changes in the positioning of three amino acids (aspartic acid, histidine, and serine) in the active site known as the catalytic triad, during the second step of the catalytic action. 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. 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 active site. This reaction, which is the third step of catalysis, breaks the bond and causes two things to happen. First, one end of the original polypeptide is freed and exits the active site. 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.

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 that performs a nucleophilic attack on the phenylalanine-serine bond, 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.

Trypsin

Trypsin is a serine protease produced in the pancreas, is found in digestive system of vertebrates to digest food proteins. Trypsin cleaves peptide chains primarily at the carboxyl side of the lysine or arginine residues. Like chymotrypsin, trypsin uses a conserved catalytic triad, Ser, His, Asp, in the active site for catalysis, and the chemical mechanism is the same as chymotrypsin as shown in Figure 6.1.1.

This page titled 6.1: Serine proteases is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

6.2: Transititon State Analogs and Catalytic Antibodies

Transition State Analog inhibitors

All chemical reactions progress through a transition state, a transient and unstable species between a substrate and a product, which contains the highest energy in the reaction progress. To reach the transition state, activation energy is required.

Without the participation of an enzyme, the activation energy is large, and as a result, the chance of the reaction to happen is small. In the presence of an enzyme, the activation energy will be lowered significantly to facilitate the reaction. This could lead to an increase in rates of some reactions by trillions of folds.

If energy does not simply appear or disappear, then, how does an enzyme lower the activation energy? This is because the enzyme could bind to the transition state extremely tightly, resulting in releasing a large amount of binding energy that is utilized to "compensate" the activation energy (to make it look smaller).

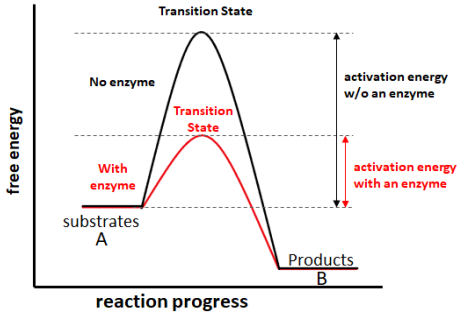


Figure 6.2.1: The energies of the stages of a chemical reaction. Without enzyme (Black line), substrates need a lot of activation energy to reach a transition state, which then decays into lower-energy products. When enzyme catalyzed (Red line), the enzyme binds the substrates (A), then stabilizes the transition state ($ES^{\ddagger}$) to reduce the activation energy required to produce products (B) which are finally released.

A transition state analog, one exhibiting the same properties such as shape and charge of the original transition molecule, may come in and bind. Although the analog displays similar properties as the original transition molecule, but it has higher affinity for the enzyme than the natural substrate and will ultimately deactivate and inhibit the enzyme and prevent it from binding to a substrate. The analogs can "function as antimetabolites".

Catalytic Antibody

Transition-state analogs have been used for generating catalytic antibodies, antibodies that catalyze chemical reactions. Enzymes lower activation energy and accelerate catalysis by tightly binding to the transition state. If an antibody could bind the transition state tightly like an enzyme, and it should also catalyze the reaction just like an enzyme. The transition state has short lifetimes (suggested to be as short as 10^{-13} S) and cannot be used to generate antibodies as antigens. Their stable mimics, the transition-state analogs, which have long half-lives and could be synthesized chemically, may function as an antigen to generate the antibody.

One example of catalytic antibody is the one that catalyzes the chelation of ferrous iron into the porphyrin plane of protoporphyrin IX which is naturally catalyzed by ferrochelatase in living organisms, the last enzyme in the synthetic pathway of heme. During catalysis, ferrochelatase significantly bends the planar porphyrin plane into a bowl-liked ring as the transition state. N-methylmesoporphyrin, a chemical causing hepatic protoporphyria mice, resembles the transition state by containing a bent and stable porphyrin ring. When N-methylmesoporphyrin was injected to animals as the antigen, monoclonal tightly binding antibodies were produced. Some of the produced antibodies were capable to distort the planary planar porphyrin plane into a bowl-liked ring and insert of iron into protoporphyrin IX, just like ferrochelatase. to facilitate the entry of ferrous iron. Using a similar technique, antibodies that catalyze ester and amide hydrolysis, transesterification, and photoinduced cleavage, among other reactions, have been developed.

Contributors

https://en.wikibooks.org/wiki/Structural_Biochemistry/Enzyme/Transition_State_Analogs

http://arjournals.annualreviews.org/...1415?cookieSet

6.2: Transititon State Analogs and Catalytic Antibodies is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

6.3: Restriction Endonuclease

Restriction Enzymes or restriction endonucleases are part of a bacterial defense system against foreign DNA, such as an infectious bacteriophage and viruses. The bacterial enzyme called **methyltransferase** methylates its own DNA with methyl groups ($-CH_3$) in a process called **DNA Methylation**. This process allows the bacteria to recognize its own DNA and destroy any foreign DNA (unmethylated viral DNA) with the help of restriction enzyme. The combined activities of the restriction endonuclease and methyltransferase are referred to as a **Restriction Modification System**. Today, most commercially available REs are not purified from their natural sources. Instead, REs are usually isolated from bacteria that overexpress large quantities of REs from plasmids. These recombinant REs have often been engineered by molecular biologists to include amino acid changes that increase the catalytic activity or stability of the RE.

To be able to sequence DNA, it is first necessary to cut it into smaller fragments. Many DNA-digesting enzymes (like those in your pancreatic fluid) can do this, but most of them are no use for sequence work because they cut each molecule randomly. This produces a heterogeneous collection of fragments of varying sizes. What is needed is a way to cleave the DNA molecule at a few precisely-located sites so that a small set of homogeneous fragments are produced. The tools for this are the restriction endonucleases. The rarer the site it recognizes, the smaller the number of pieces produced by a given restriction endonuclease.

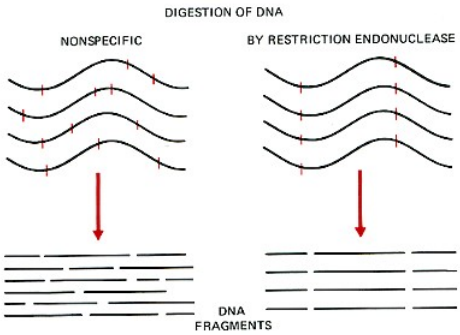


Figure 6.3.1: Restriction Digest

A restriction enzyme recognizes and cuts DNA only at a particular sequence of nucleotides. For example, the bacterium *Hemophilus aegypticus* produces an enzyme named HaeIII that cuts DNA wherever it encounters the sequence

5'GGCC3'

3'CCGG5'

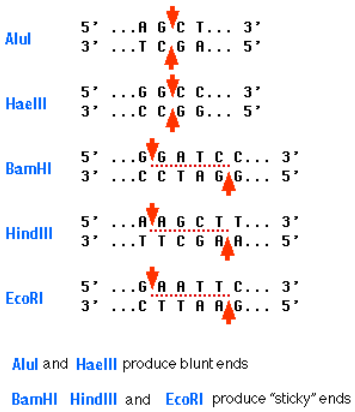


Figure 6.3.2: Restriction Enzymes

Restriction enzymes hydrolyze covalent phosphodiester bonds of the DNA to leave either "**sticky/cohesive" ends** or "**blunt" ends.**This distinction in cutting is important because an *EcoRI* sticky end can be used to match up a piece of DNA cut with the same enzyme in order to glue or ligate them back together. While endonucleases cut DNA, **DNA ligases** join them back together. DNA digested with *EcoRI* can be ligated back together with another piece of DNA digested with *EcoRI*, but not to a piece digested with *SmaI*. Another blunt cutter is *EcoRV* with a recognition sequence of GAT | ATC.

EcoRI restriction enzyme recognition site SmaI restriction enzyme recognition site

***EcoRI** generates sticky of cohesive ends **SmaI** generates blunt ends*

Contributors

Template:ContribAhern

Contributed by Clare M. O'Connor, Associate Professor Emeritus (Biology) at Boston Collegehttp://www.biology-pages.info/This content is distributed under a Creative Commons Attribution 3.0 Unported (CC BY 3.0) license and made possible by funding from The Saylor Foundation.

Dr. Todd Nickle and Isabelle Barrette-Ng (Mount Royal University) The content on this page is licensed under CC SA 3.0 licensing guidelines

---

6.3: Restriction Endonuclease is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7: Regulation of Enzyme Activity

In living cells, there are hundreds of different enzymes working together in a coordinated manner, and since cells neither synthesize nor break down more material than is required for normal metabolism and growth, precise enzyme regulation is required for turning metabolic reactions on and off. There is tremendous diversity in the mechanisms bacteria use to regulate enzyme synthesis and enzyme activity.

This page titled 7: Regulation of Enzyme Activity is shared under a CC BY license and was authored, remixed, and/or curated by Gary Kaiser.

7.1: Control of Metabolism Through Enzyme Regulation

Control of Metabolism Through Enzyme Regulation

Cellular needs and conditions vary from cell to cell and change within individual cells over time. For example, a stomach cell requires a different amount of energy than a skin cell, fat storage cell, blood cell, or nerve cell. The same stomach cell may also need more energy immediately after a meal and less energy between meals.

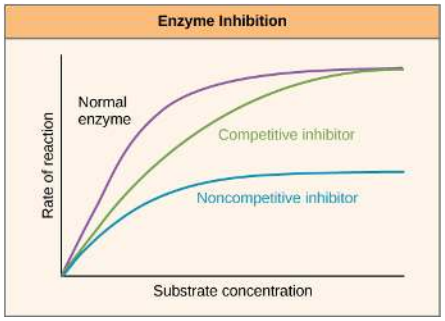


Figure 7.1.1: Competitive and noncompetitive inhibition affect the rate of reaction differently. Competitive inhibitors affect the initial rate, but do not affect the maximal rate, whereas noncompetitive inhibitors affect the maximal rate.

A cell's function is encapsulated by the chemical reactions it can carry out. Enzymes lower the activation energies of chemical reactions; in cells, they promote those reactions that are specific to the cell's function. Because enzymes ultimately determine which chemical reactions a cell can carry out and the rate at which they can proceed, they are key to cell functionality.

Competitive and Noncompetitive Inhibition

The cell uses specific molecules to regulate enzymes in order to promote or inhibit certain chemical reactions. Sometimes it is necessary to inhibit an enzyme to reduce a reaction rate, and there is more than one way for this inhibition to occur. In competitive inhibition, an inhibitor molecule is similar enough to a substrate that it can bind to the enzyme's active site to stop it from binding to the substrate. It "competes" with the substrate to bind to the enzyme.

In noncompetitive inhibition, an inhibitor molecule binds to the enzyme at a location other than the active site (an allosteric site). The substrate can still bind to the enzyme, but the inhibitor changes the shape of the enzyme so it is no longer in optimal position to catalyze the reaction.

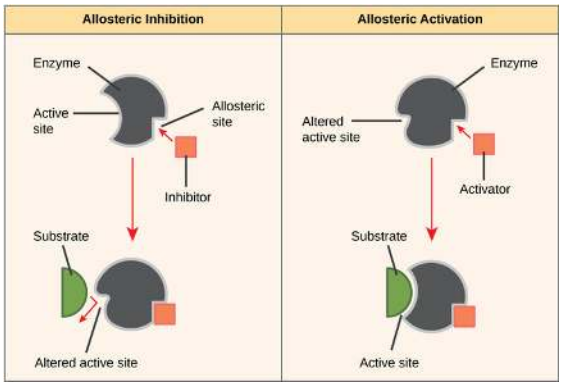


Figure 7.1.2: Allosteric inhibitors modify the active site of the enzyme so that substrate binding is reduced or prevented. In contrast, allosteric activators modify the active site of the enzyme so that the affinity for the substrate increases.

Allosteric Inhibition and Activation

In noncompetitive allosteric inhibition, inhibitor molecules bind to an enzyme at the allosteric site. Their binding induces a conformational change that reduces the affinity of the enzyme's active site for its substrate. The binding of this allosteric inhibitor changes the conformation of the enzyme and its active site, so the substrate is not able to bind. This prevents the enzyme from lowering the activation energy of the reaction, and the reaction rate is reduced.

However, allosteric inhibitors are not the only molecules that bind to allosteric sites. Allosteric activators can increase reaction rates. They bind to an allosteric site which induces a conformational change that *increases* the affinity of the enzyme's active site for its substrate. This increases the reaction rate.

Cofactors and Coenzymes

Many enzymes only work if bound to non-protein helper molecules called cofactors and coenzymes. Binding to these molecules promotes optimal conformation and function for their respective enzymes. These molecules bind temporarily through ionic or hydrogen bonds or permanently through stronger covalent bonds.

Cofactors are inorganic ions such as iron (Fe^{2+}) and magnesium (Mg^{2+}). For example, DNA polymerase requires a zinc ion (Zn^{2+}) to build DNA molecules. **Coenzymes** are organic helper molecules with a basic atomic structure made up of carbon and hydrogen.

The most common coenzymes are dietary vitamins. Vitamin C is a coenzyme for multiple enzymes that take part in building collagen, an important component of connective tissue. Pyruvate dehydrogenase is a complex of several enzymes that requires one cofactor and five different organic coenzymes to catalyze its chemical reaction. The availability of various cofactors and coenzymes regulates enzyme function.

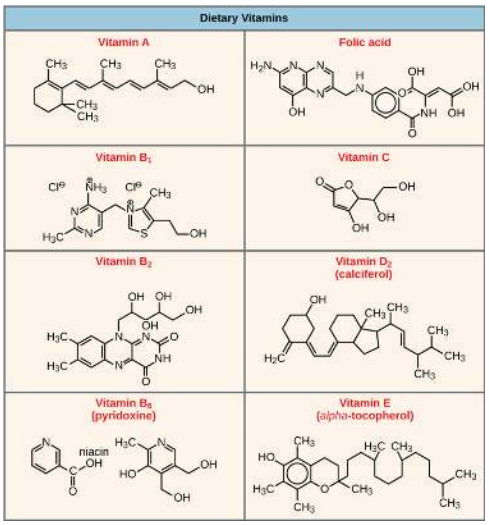


Figure 7.1.3: Vitamins are important coenzymes or precursors of coenzymes and are required for enzymes to function properly. Multivitamin capsules usually contain mixtures of all the vitamins at different percentages.

Enzyme Compartmentalization

In eukaryotic cells, molecules such as enzymes are usually compartmentalized into different organelles. This organization contributes to enzyme regulation because certain cellular processes are contained in separate organelles. For example, the enzymes involved in the later stages of cellular respiration carry out reactions exclusively in the mitochondria. The enzymes involved in the digestion of cellular debris and foreign materials are located within lysosomes.

Feedback Inhibition in Metabolic Pathways

Feedback inhibition is when a reaction product is used to regulate its own further production. Cells have evolved to use feedback inhibition to regulate enzyme activity in metabolism, by using the products of the enzymatic reactions to inhibit further enzyme activity. Metabolic reactions, such as anabolic and catabolic processes, must proceed according to the demands of the cell. In order to maintain chemical equilibrium and meet the needs of the cell, some metabolic products inhibit the enzymes in the chemical pathway while some reactants activate them.

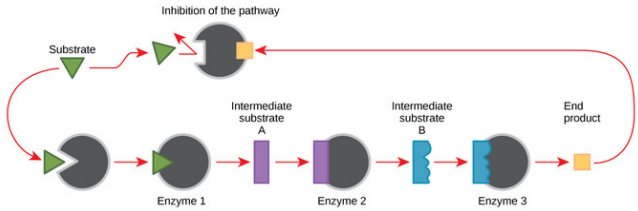


Figure 7.1.3 **:** Metabolic pathways are a series of reactions catalyzed by multiple enzymes. Feedback inhibition, where the end product of the pathway inhibits an earlier step, is an important regulatory mechanism in cells.

The production of both amino acids and nucleotides is controlled through feedback inhibition. For an example of feedback inhibition, consider ATP. It is the product of the catabolic metabolism of sugar (cellular respiration), but it also acts as an allosteric regulator for the same enzymes that produced it. ATP is an unstable molecule that can spontaneously dissociate into ADP; if too much ATP were present, most of it would go to waste. This feedback inhibition prevents the production of additional ATP if it is already abundant. However, while ATP is an inhibitor, ADP is an allosteric activator. When levels of ADP are high compared to ATP levels, ADP triggers the catabolism of sugar to produce more ATP.

Key Points

- In competitive inhibition, an inhibitor molecule competes with a substrate by binding to the enzyme ‘s active site so the substrate is blocked.
- In noncompetitive inhibition (also known as allosteric inhibition), an inhibitor binds to an allosteric site; the substrate can still bind to the enzyme, but the enzyme is no longer in optimal position to catalyze the reaction.
- Allosteric inhibitors induce a conformational change that changes the shape of the active site and reduces the affinity of the enzyme’s active site for its substrate.
- Allosteric activators induce a conformational change that changes the shape of the active site and increases the affinity of the enzyme’s active site for its substrate.
- Feedback inhibition involves the use of a reaction product to regulate its own further production.
- Inorganic cofactors and organic coenzymes promote optimal enzyme orientation and function.
- Vitamins act as coenzymes (or precursors to coenzymes) and are necessary for enzymes to function.

Key Terms

- **coenzyme**: An organic molecule that is necessary for an enzyme to function.
- **allosteric site**: A site other than the active site on an enzyme.
- **cofactor**: An inorganic molecule that is necessary for an enzyme to function.

This page titled 7.1: Control of Metabolism Through Enzyme Regulation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Boundless.

7.2: Amino Acids, Proteins, and Enzymes (Summary)

A **protein** is a large biological polymer synthesized from **amino acids**, which are carboxylic acids containing an α-amino group. Proteins have a variety of important roles in living organisms, yet they are made from the same 20 L-amino acids. About half of these amino acids, the **essential amino acids**, cannot be synthesized by the human body and must be obtained from the diet. In the solid state and in neutral solutions, amino acids exist as **zwitterions**, species that are charged but electrically neutral. In this form, they behave much like inorganic salts. Each amino acid belongs to one of four classes depending on the characteristics of its R group or amino acid side chain: nonpolar, polar but neutral, positively charged, and negatively charged. Depending on the conditions, amino acids can act as either acids or bases, which means that proteins act as buffers. The pH at which an amino acid exists as the zwitterion is called the **isoelectric point (pI)**.

The amino acids in a protein are linked together by **peptide bonds**. Protein chains containing 10 or fewer amino acids are usually referred to as **peptides**, with a prefix such as di- or tri- indicating the number of amino acids. Chains containing more than 50 amino acid units are referred to as *proteins* or **polypeptides**. Proteins are classified globular or fibrous, depending on their structure and resulting solubility in water. **Globular proteins** are nearly spherical and are soluble in water; fibrous proteins have elongated or fibrous structures and are not soluble in water.

Protein molecules can have as many as four levels of structure. The **primary structure** is the sequence of amino acids in the chain. The **secondary structure** is the arrangement of adjacent atoms in the peptide chain; the most common arrangements are α-helices or β-pleated sheets. The **tertiary structure** is the overall three-dimensional shape of the molecule that results from the way the chain bends and folds in on itself. Proteins that consist of more than one chain have **quaternary structure**, which is the way the multiple chains are packed together.

Four types of intramolecular and intermolecular forces contribute to secondary, tertiary, and quaternary structure: (1) **hydrogen bonding** between an oxygen or a nitrogen atom and a hydrogen atom bound to an oxygen atom or a nitrogen atom, either on the same chain or on a neighboring chain; (2) **ionic bonding** between one positively charged side chain and one negatively charged side chain; (3) **disulfide linkages** between cysteine units; and (4) **dispersion forces** between nonpolar side chains.

Because of their complexity, protein molecules are delicate and easy to disrupt. A *denatured* protein is one whose conformation has been changed, in a process called **denaturation**, so that it can no longer do its physiological job. A variety of conditions, such as heat, ultraviolet radiation, the addition of organic compounds, or changes in pH can denature a protein.

An **enzyme** is an organic catalyst produced by a living cell. Enzymes are such powerful catalysts that the reactions they promote occur rapidly at body temperature. Without the help of enzymes, these reactions would require high temperatures and long reaction times.

The molecule or molecules on which an enzyme acts are called its **substrates**. An enzyme has an **active site** where its substrate or substrates bind to form an enzyme-substrate complex. The reaction occurs, and product is released:

$$E + S \rightarrow [ES] \rightarrow E + P \tag{7.2.1}$$

The original **lock-and-key model** of enzyme and substrate binding pictured a rigid enzyme of unchanging configuration binding to the appropriate substrate. The newer **induced-fit model** describes the enzyme active site as changing its conformation after binding to the substrate.

Most enzymes have maximal activity in a narrow pH range centered on an **optimum pH**. In this pH range, the enzyme is correctly folded, and catalytic groups in the active site have the correct charge (positive, negative, or neutral). For most enzymes, the optimum pH is between 6 and 8.

Substances that interfere with enzyme function are called inhibitors. An **irreversible inhibitor** inactivates enzymes by forming covalent bonds to the enzyme, while a **reversible inhibitor** inactivates an enzyme by a weaker, noncovalent interaction that is easier to disrupt. A **competitive inhibitor** is a reversible inhibitor that is structurally similar to the substrate and binds to the active site. When the inhibitor is bound, the substrate is blocked from the active site and no reaction occurs. Because the binding of such an inhibitor is reversible, a high substrate concentration will overcome the inhibition because it increases the likelihood of the substrate binding. A **noncompetitive inhibitor** binds reversibly at a site distinct from the active site. Thus, it can bind to either the enzyme or the enzyme-substrate complex. The inhibitor changes the conformation of the active site so that the enzyme cannot

function properly. Noncompetitive inhibitors are important in **feedback inhibition**, in which the amount of product produced by a series of reactions is carefully controlled. The final product in a series of reactions acts as a noncompetitive inhibitor of the initial enzyme.

Simple enzymes consist entirely of one or more amino acid chains. Complex enzymes are composed of one or more amino acid chains joined to **cofactors**—inorganic ions or organic **coenzymes**. **Vitamins** are organic compounds that are essential in very small amounts for the maintenance of normal metabolism and generally cannot be synthesized at adequate levels by the body. Vitamins are divided into two broad categories: *fat-soluble* vitamins and *water-soluble* vitamins. Many of the water-soluble vitamins are used for the synthesis of coenzymes.

7.2: Amino Acids, Proteins, and Enzymes (Summary) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

7.3: Exercises

Multiple Choice

An enzyme has a maximum velocity expressed as V_{max}. What is the initial velocity of an enzyme reaction when the substrate concentration is at 20 times the K_m?

a. 0.2 V_{max}

b. 0.95 V_{max}

c. 1.5 V_{max}

d. 10 V_{max}

e. Insufficient information to determine V.

b: apply M-M equation.

Methanol and ethanol are both catalyzed by alcohol dehydrogenase (ADH). After ADH catalysis, methanol becomes formaldehyde, which is highly toxic. For a person with 60L body water where methanol and ethanol can rapidly distribute. Assume the K_m of ADH for both ethanol and methanol is 10^{-5} M, which approximately equal to their K_d. The density of both methanol and ethanol are 0.8 g/mL. If the person consumed 20 mL of methanol (50 - 100 mL can be fatal to a full grown adult) and 50% ethanol is used to compete with methanol. How much 50% ethanol has to be consumed to have ~ 50% of ADH working on methanol? (methanol has a molar mass of 32 g/mol; ethanol has a molar mass of 46 g/mol

a. 300 mL

b. 100 mL

c. 60 mL

d. 30 mL

e. 20 mL

c: if methanol was distributed equally in 60L, it has a concentration of (20 mL x 0.8) g/mL / (32g/mol x 60 L) = 8 x 10^{-3} M, which is >> than Km. This suggest that ~ 100 % ADH is working on methonal at this moment. To have ~50% ADH working on methanol, which has the same Km to ethonal, you will need equal molar amount off ethonal taken.

Two people have the same k after drinking alcohol. Please calculate the substrate concentration in Person 1.

Aldehyde dehydrogenase	Person 1	Person 2
k_{cat}(S^{-1})	10	1
K_m (M)	0.1	0.05
[S] (M)	?	0.05

a. 0.005

b. 0.01

c. 0.05

d. 0.1

e. 0.5

a. Use M-M equation

Which of the following describes the Bohr effect correctly?

a. the ability of hemoglobin to retain O_2 in the lungs

b. the regulation of hemoglobin's oxygen binding affinity by protons and CO_2

c. the alteration of hemoglobin conformation explained by sequential or concerted model

d. All of the above are correct.

e. None of the above is correct.

b

How is specificity determined by chymotrypsin?

a. interaction of the catalytic amino acids with the substrate

b. binding of the N-terminal amino acid at the active site

c. covalent binding of a His residue to the substrate

d. large conformational change of a P-loop upon binding of substrate

e. binding of the proper amino acid into a deep pocket of the enzyme

e

Restriction endonucleases are found in prokaryotic organisms to defend against foreign DNA. How does the host protect its own DNA from cleavage?

a. packaging DNA in nucleus

b. deamination of its DNA

c. methylation of its DNA

d. regulation of endonucleases activity with inhibitors

e. all of above are correct

c

2,3-BPG regulates the oxygen affinity for hemoglobin. 2,3-BPG:

a. promotes the transition to T-state.

b. binds to the C-terminal carboxylate group of hemoglobin.

c. preferentially binds to oxyhemoglobin to stabilize it

d. binds at hydrophobic patches of α-chains.

e. binds in an amount up to four molecules per hemoglobin molecule.

a

Which type of enzymes displays sigmoidal reaction kinetics (a plot of velocity versus [substrate])?

a. isozymes

b. apoenzymes

c. allosteric enzymes

d. holoenzyme

e. all enzymes

c

a.	
____un-competitive inhibitor	a. does not change K_m
____competitive inhibitor	b. does not change V_{max}
____suicide inhibitor	c. does not change V_{max}/K_m
____feedback inhibition	d. forms covalent bond with enzyme
____non-competitive inhibitor	e. inhibition by product

cbdea

b.	
____ Vitamin K	a. protein kinase A activation
____ ATP	b. part of heme
____ cAMP	c. O_2 binding affinity
____ porphyrin	d. thrombin binding
____2,3 BPG	e. phosphate carrier

deabc

Sample Questions

1. The form in which the ping pong mechanism binds substrates is identified as which type of mechanism?
2. What are two characteristics of an enzyme that catalyzes a reaction through the ping-pong mechanism?
3. What are the characteristics of an irreversible inhibitor?
4. In what ways does a competitive inhibitor differ from a noncompetitive inhibitor?
5. The activity of a purified enzyme is measured at a substrate concentration of 1.0 μM and found to convert 49 μmol of substrate to product in 1 min. The activity is measured at 2.0 μM substrate and found to convert 98 μmol of substrate to product/minute.

a) When the substrate concentration is 100 μM, how much substrate would you predict is converted to product in 1 min? What if the substrate concentration were increased to 1,000 μM (1.0 mM)?

b) The activities actually measured are 676 μmol product formed/minute at a substrate concentration of 100 μM and 698 μmol product formed/minute at 1,000 μM (1.0 mM) substrate. Is there any discrepancy between these values and those you predicted in Exercise 15a? Explain.

Ans a) at 100 μM, you would predict that the rate would increase 100 times to 4,900 μmol of substrate to product in 1 min; at 1.0 mM, you would predict an increase to 49,000 μmol of substrate to product in 1 min.

Ans b) There is a great discrepancy between the predicted rates and actual rates; this occurs because the enzyme becomes saturated with substrate, preventing a further increase in the rate of the reaction (the reaction is no longer linear with respect to substrate concentration because it is at very low concentrations).

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

CHAPTER OVERVIEW

8: Carbohydrate Structures, Stereochemistry, and Glycosides

Carbohydrates, also known as sugars, are found in all living organisms. They are essential to the very source of life (ex. Ribose sugars in DNA and RNA) or sustaining life itself (ex. Metabolic conversion of carbohydrates into usable biochemical energy, ATP). Another important role of carbohydrates is structural (ex. Cellulose in plants).

8.1: Carbohydrates Fundamentals

8.2: Monosaccharides

8.3: Disaccharides

8.4: Oligosaccharides

8.5: Polysaccharides

8.6: Exercises

Index

8: Carbohydrate Structures, Stereochemistry, and Glycosides is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8.1: Carbohydrates Fundamentals

Carbohydrates, also known as sugars, are found in all living organisms. They are **essential to the very source of life** (ex. Ribose sugars in DNA and RNA) or **sustaining life itself** (e.g., Metabolic conversion of carbohydrates into usable biochemical energy, ATP). Another important role of carbohydrates is **structural** (ex. Cellulose in plants). General names for carbohydrates include sugars, starches, saccharides, and polysaccharides. The term saccharide is derived from the Latin word "sacchararum" from the sweet taste of sugars. The name "carbohydrate" means a "hydrate of carbon." The name derives from the general formula of carbohydrate is $C_x(H_2O)_y$ - x and y may or may not be equal and range in value from 3 to 12 or more. For example glucose is: $C_6(H_2O)_6$ or is more commonly written, $C_6H_{12}O_6$.

Table 1: Common Carbohydrates

Name	Derivation of name and Source
Monosaccharides	
Glucose	From Greek word for sweet wine; grape sugar, blood sugar, dextrose.
Galactose	Greek word for milk--"galact", found as a component of lactose in milk.
Fructose	Latin word for fruit--"fructus", also known as levulose,found in fruits and honey; sweetest sugar.
Ribose	Ribose and Deoxyribose are found in the backbone structure of RNA and DNA, respectively.
Disaccharides - contain two monosaccharides	
Sucrose	French word for sugar--"sucre", a disaccharide containing **glucose and fructose**; table sugar, cane sugar, beet sugar.
Lactose	Latin word for milk--"lact"; a disaccharide found in milk containing **glucose and galactose**.
Maltose	French word for "malt"; a disaccharide containing **two units of glucose**; found in germinating grains, used to make beer.
Common Polysaccharides	
Starch	Plants store glucose as the polysaccharide starch. The cereal grains (wheat, rice, corn, oats, barley) as well as tubers such as potatoes are rich in starch.
Cellulose	The major component in the rigid cell walls in plants is cellulose and is a linear polysaccharide polymer with many glucose monosaccharide units.
Glycogen	This is the storage form of glucose in animals and humans which is analogous to the starch in plants. **Glycogen is synthesized and stored mainly in the liver and the muscles**.

Contributors

Charles Ophardt, Professor Emeritus, Elmhurst College; Virtual Chembook

Template:HideTOC

8.1: Carbohydrates Fundamentals is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

8.2: Monosaccharides

A monosaccharide is a carbohydrate consisting of one sugar unit. Common examples of simple sugars or monosaccharides are glucose and fructose. Both of these monosaccharides are referred to as hexoses since they have six carbons. Glucose is abundant in many plant sources and makes up sweeteners such as corn sugar or grape sugar. Fructose found in honey and fruits. These sugars are structural isomers of one another, with the difference being that glucose contains an aldehyde functional group whereas fructose contains a ketone functional group.

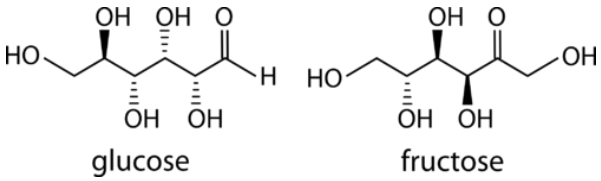


Figure 8.2.1:Glucose and fructose are monosaccharides, or simple sugars.

Glucose and fructose are both soluble in water. In aqueous solution, the predominant forms are not the straight-chain structure shown above. Rather, they adopt a cyclic structure (see figure below). Glucose is six membered ring, while fructose is a five-membered ring. Both rings contain an oxygen atom.

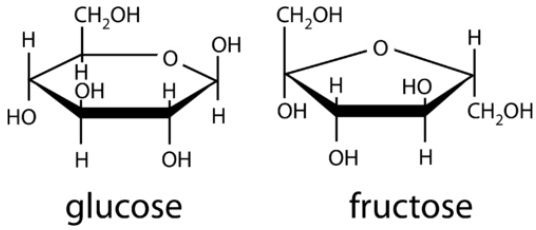


Figure 8.2.1: The cyclic form of sugars is the favored form in aqueous solution.

Glucose and other 5C and 6C sugars can cyclize through intramolecular nucleophilic attack of one of the OH's on the carbonyl C 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 which points either below the ring (a anomer) or above the ring (b anomer).

Monosaccharides in solution exist as equilibrium mixtures of the straight and cyclic forms. In solution, glucose is mostly in the pyranose form, fructose is 67% pyranose and 33% furanose, and ribose is 75% furanose and 25% pyranose.

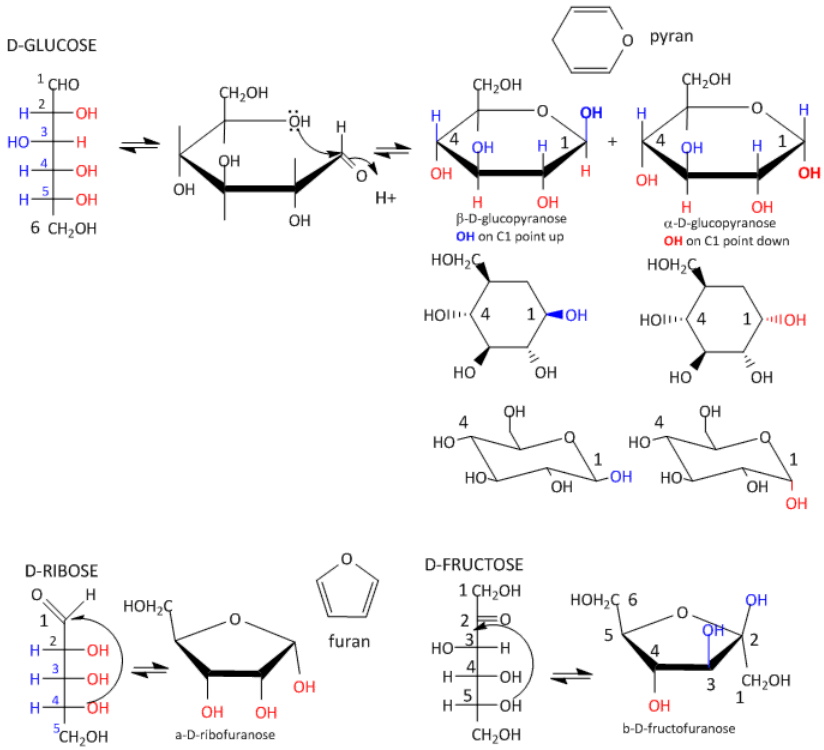
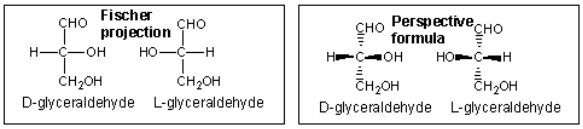


Figure 8.2.3: Sugar Ring Formation and Representations

Sugars can be drawn in the straight chain form as either Fisher projections or perspective structural formulas.



In the Fisher projection, the vertical bonds point down into the plane of the paper. That's easy to visualize for 3C molecules. but more complicated for bigger molecules. 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 Fisher project, it should be pointing to the right in the wedge and dash drawing, as shown below for D-erthyrose and D-glucose.

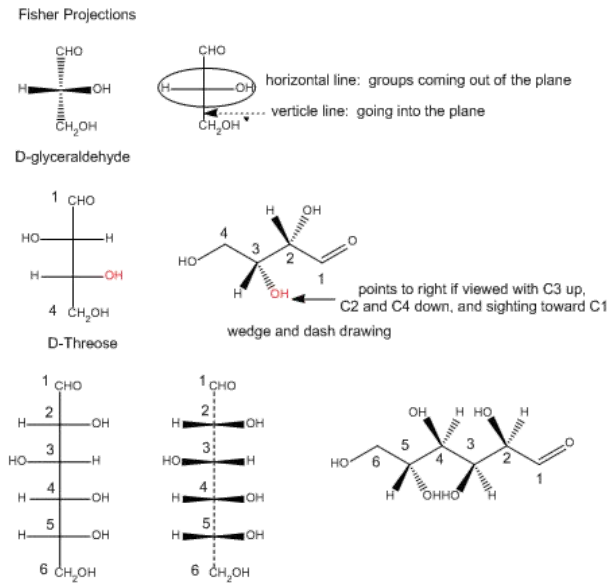


Figure 8.2.4: Orienting OH groups in wedge and dashing drawings of simple straight chain sugars

Cyclic forms can be drawn either as the **Haworth projections**, which shows 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, for example, b-D-glucopyranose is the only aldohexose which can be drawn with all its bulky substituents (OH and CH_2OH) in equatorial positions, which probably accounts for its widespread prevalence in nature.

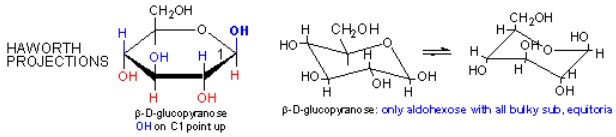
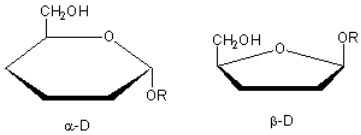


Figure 8.2.5: *Haworth projections*

Haworth projections are more realistic than the Fisher projections, but you should be able to draw both structures. In general, if a substituent points to the right in the Fisher structure, it points down in the Haworth. if it points left, it points up. In general, the OH on the a-anomer points down (ants down) while on the b-anomer it points up (butterflies up).



Another important group of monosaccharides are the **pentoses,** containing five carbons in the chain. Ribose and deoxyribose are two pentoses that are components of the structures of DNA and RNA.

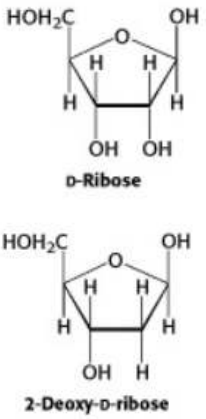


Figure 8.2.6: *Ribose and deoxyribose*

Glycosides

Acetal derivatives formed when a monosaccharide reacts with an alcohol in the presence of an acid catalyst are called glycosides. This reaction is illustrated for glucose and methanol in the diagram below. In naming of glycosides, the "ose" suffix of the sugar name is replaced by "oside", and the alcohol group name is placed first. As is generally true for most acetals, glycoside formation involves the loss of an equivalent of water. The diether product is stable to base and alkaline oxidants such as Tollen's reagent. Since acid-catalyzed aldolization is reversible, glycosides may be hydrolyzed back to their alcohol and sugar components by aqueous acid.

The anomeric methyl glucosides are formed in an equilibrium ratio of 66% alpha to 34% beta. From the structures in the previous diagram, we see that pyranose rings prefer chair conformations in which the largest number of substituents are equatorial. In the case of glucose, the substituents on the beta-anomer are all equatorial, whereas the C-1 substituent in the alpha-anomer changes to axial. Since substituents on cyclohexane rings prefer an equatorial location over axial (methoxycyclohexane is 75% equatorial), the preference for alpha-glycopyranoside formation is unexpected, and is referred to as the anomeric effect.

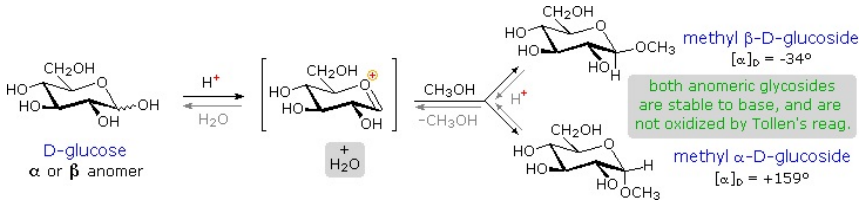


Figure 8.2.7: *Formation of acetal derivatives by the action of glycosides on glucose and methanol.*

Glycosides abound in biological systems. By attaching a sugar moiety to a lipid or benzenoid structure, the solubility and other properties of the compound may be changed substantially. Because of the important modifying influence of such derivatization, numerous enzyme systems, known as glycosidases, have evolved for the attachment and removal of sugars from alcohols, phenols and amines. Chemists refer to the sugar component of natural glycosides as the glycon and the alcohol component as the aglycon.

Two examples of naturally occurring glycosides and one example of an amino derivative are displayed above. Salicin, one of the oldest herbal remedies known, was the model for the synthetic analgesic aspirin. A large class of hydroxylated, aromatic oxonium cations called anthocyanins provide the red, purple and blue colors of many flowers, fruits and some vegetables. Peonin is one example of this class of natural pigments, which exhibit a pronounced pH color dependence. The oxonium moiety is only stable in acidic environments, and the color changes or disappears when base is added. The complex changes that occur when wine is fermented and stored are in part associated with glycosides of anthocyanins. Finally, amino derivatives of ribose, such as cytidine play important roles in biological phosphorylating agents, coenzymes and information transport and storage materials.

Boat/Chair Conformations

Independent of stereoisomerization, sugars in ring form of a given type (such as glucose) can “twist" themselves into alternative conformations called boat and chair. Note that this rearrangement does not change the relative positions of hydroxyl groups. All that has changed is the shape of the molecule. As shown for glucose, one can see that the beta-hydroxyl of glucose is closer to the

CH_2OHCH_2OH (carbon #6) in the boat form than it is in the chair form. Steric hindrance can be a factor in favoring one configuration over another.

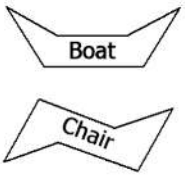


Figure 8.2.8: *Boat and Chair*

Contributors

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

Prof. Steven Farmer (Sonoma State University)

William Reusch, Professor Emeritus (Michigan State U.), Virtual Textbook of Organic Chemistry

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

Thumbnail Ball-and-stick model of the α-D-glucose molecule, $C_6H_{12}O_6$. Image used with permission (Public Domain; Ben Mills).

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

8.3: Disaccharides

Two monosaccharide molecules may chemically bond to form a disaccharide. The name given to the covalent bond between the two monosaccharides is a **glycosidic bond**. Glycosidic bonds form between hydroxyl groups of the two saccharide molecules.

Common disaccharides are the grain sugar maltose, made of two glucose molecules; the milk sugar lactose, made of a galactose and a glucose molecule; and the table sugar sucrose, made of a glucose and a fructose molecule (Figure 8.3.1).

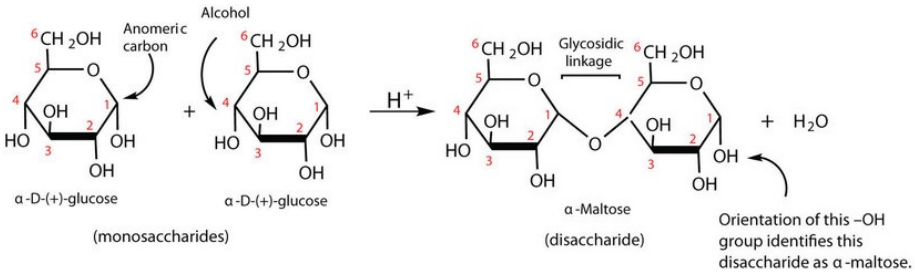


Figure 8.3.1: Formation of disaccharide from monosaccharide through glucosidic linkage

Maltose

Maltose occurs to a limited extent in sprouting grain. It is formed most often by the partial hydrolysis of starch and glycogen. In the manufacture of beer, maltose is liberated by the action of malt (germinating barley) on starch; for this reason, it is often referred to as *malt sugar*. Maltose is about 30% as sweet as sucrose. The human body is unable to metabolize maltose or any other disaccharide directly from the diet because the molecules are too large to pass through the cell membranes of the intestinal wall. Therefore, an ingested disaccharide must first be broken down by hydrolysis into its two constituent monosaccharide units. In the body, such hydrolysis reactions are catalyzed by enzymes such as *maltase*.

Maltose is a **reducing sugar**. Thus, its two glucose molecules must be linked in such a way as to leave one anomeric carbon that can open to form an aldehyde group. The glucose units in maltose are joined in a *head-to-tail* fashion through an α-linkage from the first carbon atom of one glucose molecule to the fourth carbon atom of the second glucose molecule (that is, an α-1,4-glycosidic linkage). The bond from the anomeric carbon of the first monosaccharide unit is directed downward, which is why this is known as an α-glycosidic linkage. The OH group on the anomeric carbon of the second glucose can be in either the α or the β position, as shown in Figure 8.3.2.

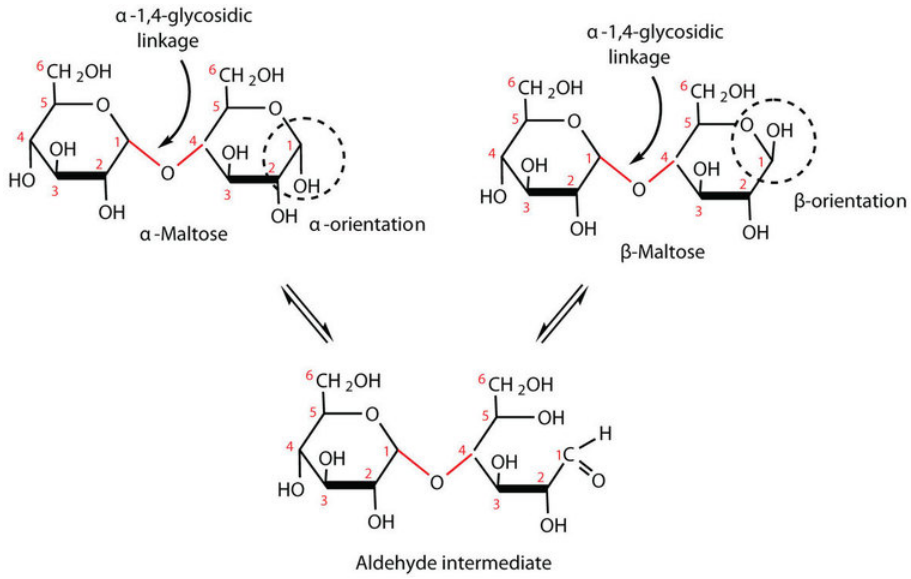


Figure 8.3.2: An Equilibrium Mixture of Maltose Isomers

Lactose

Lactose is known as *milk sugar* because it occurs in the milk of humans, cows, and other mammals. In fact, the natural synthesis of lactose occurs only in mammary tissue, whereas most other carbohydrates are plant products. Human milk contains about 7.5% lactose, and cow's milk contains about 4.5%. This sugar is one of the lowest ranking in terms of sweetness, being about one-sixth as sweet as sucrose. Lactose is produced commercially from whey, a by-product in the manufacture of cheese. It is important as an infant food and in the production of penicillin.

Lactose is a reducing sugar composed of one molecule of D-galactose and one molecule of D-glucose joined by a β-1,4-glycosidic bond (the bond from the anomeric carbon of the first monosaccharide unit being directed upward).

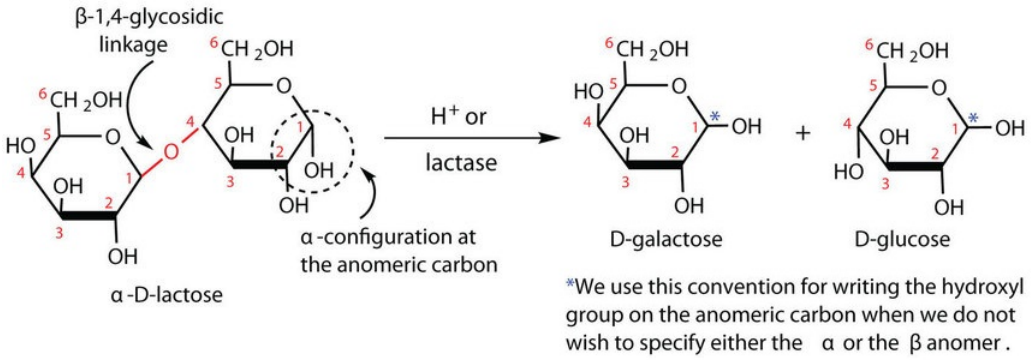


Figure 8.3.3: *The two monosaccharides are obtained from lactose by acid hydrolysis or the catalytic action of the enzyme lactase*

Many adults and some children suffer from a deficiency of **lactase.** These individuals are said to be lactose intolerant because they cannot digest the lactose found in milk. A more serious problem is the genetic disease **galactosemia,** which results from the absence of an enzyme needed to convert galactose to glucose. Certain bacteria can metabolize lactose, forming lactic acid as one of the products. This reaction is responsible for the "souring" of milk.

To Your Health:

Lactose makes up about 40% of an infant's diet during the first year of life. Infants and small children have one form of the enzyme lactase in their small intestines and can digest the sugar easily; however, adults usually have a less active form of the enzyme, and about 70% of the world's adult population has some deficiency in its production. As a result, many adults experience a reduction in the ability to hydrolyze lactose to galactose and glucose in their small intestine. For some people the inability to synthesize sufficient enzyme increases with age. Up to 20% of the US population suffers some degree of **lactose intolerance.**

Sucrose

Sucrose, probably the largest-selling pure organic compound in the world, is known as *beet sugar*, *cane sugar*, *table sugar*, or simply *sugar*. Most of the sucrose sold commercially is obtained from sugar cane and sugar beets (whose juices are 14%–20% sucrose) by evaporation of the water and recrystallization. The dark brown liquid that remains after the recrystallization of sugar is sold as molasses.

The sucrose molecule is unique among the common disaccharides in having an α-1,β-2-glycosidic (head-to-head) linkage. Because this glycosidic linkage is formed by the OH group on the anomeric carbon of α-D-glucose and the OH group on the anomeric carbon of β-D-fructose, it ties up the anomeric carbons of both glucose and fructose.

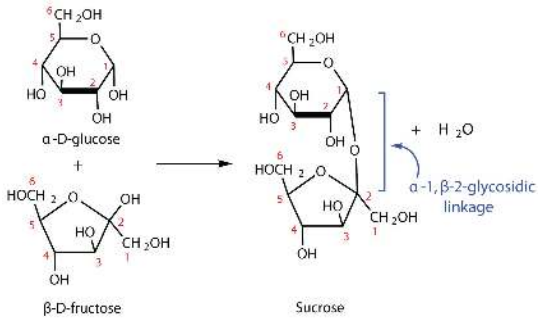


Figure 8.3.4: formation of α-1,β-2-glycosidic (head-to-head) linkage between the OH group on the anomeric carbon of α-D-glucose and the OH group on the anomeric carbon of β-D-fructose

This linkage gives sucrose certain properties that are quite different from those of maltose and lactose. As long as the sucrose molecule remains intact, neither monosaccharide "uncyclizes" to form an open-chain structure. Thus, sucrose is incapable of **mutarotation** and exists in only one form both in the solid state and in solution. In addition, sucrose does not undergo reactions that are typical of aldehydes and ketones. Therefore, sucrose is a nonreducing sugar.

The hydrolysis of sucrose in dilute acid or through the action of the enzyme *sucrase* (also known as invertase) gives an equimolar mixture of glucose and fructose. This 1:1 mixture is referred to as *invert sugar* because it rotates plane-polarized light in the opposite direction than sucrose. The hydrolysis reaction has several practical applications. Sucrose readily recrystallizes from a solution, but invert sugar has a much greater tendency to remain in solution. In the manufacture of jelly and candy and in the canning of fruit, the recrystallization of sugar is undesirable. Therefore, conditions leading to the hydrolysis of sucrose are employed in these processes. Moreover, because fructose is sweeter than sucrose, the hydrolysis adds to the sweetening effect. Bees carry out this reaction when they make honey.

The average American consumes more than 100 lb of sucrose every year. About two-thirds of this amount is ingested in soft drinks, presweetened cereals, and other highly processed foods. The widespread use of sucrose is a contributing factor to obesity and tooth decay. Carbohydrates such as sucrose, are converted to fat when the caloric intake exceeds the body's requirements, and sucrose causes tooth decay by promoting the formation of plaque that sticks to teeth.

Contributors

Prof. Steven Farmer (Sonoma State University)

William Reusch, Professor Emeritus (Michigan State U.), Virtual Textbook of Organic Chemistry

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

SECTION OVERVIEW

8.4: Oligosaccharides

An oligosaccharide is a carbohydrate whose molecule, upon hydrolysis, yields two to ten Monosaccharid molecules. Oligosaccharides are classified into subclasses based on the number of monosaccharide molecules that form when one molecule of the oligosaccharide is hydrolyzed.

Oligosaccharides are not commonly found free in cells, but instead are found **covalently attached to proteins**, which are then said to be glycosylated. Oligosaccharides attached to proteins may be **N-linked** (through **asparagine**) or **O-linked** (though **serine or threonine**). O-linked sugars are added only in the Golgi apparatus while N-linked sugars are attached starting in the endoplasmic reticulum and then completed in the Golgi.

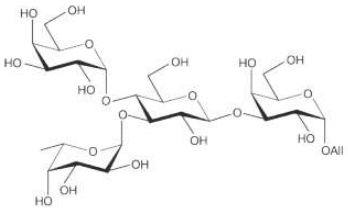


Figure 8.4.1: *An oligosaccharide.*

Glycoproteins

Membrane proteins are often covalently linked to **oligosaccharides**, which are branched *glycoside-linked* sugars (averaging around 15 sugar residues). As **glycans**, they are the sugars linked to **glycoproteins**. Glycoproteins are rare in the cytosol, but common on secreted and membrane proteins. Oligosaccharides are typically linked to proteins via the hydroxyl group on *serine* or *threonine*. Occasional linkages are to modified amino acids like hydroxylysine or hydroxyproline (*O-glycosylation*), and to the amide nitrogen on asparagine (*N-glycosylation*). The oligosaccharide domains of glycoproteins often play a major role in membrane protein function. For example, the glycoproteins, along with the polar domains of integral and peripheral proteins and glycolipids, are a major feature of the glycocalyx.

Glycosylation

Sugars are commonly attached to proteins in a process called **glycosylation.** Typically the attachment is to a hydroxyl or other functional group. The majority of proteins synthesized in the endoplasmic reticulum are glycosylated. Five classes of glycosylated products (called glycans if multiple carbohydrates are attached via glycosidic bonds) are known. They include:

• N-linked glycans - carbohydrate attached to N group of asparagine or arginine side chain

• O-linked glycans - carbohydrate attached to OH of serine, threonine, tyrosine, hydroxyproline, hydroxylysine, or lipids.

• Phosphoglycans - attachment to a phosphoserine

• Glypiation - linkage of a phosphatidyl inositol to link proteins to lipids via glycan linkages

• C-linked glycans - sugar attached to a carbon on a tryptophan side chain.

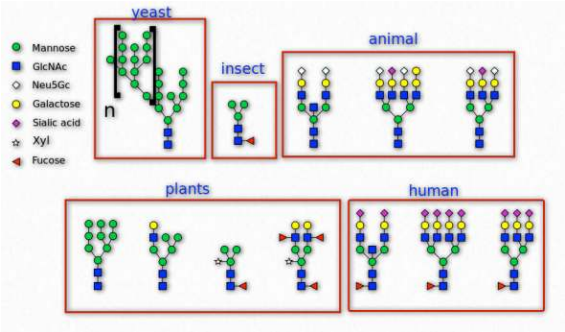


Figure 8.4.2: - *N-linked glycosylation in various organisms.* (https://en.wikipedia.org/wiki/N-linked_glycosylation)

Glycosylation has several molecular/ cellular functions. Some proteins require glycosylation to **fold properly** or to be stable. Glycosylated proteins on the plasma membrane serve as **cellular identifiers**. Blood types, for example, arise from differential glycosylation of a blood cell membrane protein. Glycosylation can also play an important role in **cell-cell adhesion** - important in the immune system.

Glycoprotein Function

The role of CHO in glycoprotein structure/function is slowly being determined. The most important seems to involve their role in directing proper folding of proteins in the endoplasmic reticulum (ER) which accounts for the observations that glycan addition to proteins in the ER is a cotranslational event. When inhibitors of ER glycosylation are added to cells, protein misfolding and aggregation are observed. The extent of misfolding depends on the particular protein and particular glycosylation sites with the protein. The polar CHO residues help promote solubility of folding intermediates, similar to the effects of many chaperone proteins.

The glycan moieties of the folding glycoprotein also lead to binding of the protein to lectins in the ER which serve as molecular chaperones. The most studied of these chaperones are involved in the calnexin-calreticulin cycle, and facilitate correct disulfide bond formation in the protein. After two glucose residues are removed by glucosidase I and II, the monoglucosylated protein binds to calnexin (CNX) and/or calreticulin (CRT), two homologous ER lectins specific for monoglucosylated proteins. Once bound, another protein, ERp57, a molecular chaperone with a disulfide bond (shown in diagram) interacts with the protein. This protein has protein disulfide isomerase activity.

If a glycoprotein has not folded completely, it is recognized by a glycoprotein glucosyltransferase, which adds a glucose to it. This then promotes reentry into the calnexin/calreticulin cycle.

Ideally, unfolded or misfolded proteins would be targeted from degradation and elimination from cells. The ER has evolved a system to accomplish this. Since folding occurs in the ER, to prevent misfolding and aggregation, the ER also contains chaperones and folding catalysts. Stress (such as through heat shock) stimulates ER chaperone activity. As a final defense mechanism, unfolded or aberrantly-folded proteins are degraded by the cytoplasmic proteasome complex. Nonnative forms of some proteins that "escape" this surveillance system can accumulate and result in disease (for example neurodegenerative diseases like Alzheimers and Parkinson's disease.

Glycosylation versus Glycation

The sugars on glycoproteins have been placed there by **glycosylation** — a precise enzymatic activity that makes a product which would otherwise not function correctly. However, sugars can also spontaneously form covalent links to proteins (and lipids) — a process called **glycation**. No enzymes are involved so the process is quite random. The products are apt to have reduced, or even no, function. The glycation of proteins and lipids is an inevitable outcome of aging. It is hastened in diabetics with their high blood sugar (glucose) levels. In fact, measuring the amount of glycation of hemoglobin is an important test for determining how well diabetes is being controlled.

Peptidoglycan

Peptidoglycan, also known as murein, is a polymer consisting of sugars and amino acids that forms a mesh-like layer outside the plasma membrane of bacteria (but not Archaea), forming the cell wall. The sugar component consists of alternating residues of β-

(1,4) linked N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM). Attached to the N-acetylmuramic acid is a peptide chain of three to five amino acids. The peptide chain can be cross-linked to the peptide chain of another strand forming the 3D mesh-like layer.

The peptidoglycan layer is substantially thicker in Gram-positive bacteria (20 to 80 nanometers) than in Gram-negative bacteria (7 to 8 nanometers), with the attachment of the S-layer. Peptidoglycan forms around 90% of the dry weight of Gram-positive bacteria but only 10% of Gram-negative strains. Thus, presence of high levels of peptidoglycan is the primary determinant of the characterization of bacteria as gram-positive. In Gram-positive strains, it is important in attachment roles and stereotyping purposes. For both Gram-positive and Gram-negative bacteria, particles of approximately 2 nm can pass through the peptidoglycan. Gram-positive and Gram-negative bacteria are sensitive to different types of antiobiotics.

Structure and Composition of Peptidoglycan

Peptidoglycan, also known as murein, is a polymer consisting of sugars and amino acids that forms a mesh-like layer outside the plasma membrane of bacteria (but not Archaea), forming the cell wall. The sugar component consists of alternating residues of β-(1,4) linked N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM). Attached to the N-acetylmuramic acid is a peptide chain of three to five amino acids. The peptide chain can be cross-linked to the peptide chain of another strand forming the 3D mesh-like layer.

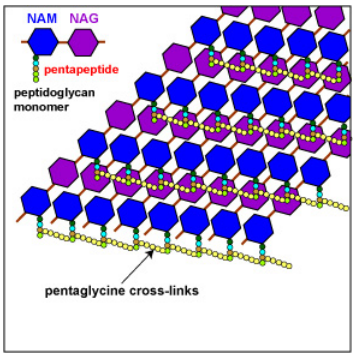


Figure 8.4.3: *Peptidoglycan is composed of cross-linked chains of peptidoglycan monomers (NAG-NAM-pentapeptide). Transglycosylase enzymes join these monomers join together to form chains. Transpeptidase enzymes then cross-link the chains to provide strength to the cell wall and enable the bacterium to resist osmotic lysis.*

Function of Peptidoglycan

Peptidoglycan prevents osmotic lysis. As seen earlier under the cytoplasmic membrane, bacteria concentrate dissolved nutrients (solute) through active transport. As a result, the bacterium's cytoplasm is usually hypertonic to its surrounding environment and the net flow of free water is into the bacterium. Without a strong cell wall, the bacterium would burst from the osmotic pressure of the water flowing into the cell.

Peptidoglycan serves a structural role in the bacterial cell wall giving it strength, as well as counteracting the osmotic pressure of the cytoplasm. A common misconception is that peptidoglycan gives the cell its shape. However, it is actually the MreB protein that facilitates cell shape. Peptidoglycan is also involved in binary fission during bacterial cell reproduction.

Antimicrobial Agents that Inhibit Peptidoglycan Synthesis Causing Bacterial Lysis

Many antibiotics work by inhibiting normal synthesis of peptidoglycan in bacteria causing them to burst as a result of osmotic lysis. In order for bacteria to increase their size following binary fission, enzymes called autolysins break the peptide cross links in the peptidoglycan, transglycosylase enzymes then insert and link new peptidoglycan monomers into the breaks in the peptidoglycan, and transpeptidase enzymes reform the peptide cross-links between the rows and layers of peptidoglycan to make the wall strong. Interference with this process results in a weak cell wall and lysis of the bacterium from osmotic pressure.

For example, penicillins and cephalosporins bind to the transpeptidase enzymes (also called penicillin-binding proteins) responsible for resealing the cell wall as new peptidoglycan monomers are added during bacterial cell growth. This blocks the transpeptidase enzymes from cross-linking the sugar chains and results in a weak cell wall and subsequent osmotic lysis of the bacterium.

Topic hierarchy

John W. Kimball. This content is distributed under a Creative Commons Attribution 3.0 Unported (CC BY 3.0) license and made possible by funding from The Saylor Foundation.

Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

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

8.5: Polysaccharides

Polysaccharides, as their name implies, are made by joining together many sugars. The functions for polysaccharides are varied. They include energy storage, structural strength, and lubrication. Polysaccharides involved in energy storage include the plant polysaccharides, amylose and amylopectin. The polysaccharide involved in energy storage in animals is called **Glycogen** and it is mostly found in the muscles and liver.

Amylose/Amylopectin

Amylose is the simplest of the polysaccharides, being comprised solely of glucose units joined in an alpha 1-4 linkage. Amylose is broken down by the enzyme alpha-amylase, found in saliva. Amylopectin is related to amylose in being composed only of glucose, but it differs in how the glucose units are joined together. Alpha 1-4 linkages predominate, but every 30-50 residues, a ‘branch’ arises from an alpha 1-6 linkage. Branches make the structure of amylopectin more complex than that of amylose.

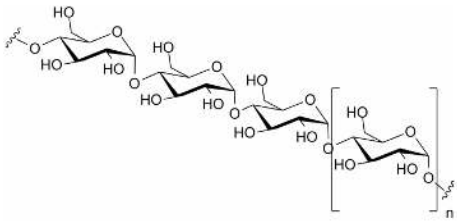


Figure 8.5.1: Amylose.

Glycogen

Glycogen is a polysaccharide that is physically related to amylopectin in being built only of glucose and in having a mix of alpha 1-4 and alpha 1-6 bonds. Glycogen, however, has many more alpha 1-6 branches than amylopectin, with such bonds occurring about every 10 residues. One might wonder why such branching occurs more abundantly in animals than in plants. A plausible explanation is based on the method by which these molecules are broken down. The breakdown of these polysaccharides is catalyzed by enzymes, known as phosphorylases, that clip glucose residues from the ends of glycogen chains and attach a phosphate to them in the process, producing glucose-1-phosphate. More highly branched polysaccharides have more ends to clip, and this translates to more glucose-1-phosphates that can be removed simultaneously by numerous phosphorylases. Since glucose is used for energy by muscles, glucose concentrations can be increased faster the more branched the glycogen is. Plants, which are immobile do not have needs for such immediate release of glucose and thus have less need for highly branched polysaccharides.

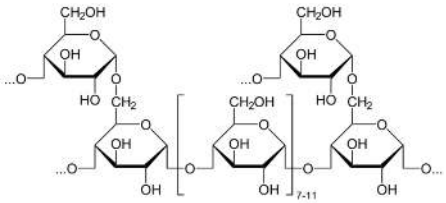


Figure 8.5.2: *Glycogen Structure*

Cellulose

Another important polysaccharide containing only glucose is cellulose. It is a polymer of glucose used to give plant cell walls structural integrity and has the individual units joined solely in a beta 1-4 configuration. That simple structural change makes a radical diff erence in its digestibility. Humans are unable to break down cellulose and it passes through the digestive system as roughage. Ruminant animals, such as cattle, however have bacteria in their rumens that contain the enzyme cellulase. It breaks the beta 1-4 links of the glucoses in cellulose to release the sugars for energy.

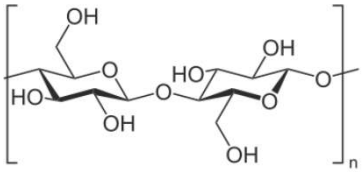


Figure 8.5.3: *The repeating structure of cellulose.*

Another polysaccharide used for structural integrity is known as **chitin**. Chitin makes up the exoskeleton of insects and is a polymer of a modified form of glucose known as N-acetyl-glucosamine.

Glycosaminoglycans

Yet another category of polysaccharides are the glycosaminoglycans (also called mucopolysaccharides), some examples of which include keratan sulfate, heparin, hyaluronic acid, and chondroitin sulfate. The polysaccharide compounds are linked to proteins, but differ from glycoproteins in having a much larger contingent of sugar residues and, further, the sugars are considerably more chemically modified. Each of them contains a repeating unit of a disaccharide that contains at least one negatively charged residue. The result is a polyanionic substance that, in its interactions with water, makes for a “slimy" feel. Glycosaminoglycans are found in snot, and in synovial fluid, which lubricates joints. Heparin is a glycosaminoglycan that helps to prevent blood from clotting.

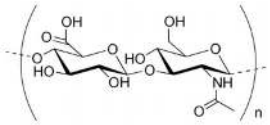


Figure 8.5.4: *The repeating unit of a glycosaminoglycan structure.*

Contributors

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

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

8.6: Exercises

Multiple Choice

1. By definition, carbohydrates contain which elements?

A. carbon and hydrogen
B. carbon, hydrogen, and nitrogen
C. carbon, hydrogen, and oxygen
D. carbon and oxygen

---

2. Monosaccharides may link together to form polysaccharides by forming which type of bond?

A. hydrogen
B. peptide
C. ionic
D. glycosidic

---

3. Which of the following is *not* a complex carbohydrate?

A. chitin

B. starch

C. disaccharide

D. none of the above

4. An example of a monosaccharide is _____.

1. fructose
2. glucose
3. galactose
4. all of the above

5. Cellulose and starch are examples of:

1. monosaccharides
2. disaccharides
3. lipids
4. polysaccharides

6. Plant cell walls contain which of the following in abundance?

1. starch
2. cellulose
3. glycogen
4. lactose

7. Lactose is a disaccharide formed by the formation of a _____ bond between glucose and _____.

1. glycosidic; lactose
2. glycosidic; galactose
3. hydrogen; sucrose
4. hydrogen; fructose

Matching

Match each polysaccharide with its description.

___chitin	A. energy storage polymer in plants
___glycogen	B. structural polymer found in plants
___starch	C. structural polymer found in cell walls of fungi and exoskeletons of some animals
___cellulose	D. energy storage polymer found in animal cells and bacteria

Short Answer

1.What are monosaccharides, disaccharides, and polysaccharides?

2. What are carbohydrates? Describe their structure.

3. Put the following carbohydrates in order from smallest to largest:

cellulose; fructose; sucrose

4. Name three carbohydrates that contain glucose as a monomer.

5.Which do you think is faster to digest — simple sugars or complex carbohydrates? Explain your answer.

6. What are the similarities and differences between muscle glycogen and liver glycogen?

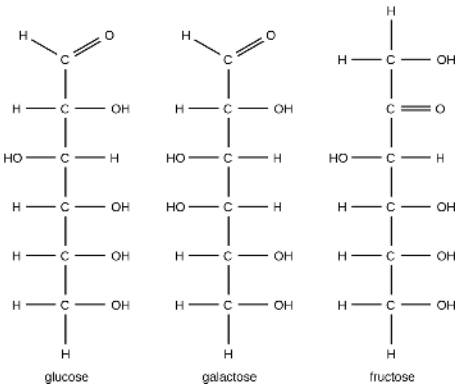
7. Which carbohydrate is used directly by the cells of living things for energy?

8. Why is it impossible for humans to digest food that contains cellulose?

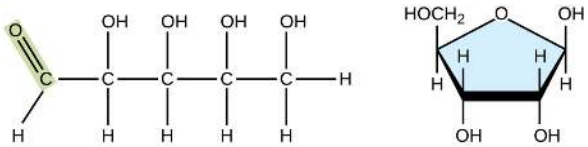
9. Describe the similarities and differences between glycogen and starch

Critical Thinking

1. The figure depicts the structural formulas of glucose, galactose, and fructose. (a) Circle the functional groups that classify the sugars either an aldose or a ketose, and identify each sugar as one or the other. (b) The chemical formula of these compounds is the same, although the structural formula is different. What are such compounds called?



2. Structural diagrams for the linear and cyclic forms of a monosaccharide are shown. (a) What is the molecular formula for this monosaccharide? (Count the C, H and O atoms in each to confirm that these two molecules have the same formula, and report this formula.) (b) Identify which hydroxyl group in the linear structure undergoes the ring-forming reaction with the carbonyl group.



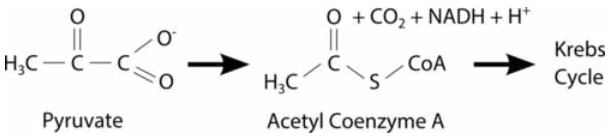
Contributor

- Template:ContribOpenSTAXMicrobiology

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

10: Pyruvate Dehydrogenase Links Glycolysis to Krebs Cycle

Aerobic respiration begins with the entry of the product of glycolysis, pyruvate, into the mitochondria. The mitochondria is an organelle in the cell. It is considered the "powerhouse" of the cell. Pyruvate is transported there via pyruvate translocase. For each initial glucose molecule, two pyruvate molecules will enter the mitochondria. Pyruvate, however, is not the molecule that enters the citric acid cycle. Prior to entry into this cycle, pyruvate must be converted into a 2-carbon acetyl-CoenzymeA (acetyl-CoA) unit.



Pyruvate dehydrogenase is a multi-enzyme complex that uses three enzymes:

1. E_1: Pyruvate dehydrogenase which uses **thiamine pyrophosphate (TPP)** as its cofactor.
2. E_2: Dihydrolipoyl transacetylase which uses **lipoamide** as its cofactor.
3. E_3: Dihydrolipoyl dehydrogenase which uses **flavin adenine dinucleotide (FAD)** as its cofactor.

The conversion of pyruvate into acetyl-CoA is referred to as the pyruvate dehydrogenase reaction. It is catalyzed by the **pyruvate dehydrogenase complex (PDH)**. This process produces one NADH electron carrier while releasing a CO_2 molecule. This step is also known as the *link reaction or transition step*, as it links glycolysis and the citric acid cycle. Of course, as two pyruvates result from glycolysis, two acetyl-CoAs are produced as are 2 NADH molecules.

Note: Prosthetic groups are molecules that are covalently bonded to an enzyme.

The net reaction of converting pyruvate into acetyl coA and CO_2 is:

$$2\,\text{pyruvate} + 2\,\text{NAD}^+ + 2\,\text{coA} \rightarrow 2\text{acetyl coA} + 2\,\text{NADH} + 2CO_2$$

The Process

This is a five step process.

1. **Step A:** Pyruvate is decarboxylated by pyruvate dehydrogenase with help from TPP.
2. **Step B**: The reactive carbon (between the N and the S of the five membered ring) of the TPP is oxidized and transferred as the acetyl group to lipoamide (which is the prosthetic group of the dihydrolipoyl transacetylase). This forms hydroxyethyl-TPP. An H^+ ion is required for the intermediate to give off CO_2.
3. **Step C**: E_2 (dihydrolipoyl transacetylase with cofactor lipoamide) oxidizes hydroxyethyl- to acetyl- and then transfers acetyl- to CoA, forming acetyl-CoA.
4. **Step D**: Acetyl CoA was made in the previous step. However, the process is incomplete. The E_2 is still attached to the acetyl CoA molecule. So, E_3 (dihydrolipoyl dehydrogenase) oxidizes the thiol groups of the dihydrolipoamide back to lipoamide.
5. **Step E**: As a side reaction, NAD^+ becomes reduced to NADH.

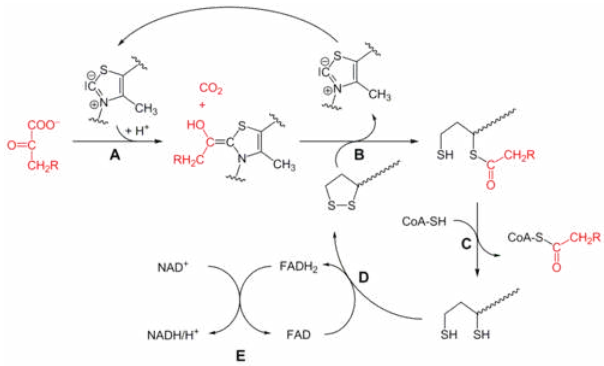


Figure 10.1.1: *The process of the PDH.* (https://en.wikipedia.org/wiki/Pyruvate_dehydrogenase_complex#/media/File:PDH_schema.png)

Supplemental Resources

Citric Acid Cycle (aka Krebs Cycle): http://virtuallabs.stanford.edu/other/biochem/TCA.swf

Krebs Cycle (aka Citric Acid Cycle): http://johnkyrk.com/krebs.html

Contributors

CK-12 Foundation by Sharon Bewick, Richard Parsons, Therese Forsythe, Shonna Robinson, and Jean Dupon.

Allison Soult, Ph.D. (Department of Chemistry, University of Kentucky)

Zubay, Geoffrey. Biochemistry. New York: Macmillan Publishing Company, 1988.

Boyer, Rodney. Concepts in Biochemistry. New Jersey: John Wiley & Sons, Inc, 2006

10: Pyruvate Dehydrogenase Links Glycolysis to Krebs Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.1: The Krebs Cycle (Citric Acid Cycle)

The primary catabolic pathway in the body is the **citric acid cycle (CAC)** because it is here that oxidation to CO_2 occurs for breakdown products of the cell's major building blocks - sugars, fatty acids, amino acids. The pathway is cyclic (***Figure 10.1***) and thus, doesn't really have a starting or ending point. All of the reactions occur in the mitochondrion, though one enzyme is embedded in the organelle's membrane. As needs change, cells may use a subset of the reactions of the cycle to produce a desired molecule rather than to run the entire cycle.

The primary catabolic pathway in the body is the citric acid cycle, also known as the tricarboxylic acid cycle and the Krebs cycle, completes the oxidation of glucose by taking the pyruvates from glycolysis (and other pathways), and completely breaking them down into CO_2 molecules, H_2O molecules, and generating additional ATP by oxidative phosphorylation. In prokaryotic cells, the citric acid cycle occurs in the cytoplasm; in eukaryotic cells the citric acid cycle takes place in the matrix of the mitochondria.

The overall reaction for the citric acid cycle is:

$$\textit{2 acetyl groups} + 6\ NAD^+ + 2\ FAD + 2\ ADP + 2\ Pi \rightarrow 4\ CO_2 + 6\ NADH + 6H^+ + 2\ FADH_2 + 2\ ATP$$

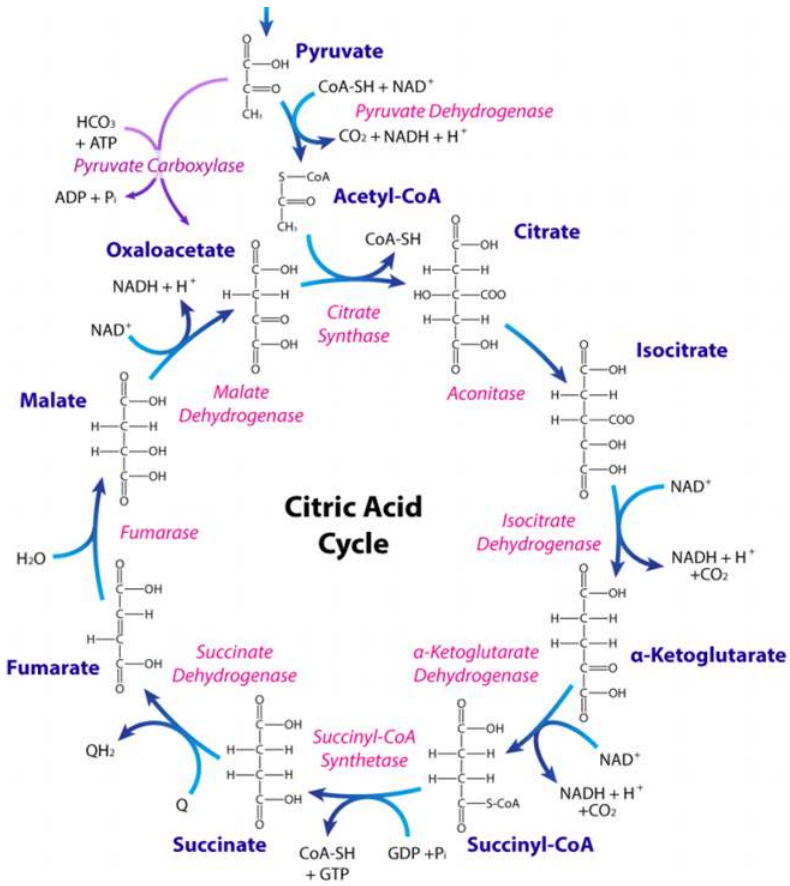


Figure 10.1: *The citric acid cycle*

Steps in the Citric Acid Cycle

Step 1. The first step is a condensation step, combining the two-carbon acetyl group (from acetyl CoA) with a four-carbon oxaloacetate molecule to form a six-carbon molecule of citrate. CoA is bound to a sulfhydryl group (-SH) and diffuses away to eventually combine with another acetyl group. This step is irreversible because it is highly exergonic. The rate of this reaction is controlled by negative feedback and the amount of ATP available. If ATP levels increase, the rate of this reaction decreases. If ATP is in short supply, the rate increases.

Step 2. Citrate loses one water molecule and gains another as citrate is converted into its isomer, isocitrate.

Steps 3 and 4. In step three, isocitrate is oxidized, producing a five-carbon molecule, α-ketoglutarate, together with a molecule of CO_2 and two electrons, which reduce NAD^+ to NADH. This step is also regulated by negative feedback from ATP and NADH and by a positive effect of ADP. Steps three and four are both oxidation and decarboxylation steps, which release electrons that reduce NAD^+ to NADH and release carboxyl groups that form CO_2 molecules. α-Ketoglutarate is the product of step three, and a succinyl group is the product of step four. CoA binds the succinyl group to form succinyl CoA. The enzyme that catalyzes step four is regulated by feedback inhibition of ATP, succinyl CoA, and NADH.

Step 5. A phosphate group is substituted for coenzyme A, and a high- energy bond is formed. This energy is used in substrate-level phosphorylation (during the conversion of the succinyl group to succinate) to form either guanine triphosphate (GTP) or ATP. There are two forms of the enzyme, called isoenzymes, for this step, depending upon the type of animal tissue in which they are found. One form is found in tissues that use large amounts of ATP, such as heart and skeletal muscle. This form produces ATP. The second form of the enzyme is found in tissues that have a high number of anabolic pathways, such as liver. This form produces GTP. GTP is energetically equivalent to ATP; however, its use is more restricted. In particular, protein synthesis primarily uses GTP.

Step 6. Step six is a dehydration process that converts succinate into fumarate with the help of an enzyme called **succinate dehydrogenate**. Two hydrogen atoms are transferred to FAD, producing $FADH_2$. The energy contained in the electrons of these atoms is insufficient to reduce NAD^+ but adequate to reduce FAD. Unlike NADH, this carrier remains attached to the enzyme and transfers the electrons to the electron transport chain directly. This process is made possible by the localization of the enzyme catalyzing this step inside the inner membrane of the mitochondrion.

Step 7. Water is added to fumarate during step seven, and malate is produced. The last step in the citric acid cycle regenerates oxaloacetate by oxidizing malate. Another molecule of NADH is produced.

Products of the Citric Acid Cycle

Two carbon atoms come into the citric acid cycle from each acetyl group, representing four out of the six carbons of one glucose molecule. Two carbon dioxide molecules are released on each turn of the cycle; however, these do not necessarily contain the most recently-added carbon atoms. The two acetyl carbon atoms will eventually be released on later turns of the cycle; thus, all six carbon atoms from the original glucose molecule are eventually incorporated into carbon dioxide. Each turn of the cycle forms three NADH molecules and one $FADH_2$ molecule. These carriers will connect with the last portion of aerobic respiration to produce ATP molecules. One GTP or ATP is also made in each cycle. Several of the intermediate compounds in the citric acid cycle can be used in synthesizing non-essential amino acids; therefore, the cycle is amphibolic (both catabolic and anabolic).

Summary

In the citric acid cycle, the acetyl group from acetyl CoA is attached to a four-carbon oxaloacetate molecule to form a six-carbon citrate molecule. Through a series of steps, citrate is oxidized, releasing two carbon dioxide molecules for each acetyl group fed into the cycle. In the process, three NAD^+ molecules are reduced to NADH, one FAD molecule is reduced to $FADH_2$, and one ATP or GTP (depending on the cell type) is produced (by substrate-level phosphorylation). Because the final product of the citric acid cycle is also the first reactant, the cycle runs continuously in the presence of sufficient reactants.

10.1: The Krebs Cycle (Citric Acid Cycle) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

10.2: Exercises

1. What are the steps involved (in order) in the conversion of pyruvate to acetyl-CoA?
2. How many net ATP is yielded during converting one glucose molecule into CO_2 using Glycolysis, PDH and CAC?
3. What is the name of the enzyme that transports the pyruvate into the PDH?
4. Provide the OVERALL reaction (net equation) of Citric Acid Cycle.
5. Write the OVERALL reaction (net equation) catalyzed by pyruvate dehydrogenase complex.
6. What is the step-wise chemical mechanism of PDH?
7. Pyruvate dehydrogenase complex and citric acid cycle together can oxidize pyruvate in CO_2. During this process (from pyruvate to CO2),

a) Which enzymes are used for regulation?

b) Which molecules are used for the activation during the regulation?

c) Which molecules are used for the inhibition during the regulation?

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

CHAPTER OVERVIEW

11: Electron Transport Chain and Oxidative Phosphorylation

Regulatory Mechanisms

A variety of mechanisms is used to control cellular respiration. Some type of control exists at each stage of glucose metabolism. Access of glucose to the cell can be regulated using the GLUT proteins that transport glucose (Figure 11.1). Different forms of the GLUT protein control passage of glucose into the cells of specific tissues.

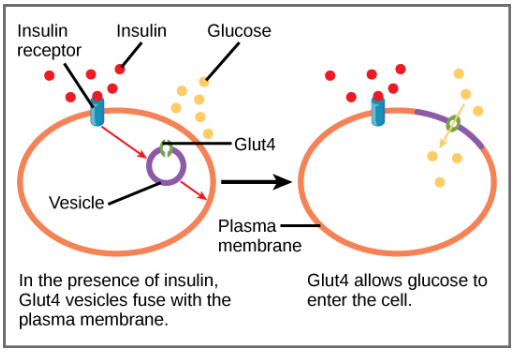


Figure 11.1: GLUT4 is a glucose transporter that is stored in vesicles. A cascade of events that occurs upon insulin binding to a receptor in the plasma membrane causes GLUT4-containing vesicles to fuse with the plasma membrane so that glucose may be transported into the cell.

Some reactions are controlled by having two different enzymes—one each for the two directions of a reversible reaction. Reactions that are catalyzed by only one enzyme can go to equilibrium, stalling the reaction. In contrast, if two different enzymes (each specific for a given direction) are necessary for a reversible reaction, the opportunity to control the rate of the reaction increases, and equilibrium is not reached.

A number of enzymes involved in each of the pathways—in particular, the enzyme catalyzing the first committed reaction of the pathway—are controlled by attachment of a molecule to an allosteric site on the protein. The molecules most commonly used in this capacity are the nucleotides ATP, ADP, AMP, NAD^+, and NADH. These regulators, allosteric effectors, may increase or decrease enzyme activity, depending on the prevailing conditions. The allosteric effector alters the steric structure of the enzyme, usually affecting the configuration of the active site. This alteration of the protein's (the enzyme's) structure either increases or decreases its affinity for its substrate, with the effect of increasing or decreasing the rate of the reaction. The attachment signals to the enzyme. This binding can increase or decrease the enzyme's activity, providing feedback. This feedback type of control is effective as long as the chemical affecting it is attached to the enzyme. Once the overall concentration of the chemical decreases, it will diffuse away from the protein, and the control is relaxed.

Control of Catabolic Pathways

Enzymes, proteins, electron carriers, and pumps that play roles in glycolysis, the citric acid cycle, and the electron transport chain tend to catalyze non-reversible reactions. Whether a particular enzyme activity is released depends upon the energy needs of the cell (as reflected by the levels of ATP, ADP, and AMP).

11.1: ETC and Oxidative Phosphorylation

11.1: Exercises

Contributors

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

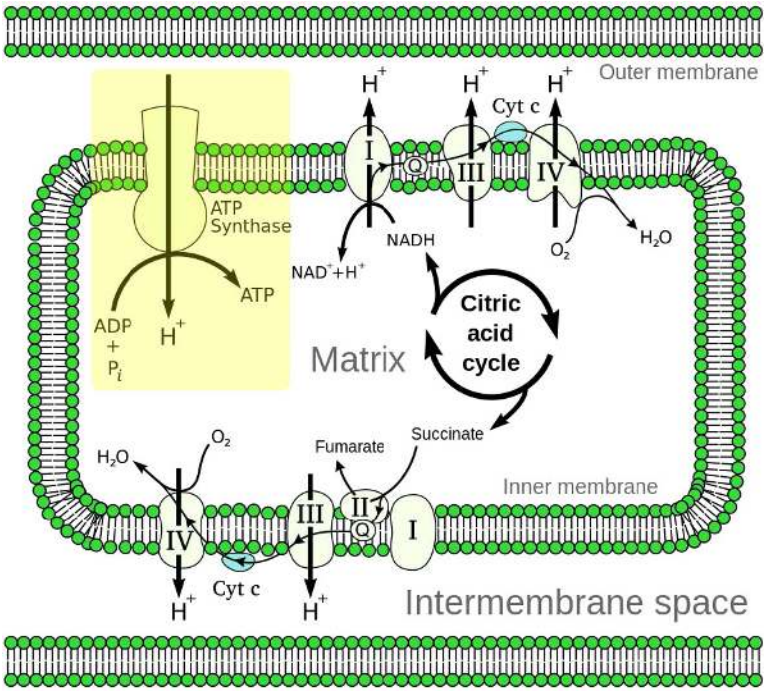
This page titled 11: Electron Transport Chain and Oxidative Phosphorylation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Henry Jakubowski.

11.1: ETC and Oxidative Phosphorylation

Source: BiochemFFA_5_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

In eukaryotic cells, the vast majority of ATP synthesis occurs in the mitochondria in a process called oxidative phosphorylation. Even plants, which generate ATP by photophosphorylation in chloroplasts, contain mitochondria for the synthesis of ATP through oxidative phosphorylation.

Oxidative phosphorylation is linked to a process known as electron transport (Figure 5.14). The electron transport system, located in the inner mitochondrial membrane, transfers electrons donated by the reduced electron carriers NADH and FADH2 (obtained from glycolysis, the citric acid cycle or fatty acid oxidation) through a series of electrons acceptors, to oxygen. As we shall see, movement of electrons through complexes of the electron transport system essentially “charges” a battery that is used to make ATP in oxidative phosphorylation. In this way, the oxidation of sugars and fatty acids is coupled to the synthesis of ATP, effectively extracting energy from food.



Chemiosmotic model

Dr. Peter Mitchell introduced a radical proposal in 1961 to explain the mechanism by which mitochondria make ATP. It is known as the chemiosmotic hypothesis and has been shown over the years to be correct. Mitchell proposed that synthesis of ATP in mitochondria depends on an electrochemical gradient, across the mitochondrial inner membrane, that arises ultimately from the energy of reduced electron carriers, NADH and FADH2.

Electron transport

Further, the proposal states that the gradient is created when NADH and FADH2 transfer their electrons to an electron transport system (ETS) located in the inner mitochondrial membrane. Movement of electrons through a series of of electron carriers is coupled to the pumping of protons out of the mitochondrial matrix across the inner mitochondrial membrane into the space between the inner and outer membranes. The result is creation of a gradient of protons whose potential energy can be used to make ATP. Electrons combine with oxygen and protons at the end of the ETS to make water.

ATP synthase

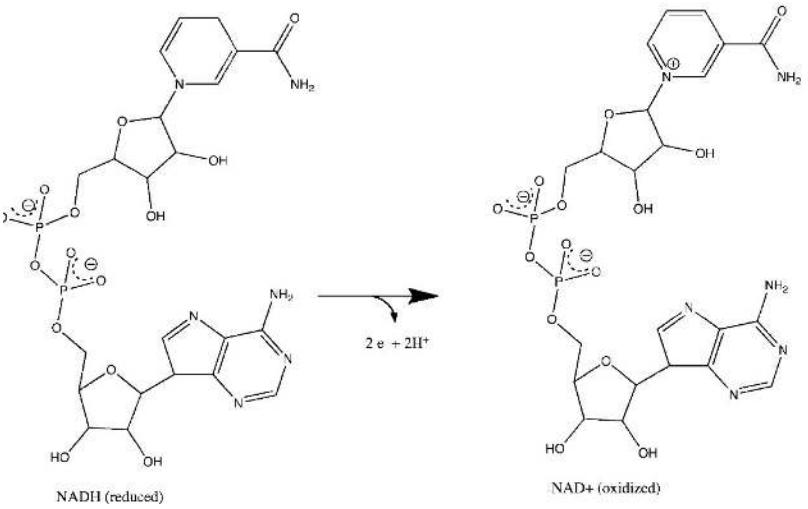
In oxidative phosphorylation, ATP synthesis is accomplished as a result of protons re-entering the mitochondrial matrix via the transmembrane ATP synthase complex, which combines ADP with inorganic phosphate to make ATP. Central to the proper functioning of mitochondria through this process is the presence of an intact mitochondrial inner membrane impermeable to protons.

Tight coupling

When this is the case, tight coupling is said to exist between electron transport and the synthesis of ATP (called oxidative phosphorylation). Chemicals which permeabilize the inner mitochondrial membrane to protons cause uncoupling, that is, they allow the protons to leak back into the mitochondrial matrix, rather than through the ATP synthase, so that the movement of electrons through the ETS is no longer linked to the synthesis of ATP.

Power plants

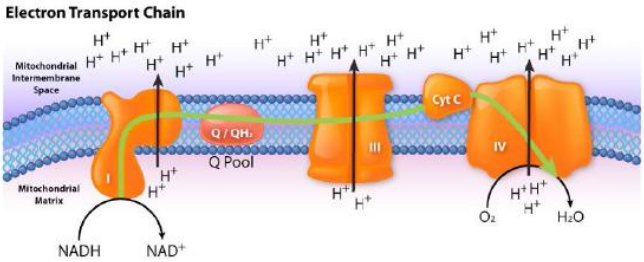
Mitochondria are called the power plants of the cell because most of a cell's ATP is produced there in the process of oxidative phosphorylation. The mechanism by which ATP is made in oxidative phosphorylation is one of the most interesting in all of biology.

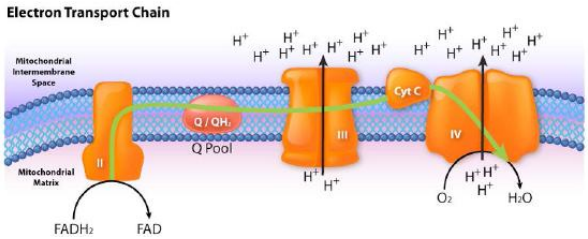


Considerations

The process has three primary considerations. The first is electrical – electrons from reduced electron carriers, such as NADH and FADH2, enter the electron transport system via Complex I and II, respectively. As seen in Figure 5.16 and Figure 5.17, electrons move from one complex to the next, not unlike the way they move through an electrical circuit. Such movement occurs a a result of

a set of reduction-oxidation (redox) reactions with electrons moving from a more negative reduction potential to a more positive one.





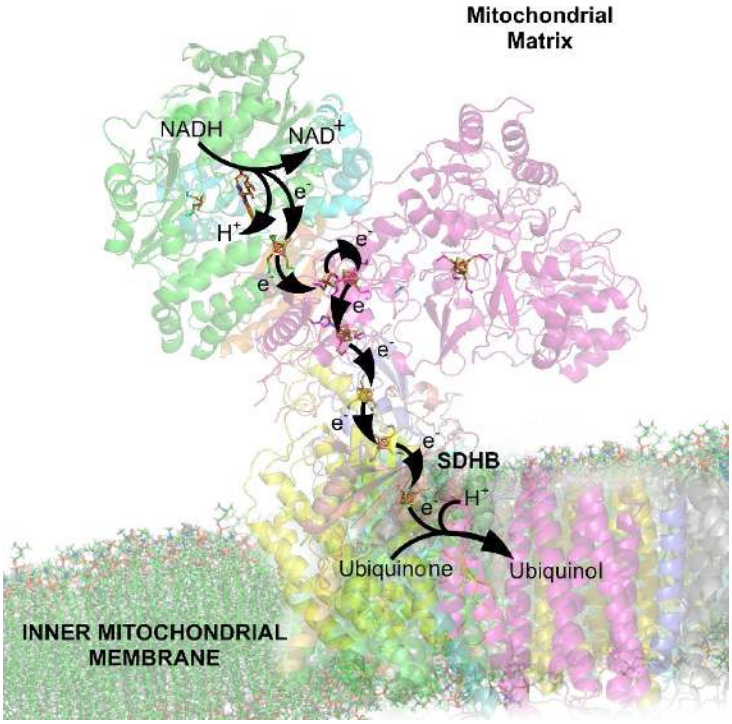
One can think of this occurring as a process where carriers “take” electrons away from complexes with lower reduction potential, much the way a bully takes lunch money from a smaller child. In this scheme, the biggest “bully” is oxygen in Complex IV. Electrons gained by a carrier cause it to be reduced, whereas the carrier giving up the electrons is oxidized.

Entry of electrons to system

Movement of electrons through the chain begins either by 1) transfer from NADH to Complex I (Figure 5.16) or 2) movement of electrons through a covalently bound FADH2 (Figure 5.17) in the membrane-bound succinate dehydrogenase (Complex II). (An alternate entry point for electrons from FADH2 is the Electron Transferring Flavoprotein via the electron-transferring-flavoprotein dehydrogenase, not shown).

Traffic cop

Both Complex I and II pass electrons to the inner membrane’s coenzyme Q (CoQ - Figures 5.18 & 5.19). In each case, coenzyme Q accepts electrons in pairs and passes them off to Complex III (CoQH2-cytochrome c reductase) singly. Coenzyme Q thus acts as a traffic cop, regulating the flow of electrons through the ETS.



Docking station

Complex III is a docking station or interchange for the incoming electron carrier (coenzyme Q) and the outgoing carrier (cytochrome c). Movement of electrons from Coenzyme Q to Complex III and then to cytochrome C occurs as a result of what is

referred to as the Q-cycle (see below).

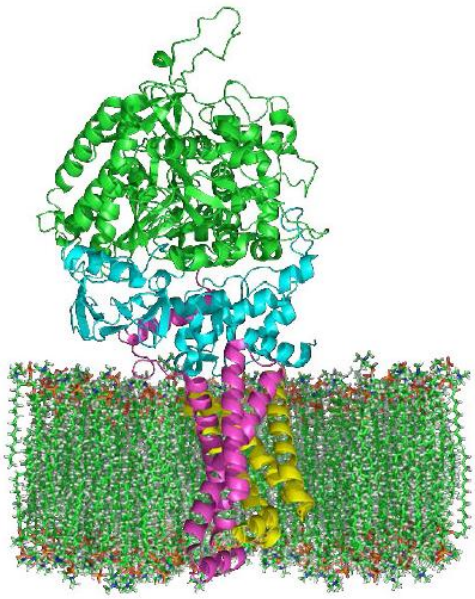
Complex III acts to ferry electrons from CoQ to cytochrome c. Cytochrome c takes one electron from Complex III and passes it to Complex IV (cytochrome oxidase). Complex IV is the final protein recipient of the electrons. It passes them to molecular oxygen (O2) to make two molecules of water. Making two water molecules requires four electrons, so Complex IV must accept, handle, and pass to molecular oxygen four separate electrons, causing the oxidation state of oxygen to be sequentially changed with addition of each electron.

Proton pumping

As electrons pass through complexes I, III, and IV, there is a release of a small amount of energy at each step, which is used to pump protons from the mitochondrial matrix (inside of mitochondrion) and deposit them in the intermembrane space (between the inner and outer membranes of the mitochondrion). The effect of this redistribution is to increase the electrical and chemical potential across the membrane.

Potential energy

As discussed earlier, electrochemical gradients have potential energy. Students may think of the process as “charging the battery.” Just like a charged battery, the potential arising from the proton differential across the membrane can be used to do things. In the mitochondrion, what the proton gradient does is facilitate the production of ATP from ADP and Pi. This process is known as oxidative phosphorylation, because the phosphorylation of ADP to ATP is dependent on the oxidative reactions occurring in the mitochondria.



Having understood the overall picture of the synthesis of ATP linked to the movement of electrons through the ETS, we will take a closer look at the individual components of the ETS.

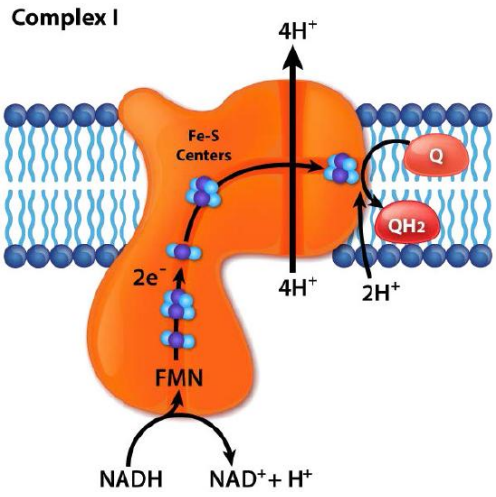
Complex I

Complex I (also called NADH:ubiquinone oxidoreductase or NADH dehydrogenase (ubiquinone)) is the electron acceptor from NADH in the electron transport chain and the largest complex found in it.

Complex I contains 44 individual polypeptide chains, numerous iron-sulfur centers, a molecule of flavin mononucleotide (FMN) and has an L shape with about 60 transmembrane domains. In the process of electron transport through it, four protons are pumped across the inner membrane into the intermembrane space and electrons move from NADH to coenzyme Q, converting it from

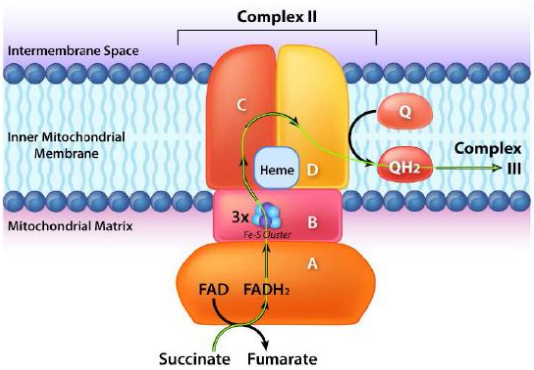
ubiquinone (no electrons) to ubiquinol (gain of two electrons). An intermediate form, ubisemiquinone (gain of one electron), is found in the Q-cycle.

Electrons travel through the complex via seven primary iron sulfur centers. The best known inhibitor of the complex, rotenone, works by binding to the CoQ binding site. Other inhibitors include ADP-ribose (binds to the NADH site) and piericidin A (rotenone analog). The process of electron transfer through complex I is reversible and when this occurs, superoxide (a reactive oxygen species) may be readily generated.



Complex II

Complex II (also called succinate dehydrogenase or succinate-coenzyme Q reductase) is a membrane bound enzyme of the citric acid cycle that plays a role in the electron transport process, transferring electrons from its covalently bound FADH2 to coenzyme Q. The process occurs, as shown in Figure 5.20 and Figure 5.21, with transfer of electrons from succinate to FAD to form FADH2 and fumarate. FADH2, in turn, donates electrons to a relay system of iron-sulfur groups and they ultimately reduce ubiquinone (CoQ) along with two protons from the matrix to ubiquinol. The role of the heme group in the process is not clear. Inhibitors of the process include carboxin, malonate, malate, and oxaloacetate. The role of citric acid cycle intermediates as inhibitors is thought to be due to inhibition of the reversal of the transfer process which can produce superoxide.

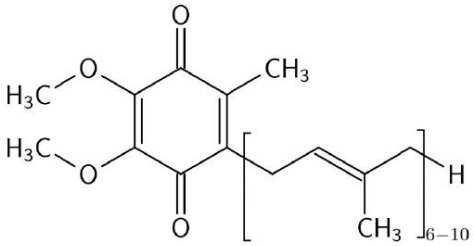


Coenzyme Q

Coenzyme Q (Figure 5.23) is a 1,4 benzoquinone whose name is often given as Coenzyme Q10, CoQ, or Q10. The 10 in the name refers to the number of isoprenyl units it contains that anchor it to the mitochondrial inner membrane. CoQ is a vitamin-like lipid substance found in most eukaryotic cells as a component of the electron transport system. The requirement for CoQ increases with increasing energy needs of cells, so the highest concentrations of CoQ in the body are found in tissues that are the most metabolically active - heart, liver, and kidney.

Three forms

CoQ is useful because of its ability to carry and donate electrons and particularly because it can exist in forms with two extra electrons (fully reduced - ubiquinol), one extra electron (semi-reduced - ubisemiquinone), or no extra electrons (fully oxidized - ubiquinone). This ability allows CoQ to provide transition between the first part of the electron transport system that moves electrons in pairs and the last part of the system that moves electrons one at a time.



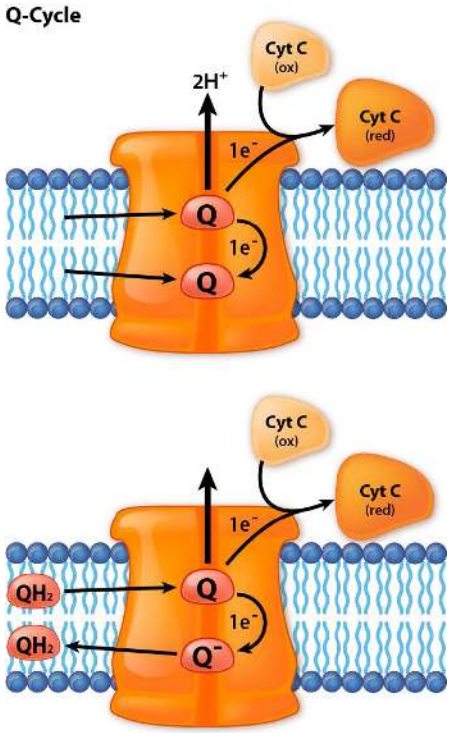
Complex III

Complex III (also known as coenzyme Q : cytochrome c — oxidoreductase or the cytochrome bc1 complex - Figure 5.24) is the third electron accepting complex of the electron transport system. It is a transmembrane protein with multiple subunits present in the mitochondria of all aerobic eukaryotic organisms and and the cell membrane of almost all bacteria. The complex contains 11 subunits, a 2-iron ferredoxin, cytochromes b and c1 and belongs to the family of oxidoreductase enzymes.

It accepts electrons from coenzyme Q in electron transport and passes them off to cytochrome c. In this cycle, known as the Q cycle, electrons arrive from CoQ in pairs, but get passed to cytochrome c individually. In the overall process, two protons are consumed from the matrix and four protons are pumped into the intermembrane space. Movement of electrons through the complex can be inhibited by antimycin A, myxothiazol, and stigmatellin. Complex III is also implicated in creation of superoxide (a reactive oxygen species) when electrons from it leak out of the chain of transfer. The phenomenon is more pronounced when antimycin A is present.

Q-cycle

In the Q-cycle, electrons are passed from ubiquinol (QH2) to cytochrome c using Complex III as an intermediary docking station for the transfer. Two pair of electrons enter from QH2 and one pair is returned to another CoQ to re-make QH2. The other pair is donated singly to two different cytochrome c molecules.



Step one

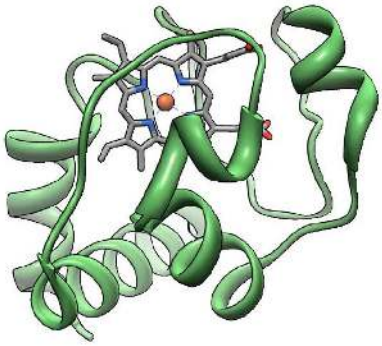
The Q-cycle happens in a two step process. First, a ubiquinol (CoQH2) and a ubiquinone (CoQ) dock at Complex III. Ubiquinol transfers two electrons to Complex III. One electron goes to a docked cytochrome c, reducing it and it exits (replaced by an oxidized cytochrome c). The other goes to the docked uniquinone to create the semi-reduced semiubiquinone (CoQ.-) and leaving behind a ubiquinone, which exits. This is the end of step 1.

Step two

The gap left behind by the ubiquinone (Q) that departed is replaced by another ubiquinol (QH2). It too donates two electrons to Complex III, which splits them. One goes to the newly docked oxidized cytochrome c, which is reduced and exits. The other goes to the ubisemiquinone. Two protons from the matrix combine with it to make another ubiquinol. It and the ubiquinone created by the electron donation exit Complex III and the process starts again. In the overall process, two protons are consumed from the matrix and four protons are pumped into the intermembrane space.

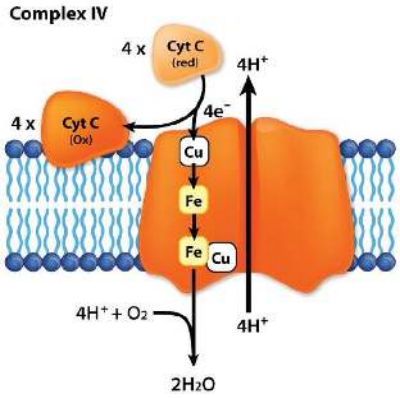
Cytochrome c

Cytochrome c (Figure 5.26) is a small (12,000 Daltons), highly conserved protein, from unicellular species to animals, that is loosely associated with the inner mitochondrial membrane where it functions in electron transport. It contains a heme group which is used to carry a single electron from Complex III to Complex IV. Cytochrome c also plays an important role in apoptosis in higher organisms. Damage to the mitochondrion that results in release of cytochrome c can stimulate assembly of the apoptosome and activation of the caspase cascade that leads to programmed cell death.



Complex IV

Complex IV, also known as cytochrome c oxidase is a 14 subunit integral membrane protein at the end of the electron transport chain (Figure 5.27). It is responsible for accepting one electron each from four cytochrome c proteins and adding them to molecular oxygen (O2) along with four protons from the mitochondrial matrix to make two molecules of water. Four protons from the matrix are also pumped into the intermembrane space in the process. The complex has two molecules of heme, two cytochromes (a and a3), and two copper centers (called CuA ad CuB). Cytochrome c docks near the CuA and donates an electron to it. The reduced CuA passes the electron to cytochrome a, which turns it over to the a3-CuB center where the oxygen is reduced. The four electrons are thought to pass through the complex rapidly resulting in complete reduction of the oxygen-oxygen molecule without formation of a peroxide intermediate or superoxide, in contrast to previous predictions.

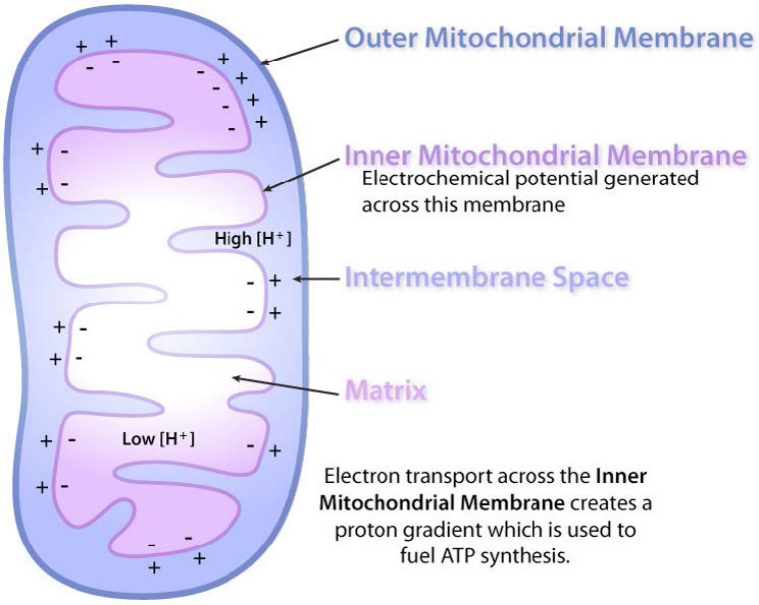


Respirasome

There has been speculation for many years that a supercomplex of electron carriers in the inner membrane of the mitochondrion may exist in cells with individual carriers making physical contact with each other. This would make for more efficient transfer reactions, minimize the production of reactive oxygen species and be similar to metabolons of metabolic pathway enzymes, for which there is some evidence. Now, evidence appears to be accumulating that complexes I, III, and IV form a supercomplex, which has been dubbed the respirasome1.

Oxidative phosphorylation

The process of oxidative phosphorylation uses the energy of the proton gradient established by the electron transport system as a means of phosphorylating ADP to make ATP. The establishment of the proton gradient is dependent upon electron transport. If electron transport stops or if the inner mitochondrial membrane's impermeability to protons is compromised, oxidative phosphorylation will not occur because without the proton gradient to drive the ATP synthase, there will be no synthesis of ATP.



ATP synthase

The protein complex harvesting energy from the proton gradient and using it to make ATP from ADP is an enzyme that has several names - Complex V, PTAS (Proton Translocating ATP Synthase), and ATP synthase (Figure 5.29). Central to its function is the movement of protons through it (from the intermembrane space back into the matrix). Protons will only provide energy to make ATP if their concentration is greater in the intermembrane space than in the matrix and if ADP is available.

It is possible, in some cases, for the concentration of protons to be greater inside the matrix than outside of it. When this happens, the ATP synthase can run backwards, with protons moving from inside to out, accompanied by conversion of ATP to ADP + Pi. This is usually not a desirable circumstance and there are some controls to reduce its occurrence.

Normally, ATP concentration will be higher inside of the mitochondrion and ADP concentration be higher outside the mitochondrion. However, when the rate of ATP synthesis exceeds the rate of ATP usage, then ATP concentrations rise outside the mitochondrion and ADP concentrations fall everywhere.

This may happen, for example, during periods of rest. It has the overall effect of reducing transport and thus lowering the concentration of ADP inside the matrix. Reducing ADP concentration in the matrix reduces oxidative phosphorylation and has effects on respiratory control (see HERE).

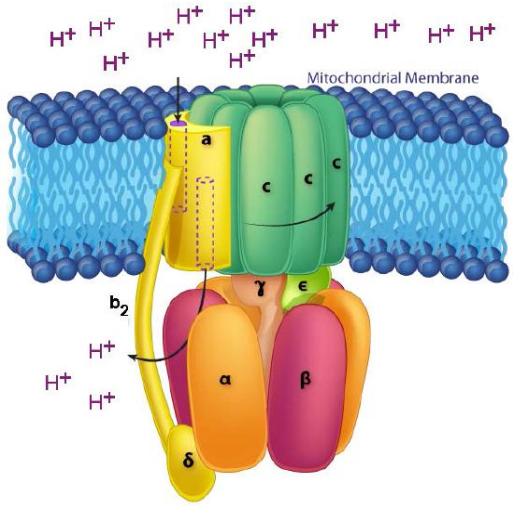
Another important consideration is that when ATP is made in oxidative phosphorylation, it is released into the mitochondrial matrix, but must be transported into the cytosol to meet the energy needs of the rest of the cell. This is accomplished by action of the adenine nucleotide translocase, an antiport that moves ATP out of the matrix in exchange for ADP moving into the matrix. This transport system is driven by the concentrations of ADP and ATP and ensures that levels of ADP are maintained within the mitochondrion, permitting continued ATP synthesis.

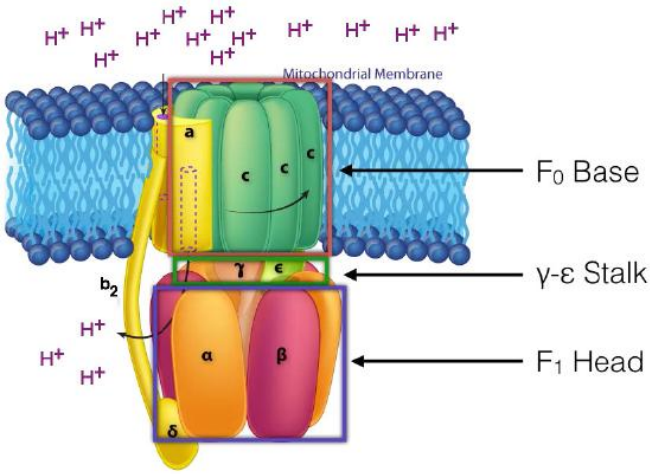
One last requirement for synthesis of ATP from ADP is that phosphate must also be imported into the matrix. This is accomplished by action of the phosphate translocase, which is a symport that moves phosphate into the mitochondrial matrix along with a proton.

There is evidence that the two translocases and ATP synthase may exist in a complex, which has been dubbed the ATP synthasome.

In summary, the electron transport system charges the battery for oxidative phosphorylation by pumping protons out of the mitochondrion. The intact inner membrane of the mitochondrion keeps the protons out, except for those that re-enter through ATP Synthase. The ATP Synthase allows protons to re-enter the mitochondrial matrix and harvests their energy to make ATP.

ATP synthase mechanism





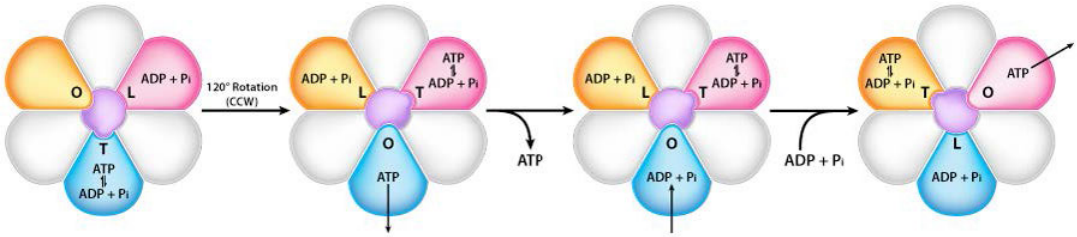
The ATP Synthase itself is an amazing nanomachine that makes ATP using a gradient of protons flowing through it from the intermembrane space back into the matrix. It is not easy to depict in a single image what the synthase does. Figure 5.31 and Figure 5.32 illustrate the multi-subunit nature of this membrane protein, which acts like a turbine at a hydroelectric dam. The movement of protons through the ATP Synthase c-ring causes it and the γ-ε stalk attached to it to turn. It is this action that is necessary for making ATP.

In ATP Synthase, the spinning components, or rotor, are the membrane portion (c ring) of the F0 base and the γ-ε stalk, which is connected to it. The γ-ε stalk projects into the F1 head of the mushroom structure. The F1 head contains the catalytic ability to make ATP. The F1 head is hexameric in structure with paired α and β proteins arranged in a trimer of dimers. ATP synthesis occurs within the β subunits.

Rotation of γ unit

Turning of the γ shaft (caused by proton flow) inside the α-β trimer of the F1 head causes each set of β proteins to change structure slightly into three different forms called Loose, Tight, and Open (L,T,O - Figure 5.31). Each of these forms has a function.

The Loose form binds ADP + Pi. The Tight form “squeezes” them together to form the ATP. The Open form releases the ATP into the mitochondrial matrix. Thus, as a result of the proton flow through the ATP synthase, from the intermembrane space into the matrix, ATP is made from ADP and Pi.



Respiratory control

When a mitochondrion has an intact inner membrane and protons can only return to the matrix by passing through the ATP synthase, the processes of electron transport and oxidative phosphorylation are said to be tightly coupled.

Interdependence

In simple terms, tight coupling means that the processes of electron transport and oxidative phosphorylation are interdependent. Without electron transport going on in the cell, oxidative phosphorylation will soon stop.

The reverse is also true, because if oxidative phosphorylation stops, the proton gradient will not be dissipated as it is being built by the electron transport system and will grow larger and larger. The greater the gradient, the greater the energy needed to pump protons out of the mitochondrion. Eventually, if nothing relieves the gradient, it becomes too large and the energy of electron transport is insufficient to perform the pumping. When pumping stops, so too does electron transport.

ADP dependence

Another relevant point is that ATP synthase is totally dependent upon a supply of ADP. In the absence of ADP, the ATP synthase stops functioning and when it stops, so too does movement of protons back into the mitochondrion. With this information, it is possible to understand the link between energy usage and metabolism. The root of this, as noted, is respiratory control.

At rest

To illustrate these links, let us first consider a person, initially at rest, who then suddenly jumps up and runs away. At first, the person's ATP levels are high and ADP levels are low (no exercise to burn ATP), so little oxidative phosphorylation is occurring and thus the proton gradient is high. Electron transport is moving slowly, if at all, so it is not using oxygen and the person's breathing is slow, as a result.

Exercise

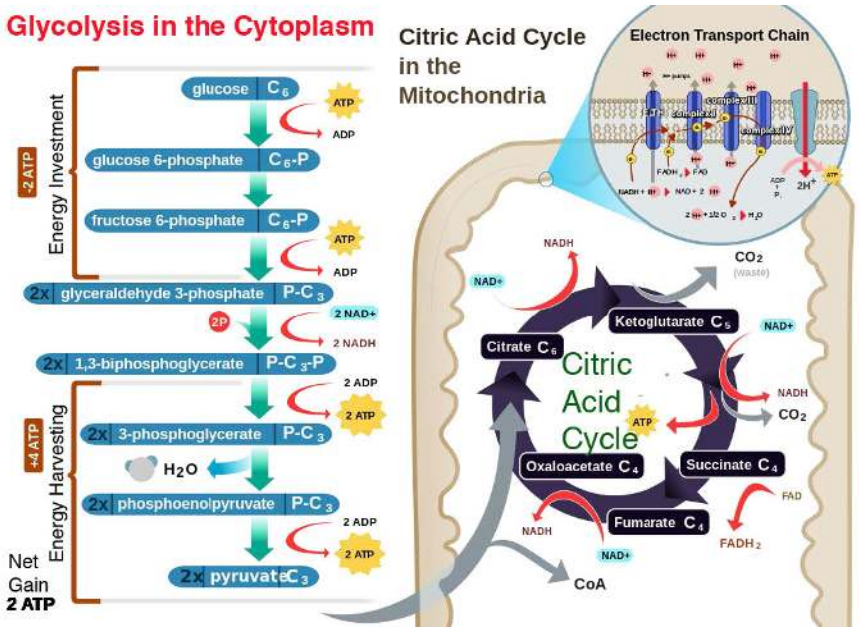
When running starts, muscular contraction, which uses energy, causes ATP to be converted to ADP. Increasing ADP in muscle cells favors oxidative phosphorylation to attempt to make up for the ATP being burned. ATP synthase begins working and protons begin to come back into the mitochondrial matrix. The proton gradient decreases, so electron transport re-starts.

Electron transport needs an electron acceptor, so oxygen use increases and when oxygen use increases, the person starts breathing more heavily to supply it. When the person stops running, ATP concentrations get rebuilt by ATP synthase. Eventually, when ATP levels are completely restored, ADP levels fall and ATP synthase stops or slows considerably. With little or no proton movement, electron transport stops because the proton gradient is too large. When electron transport stops, oxygen use decreases and the rate of breathing slows down.

Electron transport critical

The really interesting links to metabolism occur relative to whether or not electron transport is occurring. From the examples, we can see that electron transport will be relatively slowed when not exercising and more rapid when exercise (or other ATP usage) is occurring. Remember that electron transport is the way in which reduced electron carriers, NADH and FADH2, donate their electrons to the ETS , becoming oxidized to NAD+ and FAD, respectively.

Oxidized carriers, such as NAD+ and FAD are needed by catabolic pathways, like glycolysis, the citric acid cycle, and fatty acid oxidation. Anabolic pathways, such as fatty acid/fat synthesis and gluconeogenesis rely on reduced electron carriers, such as FADH2, NADH, and the related carrier, NADPH.



Links to exercise

High levels of NADH and FADH2 prevent catabolic pathways from operating, since NAD+ and FAD levels will be low and these are needed to accept the electrons released during catabolism by the oxidative processes.

Thanks to respiratory control, when one is exercising, NAD+ and FAD levels increase (because electron transport is running), so catabolic pathways that need NAD+ and FAD can function. The electrons lost in the oxidation reactions of catabolism are captured by NAD+ and FAD to yield NADH and FADH2, which then supply electrons to the electron transport system and oxidative phosphorylation to make more needed ATP.

Thus, during exercise, cells move to a mode of quickly cycling between reduced electron carriers (NADH/FADH2) and oxidized electron carriers (NAD+/FAD). This allows rapidly metabolizing tissues to transfer electrons to NAD+/FAD and it allows the reduced electron carriers to rapidly become oxidized, allowing the cell to produce ATP.

Rest

When exercise stops, NADH and FADH2 levels rise (because electron transport is slowing) causing catabolic pathways to slow/stop. If one does not have the proper amount of exercise, reduced carriers remain high in concentration for long periods of time. This means we have an excess of energy and then anabolic pathways, particularly fatty acid synthesis, are favored, so we get fatter.

Altering respiratory control

One might suspect that altering respiratory control could have some very dire consequences and that would be correct. Alterations can take the form of either inhibiting electron transport/oxidative phosphorylation or uncoupling the two . These alterations can be achieved using compounds with specific effects on particular components of the system.

All of the chemicals described here are laboratory tools and should never be used by people. The first group for discussion are the inhibitors. In tightly coupled mitochondria, inhibiting either electron transport or oxidative phosphorylation has the effect of inhibiting the other one as well.

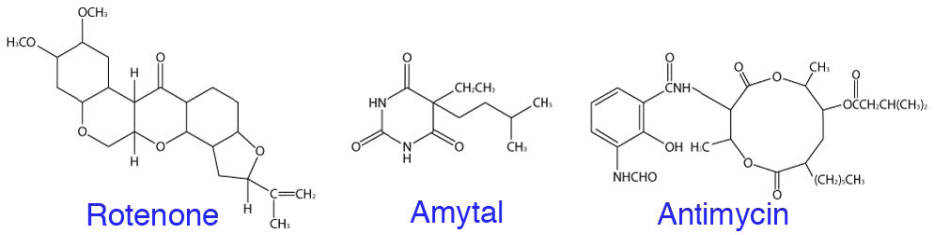
Electron transport inhibitors

Common inhibitors of electron transport include **rotenone** and **amytal**, which stop movement of electrons past Complex I, malonate, **malate,** and **oxaloacetat**e, which inhibit movement of electrons through Complex II, **antimycin A** which stops movement of electrons past Complex III, and **cyanide, carbon monoxide, azide**, and hydrogen sulfide, which inhibit electron movement through Complex IV (Figure 5.33). All of these compounds can stop electron transport directly (no movement of

electrons) and oxidative phosphorylation indirectly (proton gradient will dissipate). While some of these compounds are not commonly known, almost everyone is aware of the hazards of carbon monoxide and cyanide, both of which can be lethal.

ATP synthase inhibitor

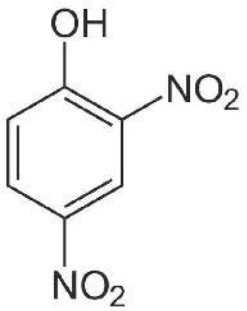
It is also possible to use an inhibitor of ATP synthase to stop oxidative phosphorylation directly (no ATP production) and electron transport indirectly (proton gradient not relieved so it becomes increasingly difficult to pump protons out of matrix). **Oligomycin A** (Figure 5.34) is an inhibitor of ATP synthase.



Rotenone

Rotenone, which is a plant product, is used as a natural insecticide that is permitted for organic farming. When mitochondria are treated with this, electron transport will stop at Complex I and so, too, will the pumping of protons out of the matrix. When this occurs, the proton gradient rapidly dissipates, stopping oxidative phosphorylation as a consequence. There are other entry points for electrons than Complex I, so this type of inhibition is not as serious as using inhibitors of Complex IV, since no alternative route for electrons is available. It is for this reason that cyanide, for example, is so poisonous.

2,4-DNP



Respiratory control can be completely destroyed by using a reagent that permeabilizes the inner mitochondrial membrane to protons. There are several such reagents, but the best known one is 2,4 dinitrophenol (2,4 DNP - Figure 5.35). Treatment of mitochondria with 2,4 DNP makes the mitochondrial inner membrane “leaky” to protons. This has the effect of providing an alternate route for protons to reenter the matrix besides going through ATP synthase, and uncouples oxidative phosphorylation from electron transport.

Imagine a dam holding back water with a turbine generating electricity through which water must flow. When all water flows through the turbine, the maximum amount of electricity can be generated. If one pokes a hole in the dam, though, water will flow through the hole and less electricity will be created. The generation of electricity will thus be uncoupled from the flow of water. If the hole is big enough, the water will all drain out through the hole and no electricity will be made.

Bypassing ATP synthase

Imagine, now, that the proton gradient is the equivalent of the water, the inner membrane is the equivalent of the dam and the ATP synthase is the turbine. When protons have an alternate route, little or no ATP will be made because protons will pass through the membrane’s holes instead of spinning the turbine of ATP synthase.

It is important to recognize, though, that uncoupling by 2,4 DNP works differently from the electron transport inhibitors or the ATP synthase inhibitor. In those situations, stopping oxidative phosphorylation resulted in indirectly stopping electron transport, since

the two processes were coupled and the inhibitors did not uncouple them. Similarly, stopping electron transport indirectly stopped oxidative phosphorylation for the same reason.

Such is not the case with 2,4 DNP. Stopping oxidative phosphorylation by destroying the proton gradient allows electron transport to continue unabated (it actually stimulates it), since the proton gradient cannot build no matter how much electron transport runs. Consequently, electron transport runs like crazy but oxidative phosphorylation stops. When that happens, NAD+ and FAD levels rise, and catabolic pathways run unabated with abundant supplies of these electron acceptors. The reason such a scenario is dangerous is because the body is using all of its nutrient resources, but no ATP is being made. Lack of ATP leads to cellular (and organismal) death. In addition, the large amounts of heat generated can raise the temperature of the body to unsafe levels.

Thermogenin

One of the byproducts of uncoupling electron transport is the production of heat. The faster metabolic pathways run, the more heat is generated as a byproduct. Since 2,4 DNP causes metabolism to speed up, a considerable amount of heat can be produced. Controlled uncoupling is actually used by the body in special tissues called brown fat. In this case, brown fat cells use the heat created to help thermoregulate the temperature of newborn children.

Permeabilization of the inner membrane is accomplished in brown fat by the synthesis of a protein called thermogenin (also known as uncoupling protein). Thermogenin binds to the inner membrane and allows protons to pass through it, thus bypassing the ATP synthase. As noted for 2,4 DNP, this results in activation of catabolic pathways and the more catabolism occurs, the more heat is generated.

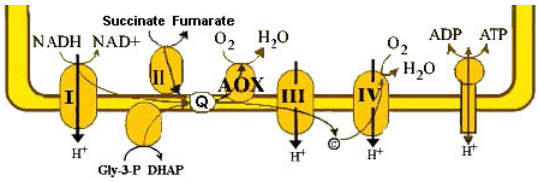
Dangerous drug

In uncoupling, whether through the action of an endogenous **uncoupling protein** or **DNP**, the energy that would have normally been captured in ATP is lost as heat. In the case of uncoupling by thermogenin, this serves the important purpose of keeping newborn infants warm. But in adults, uncoupling merely wastes the energy that would have been harvested as ATP. In other words, it mimics starvation, even though there is plenty of food, because the energy is dissipated as heat.

This fact, and the associated increase in metabolic rate, led to DNP being used as a weight loss drug in the 1930s. Touted as an effortless way to lose weight without having to eat less or exercise more, it was hailed as a magic weight loss pill. It quickly became apparent, however, that this was very dangerous. Many people died from using this drug before laws were passed to ban the use of DNP as a weight loss aid.

Alternative oxidase

Another approach to generating heat that doesn't involve breaking respiratory control is taken by some fungi, plants, and protozoa. They use an alternative electron transport. In these organisms, there is an enzyme called alternative oxidase (Figure 5.36). Alternative oxidase is able to accept electrons from CoQ and pass them directly to oxygen.



The process occurs in coupled mitochondria. Its mechanism of action is to reduce the yield of ATP, since fewer protons are being pumped per reduced electron carrier. Thus NAD+ concentrations increase, oxygen consumption increases, and the efficiency of ATP production decreases.

Organisms using this method must activate catabolic pathways by the increase in NAD+ concentration. This, in turn produces quantities of NADH and FADH2 necessary to make sufficient amounts of ATP. The byproduct of this increased catabolism is more heat. Not surprisingly, the alternative oxidase pathway can be activated by cold temperatures.

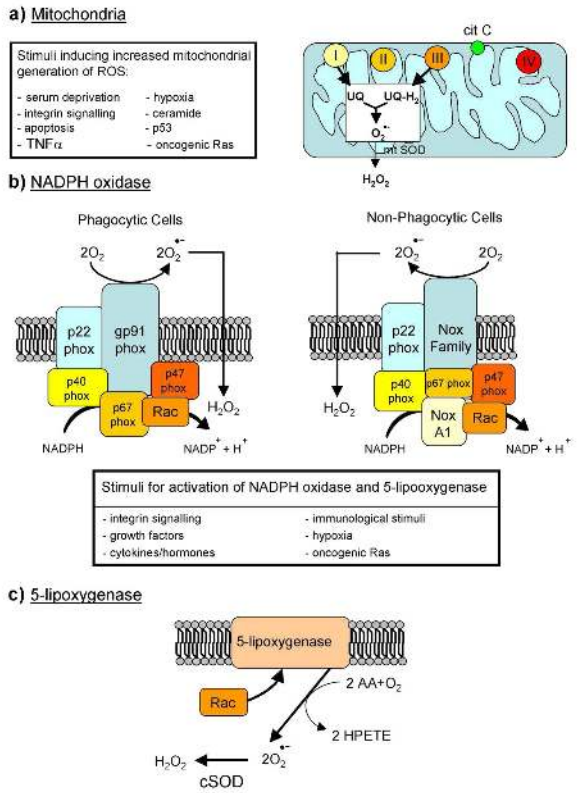
Energy efficiency

Cells are not 100% efficient in energy use. Nothing we know is. Consequently, cells do not get as much energy out of catabolic processes as they put into anabolic processes. A good example is the synthesis and breakdown of glucose, something liver cells are frequently doing. The complete conversion of glucose to pyruvate in glycolysis (catabolism) yields two pyruvates plus 2 NADH plus 2 ATPs. Conversely, the complete conversion of two pyruvates into glucose by gluconeogenesis (anabolism) requires 4 ATPs, 2 NADH, and 2 GTPs. Since the energy of GTP is essentially equal to that of ATP, gluconeogenesis requires a net of 4 ATPs more than glycolysis yields. This difference must be made up in order for the organism to meet its energy needs. It is for this reason that we eat. In addition, the inefficiency of our capture of energy in reactions results in the production of heat and helps to keep us warm, as noted. You can read more about glycolysis (HERE) and gluconeogenesis (HERE).

Metabolic controls of energy

It is also noteworthy that cells do not usually have both catabolic and anabolic processes for the same molecules occurring simultaneously inside of them (for example, breakdown of glucose and synthesis of glucose) because the cell would see no net production of anything but heat and a loss of ATPs with each turn of the cycle. Such cycles are called futile cycles and cells have controls in place to limit the extent to which they occur. Since futile cycles can, in fact, yield heat, they are used as sources of heat in some types of tissue. Brown adipose tissue of mammals uses this strategy, as described earlier. See also HERE for more on heat generation with a futile cycle.

Reactive oxygen species



Reactive oxygen species (ROS - Figure 5.37) are oxygen containing molecules, such as peroxides, hydroxyl radical, superoxide, peroxynitrite, and others that are very chemically reactive. Though some ROS, such as peroxide and nitric oxide have important

biological functions in signaling, increases in reactive oxygen species in times of stress can cause significant damage in the cell. Exogenous sources of ROS, such as pollution, tobacco, smoke, radiation or drugs can also cause significant problems.

Endogenous production of ROS is directed towards intracellular signaling (H2O2 and nitric oxide, for example) and defense. Many cells, for example, have NADPH oxidase (Figure 5.38) embedded in the exterior portion of the plasma membranes, in peroxisomes, and endoplasmic reticulum. It produces superoxides in the reaction below to kill bacteria .

In the immune system, cells called phagocytes engulf foreign cells and then use ROS to kill them. ROS can serve as signals for action. In zebrafish, damaged tissues have increased levels of H2O2 and this is thought to be a signal for white blood cells to converge on the site. In fish lacking the genes to produce hydrogen peroxide, white blood cells do not converge at the damage site. Sources of hydrogen peroxide include peroxisomes, which generate it as a byproduct of oxidation of long chain fatty acids.

Aging

Reactive oxygen species are at the heart of the free radical theory of aging, which states that organisms age due to the accumulation of damage from free radicals in their cells. In yeast and Drosophila, there is evidence that reducing oxidative damage can increase lifespan. In mice, increasing oxidative damage decreases life span, though in Caenorhabditis, blocking production of superoxide dismutase actually increases lifespan, so the role of ROS in aging is not completely clear.

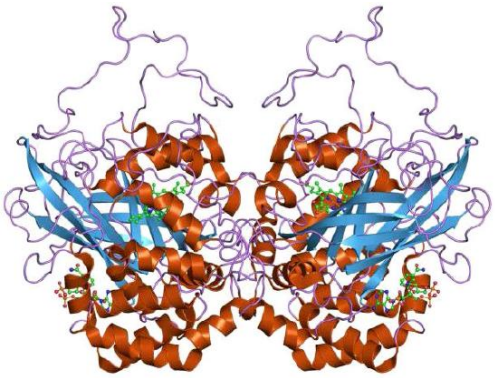
It is clear, though, that accumulation of mitochondrial damage is problematic for individual cells. Bcl-2 proteins on the surface of mitochondria monitor damage and if they detect it, will activate proteins called Bax to stimulate the release of cytochrome c from the mitochondrial membrane, stimulating apoptosis (programmed cell death). Eventually the dead cell will be phagocytosed.

A common endogenous source of superoxide is the electron transport chain. Superoxide can be produced when movement of electrons into and out of the chain don't match well. Under these circumstances, semi-reduced CoQ can donate an electron to O2 to form superoxide (O2-). Superoxide can react with many molecules, including DNA where it can cause damage leading to mutation. If it reacts with the aconitase enzyme, ferrous iron (Fe++) can be released which, in turn, can react in the Fenton reaction to produce another reactive oxygen species, the hydroxyl radical (Figure 5.39) .

Countering the effects of ROS are enzymes, such as catalase, superoxide dismutase, and anti-oxidants, such as glutathione and vitamins C and E.

Glutathione protects against oxidative damage by being a substrate for the enzyme glutathione peroxidase. Glutathione peroxidase catalyzes the conversion of hydrogen peroxide to water (next page).

Catalase



Catalase (Figure 5.40) is an important enzyme for cells of all types that live in an oxygen environment. A first line of defense against reactive oxygen species, catalase catalyzes the breakdown of hydrogen peroxide into water and oxygen.

2 H2O2 <=> 2 H2O + O2

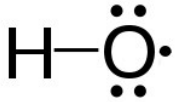
The enzyme, which employs four heme groups in its catalysis, works extremely rapidly, converting up to 40,000,000 molecules of hydrogen peroxide to water and oxygen per enzyme per second. It is abundantly found in peroxisomes.

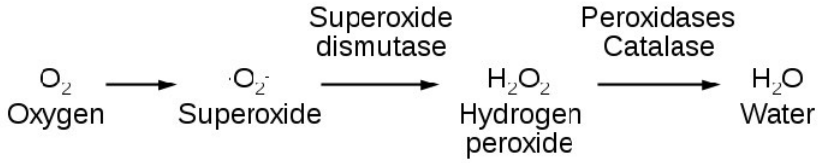
In addition to catalase's ability to break down hydrogen peroxide, the enzyme can also use hydrogen peroxide to oxidize a wide variety organic compounds, including phenols, formic acid, formaldehyde, acetaldehyde, and alcohols, but with much lower efficiency.

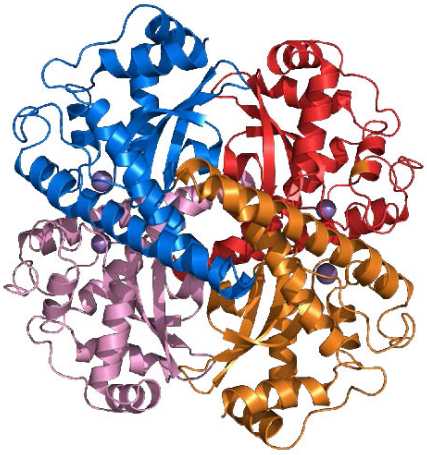
Health

The importance of catalase for health is uncertain. Mice deficient in the enzyme appear healthy and humans with low levels of the enzyme display few problems. On the other hand, mice engineered to produce higher levels of catalase, in at least one study, lived longer. The ability of organisms to live with lower levels or no catalase may arise from another group of enzymes, the peroxiredoxins, which also act on hydrogen peroxide and may make up for lower quantities of catalase. Last, there is evidence that reduced levels of catalase with aging may be responsible for the graying of hair. Higher levels of H2O2 with reduced catalase results in a bleaching of hair follicles.

Superoxide dismutase



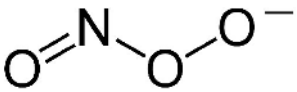




Another important enzyme for protection against reactive oxygen species is superoxide dismutase (SOD), which is found, like catalase, in virtually all organisms living in an oxygen environment. Superoxide dismutase, also like catalase, has a very high Kcat value and, in fact, has the highest Kcat/Km known for any known enzyme. It catalyzes the reactions at the top of the next column (superoxides shown in red):

The enzyme thus works by a ping-pong (double displacement) mechanism (see HERE), being converted between two different forms.

The hydrogen peroxide produced in the second reaction is easily handled by catalase and is also less harmful than superoxide, which can react with nitric oxide (NO) to form very toxic peroxynitrite ions (Figure 5.43). Peroxynitrite has negative effects on cells.



In addition to copper, an ion of Zn++ is also bound by the enzyme and likely plays a role in the catalysis. Forms of superoxide dismutase that use manganese, nickel, or iron are also known and are mostly found in prokaryotes and protists, though a manganese SOD is found in most mitochondria. Copper/zinc enzymes are common in eukaryotes.

Three forms of superoxide dismutase are found in humans and localized to the cytoplasm (SOD1), mitochondria (SOD2), and extracellular areas (SOD3). Mice lacking any of the three forms of the enzyme are more sensitive to drugs, such as paraquat. Deficiency of SOD1 in mice leads to hepatocellular carcinoma and early loss of muscle tissue related to aging. Drosophila lacking SOD2 die before birth and those lacking SOD1 prematurely age.

In humans, superoxide dismutase mutations are associated with the genetically-linked form of Amyotrophic Lateral Sclerosis (ALS) and over-expression of the gene is linked to neural disorders associated with Down syndrome.

Mixed function oxidases

Other enzymes catalyzing reactions involving oxygen include the mixed function oxidases. These enzymes use molecular oxygen for two different purposes in one reaction. The mixed function part of the name is used to indicate reactions in which two different substrates are being oxidized simultaneously. Monooxygenases are examples of mixed function oxidases. An example of a mixed function oxidase reaction is shown below.

AH + BH2 + O2 <=> AOH + B + H2O

In this case, the oxygen molecule has one atom serve as an electron acceptor and the other atom is added to the AH, creating an alcohol.

Cytochrome P450 enzymes

Cytochrome P450 enzymes (called CYPs) are family of heme-containing mixed function oxidase enzymes found in all domains of life. Over 21,000 CYP enzymes are known. The most characteristic reaction catalyzed by these enzymes follows

Monooxygenase reactions such as this are relatively rare in the cell due to their use of molecular oxygen. CYPs require an electron donor for reactions like the one shown here and frequently require a protein to assist in transferring electrons to reduce the heme iron. There are six different classes of P450 enzymes based on how they get electrons

1. Bacterial P450 - electrons from ferredoxin reductase and ferredoxin

2. Mitochondrial P450 - electrons from adrenodoxin reductase and adrenodoxin

3. CYB5R/cyb5 - electrons come from cytochrome b5

4. FMN/Fd - use a fused FMN reductase

5. Microsomal P450 - NADPH electrons come via cytochrome P450 reductase or from cytochrome b5 and cytochrome b5 reductase

6. P450 only systems - do not require external reducing power

The CYP genes are abundant in humans and catalyze thousand of reactions on both cellular and extracellular chemicals. There are 57 human genes categorized into 18 different families of enzymes. Some CYPs are specific for one or a few substrates, but others can act on many different substrates.

CYP enzymes are found in most body tissues and perform important functions in synthesis of steroids (cholesterol, estrogen, testoterone, Vitamin D, e.g.), breakdown of endogenous compounds (bilirubin), and in detoxification of toxic compounds including drugs. Because they act on many drugs, changes in CYP activity can produce unexpected results and cause problems with drug interactions.

Bioactive compounds, for example, in grapefruit juice, can inhibit CYP3A4 activity, leading to increased circulating concentrations of drugs that would normally have been acted upon by CYP3A4. This is the reason that patients prescribed drugs that are known to be CYP3A4 substrates are advised to avoid drinking grapefruit juice while under treatment. St. Johns Wort, an herbal treatment, on the other hand, induces CYP3A4 activity, but inhibits CYP1A1, CYP1B1, and CYP2D6. Tobacco smoke induces CYP1A2 and watercress inhibits CYP2E1.

Cytochromes

Cytochromes are heme-containing proteins that play major roles in the process of electron transport in the mitochondrion and in photosynthesis in the chloroplast. They exist either as monomers (cytochrome c) or as subunits within large redox complexes (Complex III and Complex IV of electron transport. An atom of iron at the center of the heme group plays a central role in the process, flipping between the ferrous (Fe++) and ferric (Fe+++) states as a result of the movement of electrons through it.

There are several different cytochromes. Cytochrome c is a soluble protein loosely associated with the mitochondrion. Cytochromes a and a3 are found in Complex IV. Complex III has cytochromes b and c1 and the plastoquinol-plastocyanin reductase of the chloroplast contains cytochromes b6 and f. Another important class of enzymes containing cytochromes is the cytochrome P450 oxidase group (see above). They get their name from the fact that they absorb light at 450 nm when their heme iron is reduced.

Iron-Sulfur Proteins

Iron-sulfur proteins contain iron-sulfur clusters in a variety of formats, including sulfide-linked di-, tri-, and tetrairon centers existing in different oxidation states (Figures 5.48 & 5.49). The clusters play a variety of roles, but the best known ones are in electron transport where they function in the redox reactions involved in the movement of electrons.

Complexes I and Complex II contain multiple Fe-S centers. Iron-sulfur proteins, though, have many other roles in cells. Aconitase uses an iron-sulfur center in its catalytic action and the ability of the enzyme to bind iron allows it to function as a barometer of iron concentration in cells. Iron-sulfur centers help to generate radicals in enzymes using S-Adenosyl Methionine (SAM) and can serve as a source of sulfur in the synthesis of biotin and lipoic acid. Some iron-sulfur proteins even help to regulate gene expression.

DNA damage theory of aging

The DNA Damage Theory of Aging is based on the observation that, over time, cells are subject to extensive oxidative events. As already noted, these afford opportunities for the formation of ROS that can damage cellular molecules, and it follows that accumulation of such damage, especially to the DNA would be deleterious to the cell. The build-up of DNA damage could, thus, be responsible for the changes in gene expression that we associate with aging.

Numerous damage events

The amount of DNA damage that can occur is considerable. In mice, for example, it is estimated that each cell experiences 40,000 to 150,000 damage events per day. The damage, which happens to nuclear as well as to mitochondrial DNA, can result in apoptosis and/or cellular senescence. DNA repair systems, of course, protect against damage to DNA, but over time, unrepairable damage may accumulate.

Antioxidants

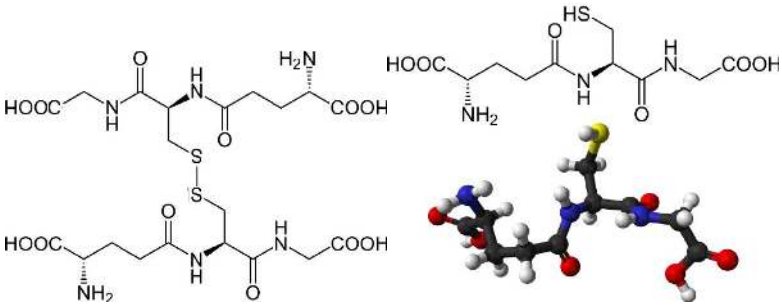
There is a growing interest in the subject of antioxidants because of health concerns raised by our knowledge of problems created as a result of spontaneous oxidation of biomolecules by Reactive Oxygen Species (ROS), such as superoxide. Antioxidants have the chemical property of protecting against oxidative damage by being readily oxidized themselves, preferentially to other biomolecules.

Biologically, cells have several lines of antioxidant defense. They include molecules, such as vitamins C, A, and E, glutathione, and enzymes that destroy ROS such as superoxide dismutase, catalase, and peroxidases.

Health effects

Oxidation by ROS is mutagenic and has been linked to atherosclerosis. Nonetheless, randomized studies of oral supplementation of various vitamin combinations have shown no protective effect against cancer and supplementation of Vitamin E and selenium has revealed no decrease in the risk of cardiovascular disease. Further, no reduction in mortality rates as a result of supplementation with these materials has been found, so the protective effects, if any, of antioxidants on ROS in human health remain poorly understood.

Glutathione



The major endogenous antioxidant found in cells spanning most living systems, glutathione is a tripeptide protecting cells against damage caused by reactive oxygen species and heavy metals (Figures 5.55 & 5.56). The three amino acids in glutathione (glutamate, cysteine, and glycine) are joined in an unusual fashion. The glutamate is joined to the center cysteine by a peptide bond between the R-group carboxyl of glutamate and the α-amine of cysteine. The bond between cysteine and glycine is a normal peptide bond between the α-carboxyl of cysteine and the α-amine of glycine.

The thiol group of cysteine is a reducing agent that reduces disulfide bonds to sulfhydryls in cytoplasmic proteins. This, in turn, is the bridge when two glutathiones get oxidized and form a disulfide bond with each other (Figure 5.56). Glutathione's two oxidative states are abbreviated as follows: GSH (reduced) and GSSG (oxidized).

Disulfide-joined glutathiones can be separated by reduction of their bonds with glutathione reductase, using electrons from NADPH for the reduction.

Essential for life

Glutathione is important for life. Mice lacking the first enzyme involved in its synthesis in the liver die in the first month after birth. In healthy cells, 90% of glutathione is in the GSH state. Higher levels of GSSG correspond to cells that are oxidatively stressed.

Besides reducing disulfide bonds in cells, glutathione is also important for the following:

• Neutralization of free radicals and reactive oxygen species.

• Maintenance of exogenous antioxidants such as vitamins C and E in their reduced forms.

Resveratrol

Categorized as a stilbenoid, resveratrol (Figure 5.57) is a phenolic compound produced in the skin of plants such as grapes, raspberries, and blueberries, in response to injury or when they are being attacked by pathogens. Numerous health benefits are claimed for the compound, though evidence of such benefits is in short supply. Resveratrol is metabolized rapidly in the body, so it is difficult to maintain levels of it.

Some data indicates resveratrol may improve the functioning of mitochondria. It also acts as an antioxidant and causes concentration of another anti-oxidant, glutathione, to increase. The compound appears to induce expression of manganese superoxide dismutase (protects against reactive oxygen species) and inhibits several phosphodiesterases. This causes an increase in cAMP which results in increases in oxidation of fatty acids, mitochondria formation, gluconeogenesis, and glycogen breakdown. It has been claimed to be the cause of the French Paradox in which drinking of red wine is supposed to give protection for the cardiovascular system. Research data is lacking in support of the claim, however. Resveratrol is known to activate Sirtuin proteins, which play roles in gene inactivation.

Summary

In summary, energy is needed for cells to perform the functions that they must carry out in order to stay alive. At its most basic level, this means fighting a continual battle with entropy, but it is not the only need for energy that cells have.

References

1. Winge, D.R., Mol Cell Biol. 2012 Jul; 32(14): 2647–2652. doi: 10.1128/MCB.00573-12

Energy: Electron Transport & Oxidative Phosphorylation

429

YouTube Lectures

by Kevin

HERE & HERE

YouTube Lectures

by Kevin

HERE & HERE

430

Figure 5.14 - Overview of electron transport (bottom left and top right) and oxidative phosphorylation (top left - yellow box) in the mitochondrion

431

Figure 5.15 - Loss of electrons by NADH to form NAD+. Relevant reactions occur in the top ring of the molecule.

432

Figure 5.16 - Flow of electrons from NADH into the electron transport system. Entry is through complex I

Image by Aleia Kim

Figure 5.17 - Flow of electrons from FADH2 into the electron transport chain. Entry is through complex II.

Image by Aleia Kim

Interactive Learning

Module

HERE

433

Figure 5.18 - Complex I embedded in the inner mitochondrial membrane. The mitochondrial matrix at at the top

Wikipedia

434

Figure 5.19 - Complex II embedded in inner mitochondrial membrane. Matrix is up.

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

435

Figure 5.20 - Movement of electrons through complex I from NADH to coenzyme Q. The mitochondrial matrix is at the bottom

Image by Aleia Kim

Figure 5.21 - Movement of electrons from succinate through complex II (A->B->C->D->Q). Mitochondrial matrix on bottom.

Image by Aleia Kim

436

Figure 5.22 - Complex II in inner mitochondrial membrane showing electron flow. Matrix is up.

Wikipedia

Figure 5.23 - Coenzyme Q

437

Movie 5.2 - The Q-cycle

Wikipedia

Figure 5.24 - The Q-Cycle Image by Aleia Kim

Figure 5.24 - Complex III

Wikipedia

438

YouTube Lectures

by Kevin

HERE & HERE

Figure 5.25 - The Q-cycle. Matrix is down.

Image by Aleia Kim

439

Figure 5.26 - Movement of electrons and protons through complex IV. Matrix is down

Image by Aleia Kim

Figure 5.25 - Cytochrome c with bound heme Group

Wikipedia

440

Figure 5.27 - Mitochondrial anatomy. Electron transport complexes and ATP synthase are embedded in the inner mitochondrial membrane

Image by Aleia Kim

441

Figure 5.28 - ATP synthase. Protons pass from intermembrane space (top) through the complex and exit in the matrix (bottom).

Image by Aleia Kim

Interactive Learning

Module

HERE

442

Movie 5.3 - ATP Synthase - ADP + Pi (pink) and ATP (red). The view is end-on from the cytoplasmic side viewing the β subunits

Movie 5.3 - ATP Synthase - ADP + Pi (pink) and ATP (red). The view is end-on from the cytoplasmic side viewing the β subunits

443

Figure 5.29 - Important structural features of the ATP synthase

Image by Aleia Kim

444

Figure 5.30 - Loose (L), Tight (T), and Open (O) structures of the F1 head of ATP synthase. Change of structure occurs by rotation of γ-protein (purple) in center as a result of proton movement. Individual α and β units do not rotate

Image by Aleia Kim

445

Figure 5.31 - Respiration overview in eukaryotic cells

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

446

Rest

ATP High / ADP Low

Oxidative Phosphorylation Low

Electron Transport Low

Oxygen Use Low

NADH High / NAD+ Low

Citric Acid Cycle Slow

Exercise

ATP Low / ADP High

Oxidative Phosphorylation High

Electron Transport High

Oxygen Use High

NADH Low / NAD+ High

Citric Acid Cycle Fast

Interactive Learning

Module

HERE

447

Figure 5.32 - Three inhibitors of electron transport

Image by Aleia Kim

448

Figure 5.33 - Oligomycin A - An inhibitor of ATP synthase

Figure 5.34 - 2,4 DNP - an uncoupler of respiratory control

449

In Cells With Tight Coupling

O2 use depends on metabolism
NAD+ levels vary with exercise
Proton gradient high with no exercise
Catabolism depends on energy needs
ETS runs when OxPhos runs and vice versa

In Cells That Are Uncoupled

O2 use high
NAD+ Levels high
Little or no proton gradient
Catabolism high
OxPhos does not run, but ETS runs rapidly

YouTube Lectures

by Kevin

HERE & HERE

450

451

Figure 5.35 - Alternative oxidase (AOX) of fungi, plants, and protozoa bypasses part of electron transport by taking electrons from CoQ and passing them to oxygen.

452

Figure 5.36 - Structure of an oxygen free radical

Wikipedia

NADPH + 2O2

NADP+ + 2O2− + H+

Figure 5.37 - Three sources of reactive oxygen species (ROS) in cells

Wikipedia

453

454

YouTube Lectures

by Kevin

HERE & HERE

Figure 5.38 A hydroxyl radical

Wikipedia

455

Reduced Glutathione (GSH) + H2O2

Oxidized Glutathione (GSSG) + H2O

Figure 5.40 - Detoxifying reactive oxygen species

Figure 5.39 - Catalase

456

1. O2- + Enzyme-Cu++

O2 + Enzyme-Cu+

2. O2- + Enzyme-Cu+ + 2H+

H2O2 + Enzyme-Cu++

Figure 5.41 - SOD2 of humans

Figure 5.42 3 - Peroxynitrite Ion

Figure 5.44 - SOD1 of humans

Wikipedia

Figure 5.45 - SOD3 of humans

457

Figure 5.43 - Peroxynitrite's effects on cells lead to necrosis or apoptosis

Wikipedia

458

RH + O2 + NADPH + H+

ROH + H2O + NADP+

459

Figure 5.46 - Cytochrome c with its heme group

460

YouTube Lectures

by Kevin

HERE & HERE

Figure 5.47 - Fe2S2 Cluster

Figure 5.48 - Redox reactions for Fe4S4 clusters

461

Figure 5.49 - Tyramine

Figure 5.50 - Phenethylamine

462

Figure 5.51 - Guanine and 8-oxo-guanine

Figure 5.52 - Adenine-8-oxo-guanine base pair. dR = deoxyribose

463

Figure 5.53 - Good antioxidant sources

464

Figure 5.55 - Oxidized glutathiones (GSSG) joined by a disulfide bond

Wikipedia

Figure 5.54 - Structure of reduced glutathione (GSH)

465

Figure 5.56 - Resveratrol

YouTube Lectures

by Kevin

HERE & HERE

466

Graphic images in this book were products of the work of several talented students. Links to their Web pages are below

Click HERE for

Martha Baker’s

Web Page

Click HERE for

Pehr Jacobson’s

Web Page

Click HERE for

Aleia Kim’s

Web Page

Click HERE for

Penelope Irving’s

Web Page

Problem set related to this section HERE

Point by Point summary of this section HERE

To get a certificate for mastering this section of the book, click HERE

Kevin Ahern’s free iTunes U Courses - Basic / Med School / Advanced

Biochemistry Free & Easy (our other book) HERE / Facebook Page

Kevin and Indira’s Guide to Getting into Medical School - iTunes U Course / Book

To see Kevin Ahern’s OSU ecampus courses - BB 350 / BB 450 / BB 451

To register for Kevin Ahern’s OSU ecampus courses - BB 350 / BB 450 / BB 451

Biochemistry Free For All Facebook Page (please like us)

Kevin Ahern’s Web Page / Facebook Page / Taralyn Tan’s Web Page

Kevin Ahern’s free downloads HERE

OSU’s Biochemistry/Biophysics program HERE

OSU’s College of Science HERE

Oregon State University HERE

Email Kevin Ahern / Indira Rajagopal / Taralyn Tan

I'm a little mitochondrion
Who gives you energy
I use my proton gradient
To make the ATPs

He's a little mitochondrion
Who gives us energy
He uses proton gradients
To make some ATPs

Electrons flow through Complex II
To traffic cop Co-Q
Whenever they arrive there in
An FADH-two

Electrons flow through Complex II
To traffic cop Co-Q
Whenever they arrive there in
An FADH-two

Tightly coupled is my state
Unless I get a hole
Created in my membrane by
Some di-ni-tro-phe-nol

Yes tightly coupled is his state
Unless he gets a hole
Created in his membrane by
Some di-ni-tro-phenol

Both rotenone and cyanide
Stop my electron flow
And halt the calculation of
My "P" to "O" ratio

Recording by Tim Karplus

Lyrics by Kevin Ahern
Recording by Tim Karplus Lyrics by Kevin Ahern

I’m a Little Mitochondrion

To the tune of “I’m a Lumberjack”

Metabolic Melodies Website HERE

In the catabolic pathways that our cells employ
Oxidations help create the ATP
While they lower Gibbs free energy
Thanks to enthalpy

If a substrate is converted from an alcohol
To an aldehyde or ketone it is clear
Those electrons do not disappear
They just rearrange – very strange

N-A-D is in my ears and in my eyes
Help-ing mol-e-cules get oxidized
Making N-A-D-H then

And the latter is a problem anaerobically
‘Cuz accumulations of it muscles hate
They respond by using pyruvate
To produce lactate

Catalyzing is essential for the cells to live
So the enzymes grab their substrates eagerly
If they bind with high affinity

Low Km you see, just as me

N-A-D is in my ears and in my eyes
Help-ing mol-e-cules get oxidized
Making N-A-D-H then

N-A-D

To the tune of “Penny Lane”

Metabolic Melodies Website HERE

Recorded by Tim Karplus

Lyrics by Kevin Ahern
Recorded by Tim Karplus Lyrics by Kevin Ahern

When oxygen’s electrons all are in the balanced state

There’s twelve of them for oh-two. The molecule is great

But problems sometimes happen on the route to complex IV

Making reactive species that the cell cannot ignore

Oh superoxide dismutase is super catalytic

Keeping cells from getting very peroxynitritic

Faster than a radical, its actions are terrific

Superoxide dismutase is super catalytic

Enzyme, enzyme deep inside

Blocking all the bad oxides

The enzyme’s main advantage is it doesn’t have to wait

By binding superoxide in a near-transition state

It turns it to an oxygen in mechanism one

Producing “h two oh two” when the cycle is all done

Oh superoxide dismutase you’re faster than all them

You’ve got the highest ratio of kcat over KM

This means that superoxide cannot cause too much mayhem

Superoxide dismutase is faster than all them

Superoxide dismutase

Stopping superoxide’s ways

The enzyme’s like a ping-pong ball that mechanistic-ly

Bounces between two copper states, plus one and two you see

So S-O-D behaves just like an anti-oxidant

Giving as much protection as a cell could ever want

Oh superoxide dismutase, the cell’s in love with you

Because you let electron transport do what it must do

Without accumulation of a radical oh two

Superoxide dismutase - that's why a cell loves you

Superoxide Dismutase

To the tune of "Supercalifragilistiexpialidocious"

Metabolic Melodies Website HERE

Lyrics by Kevin Ahern

No Recording Yet For This Song

11.1: ETC and Oxidative Phosphorylation is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

11.1: Exercises

Problems

1. How does ETC complexes transfer electrons?

2. Where does the NADH and [$FADH_2$] come from that drives ETC.

3.When the proton gradient is created because of the ETC where is the highest concentration of protons found?

4. What are the three electron carriers responsible for transporting protons?

5. Describe the purpose and name each electron carriers in the ETC

6. How many protons does it take to produce one ATP via oxidative phosphorolation?

7. Briefly describe the chemiosmotic theory of generation of ATP as a result of an electron transport chain.

8. Compare where the electron transport chain occurs in prokaryotic cells and in eukaryotic cells.

9. State the function of ATP synthases in chemiosmosis.

10. State the final electron acceptor and the end product formed at the end of aerobic respiration.

11. Fill in the blanks with appropriate choices from the given list.

____a protein that forms a channel to separate oxidative phosphorylation from ATP synthesis

____inhibitor of complex 1

____inhibitor of complex 3

____inhibitor of complex 4

____directly receiving electrons from quinone (reduced form)

____directly receiving electrons from complex 1

____rotational catalysis with inequivalent nucleotide binding sites

____scavenging radicals to counteract oxidative damage

A. DNP

B. succcinate dehydrogenase

C. UCP-1

D. rotenone

E. cytochrome c oxidase complex

F. ATP synthase

G. ubiquinone

H. superoxide dismutase

I. antimycin A

J. CN^-

K. NADH-Q reductase complex

L. cytochrome c reductase (oxidoreductase)

12. Which of the following is true regarding chemiosmosis?

A. The energy from a proton gradient is used to make ATP.

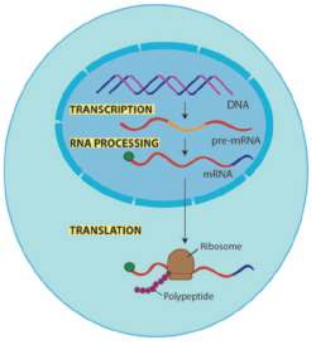
B. Chemiosmosis regenerates electron carriers like NADH AND $FADH_2$.

C. ATP synthesis creates a proton gradient that causes electron flow through an electron transport chain (ETC).

D. A temperature gradient drives ATP synthesis.

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

12: The Flow of Genetic Information: from DNA to RNA and Proteins



DNA $\xrightarrow{\text{Transcription}}$ **RNA** $\xrightarrow{\text{Translation}}$ **Protein**

12: The Flow of Genetic Information: from DNA to RNA and Proteins is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

12.1: The Structure of DNA

A quick look at the whole structure of DNA

DNA is a complex molecule which carries the genetic code. The diagram below shows a tiny bit of a DNA double helix.

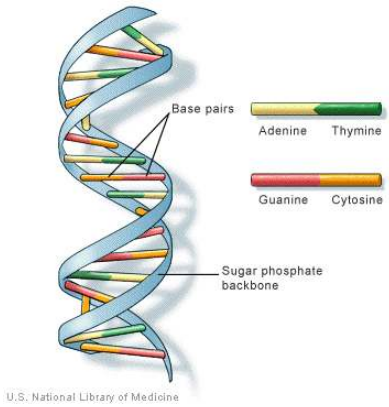
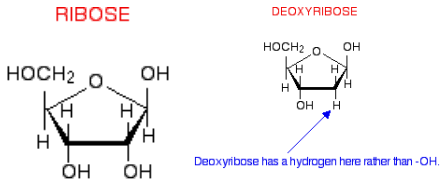
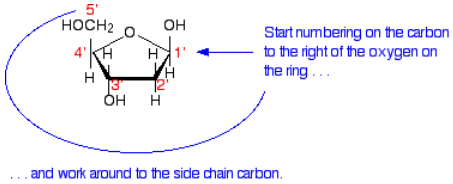


Figure 12.1.1: *The double helical structure of DNA*

The backbone of DNA is based on a repeated pattern of a sugar group and a phosphate group. The full name of DNA, deoxyribonucleic acid, because of the presence of sugar - deoxyribose. Deoxyribose is a modified form of another sugar called ribose. Ribose is the sugar in the backbone of RNA, ribonucleic acid.

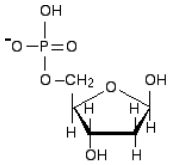


The carbon atom to the right of the oxygen as we have drawn the ring is given the number 1, and then you work around to the carbon on the CH_2OH side group which is number 5.



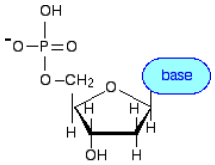
Attaching a phosphate group

The other repeating part of the DNA backbone is a phosphate group. A phosphate group is attached to the sugar molecule in place of the -OH group on the 5' carbon.

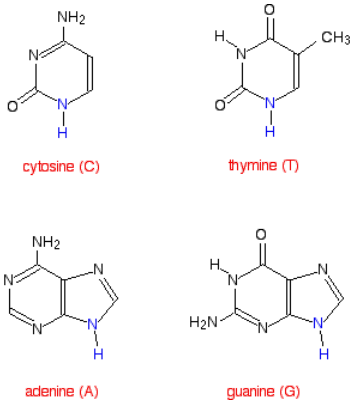


Attaching a base and making a nucleotide

In DNA, these bases are cytosine (C), thymine (T), adenine (A) and guanine (G).These bases attach in place of the -OH group on the 1' carbon atom in the sugar ring.

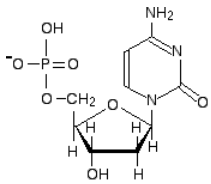


What we have produced is known as a nucleotide. We now need a quick look at the four bases.



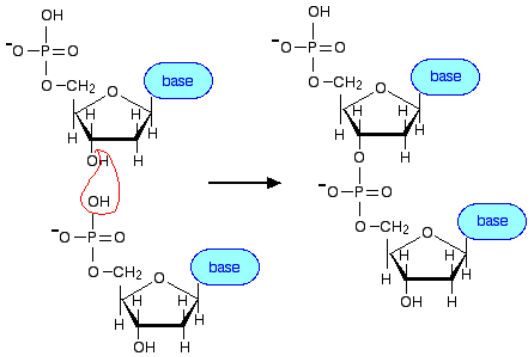
The nitrogen and hydrogen atoms shown in blue on each molecule show where these molecules join on to the deoxyribose. In each case, the hydrogen is lost together with the -OH group on the 1' carbon atom of the sugar. This is a condensation reaction - two molecules joining together with the loss of a small one (not necessarily water).

For example, here is what the nucleotide containing cytosine would look like:



Joining the nucleotides into a DNA strand

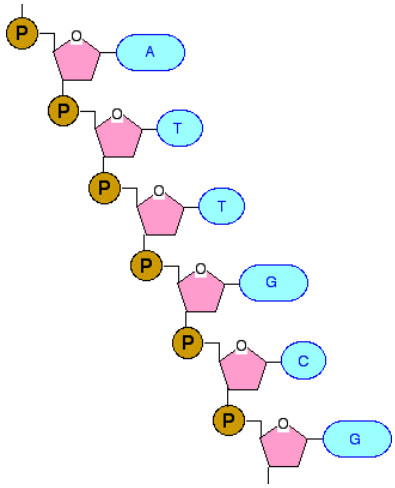
A DNA strand is simply a string of nucleotides joined together. The phosphate group on one nucleotide links to the 3' carbon atom on the sugar of another one. In the process, a molecule of water is lost - another condensation reaction.



Building a DNA chain concentrating on the essentials

What matters in DNA is the sequence the four bases take up in the chain. We aren't particularly interested in the backbone, so we can simplify that down. For the moment, we can simplify the precise structures of the bases as well. We can build the chain based on this fairly obvious simplification:

The diagram below is a bit from the middle of a chain. Notice that the individual bases have been identified by the first letters of the base names. (A = adenine, etc). Notice also that there are two different sizes of base. Adenine and guanine are bigger because they both have two rings. Cytosine and thymine only have one ring each.



Joining the two DNA chains together

If you look at this carefully, you will see that an adenine on one chain is always paired with a thymine on the second chain. And a guanine on one chain is always paired with a cytosine on the other one.

So how exactly does this work?

The first thing to notice is that a smaller base is always paired with a bigger one. The effect of this is to keep the two chains at a fixed distance from each other all the way along.

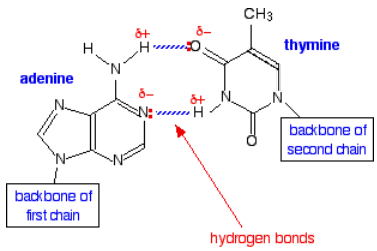
But, more than this, the pairing has to be exactly . . .

- adenine (A) pairs with thymine (T);
- guanine (G) pairs with cytosine (C).

That is because these particular pairs fit exactly to form very effective hydrogen bonds with each other. It is these hydrogen bonds which hold the two chains together.

The base pairs fit together as follows.

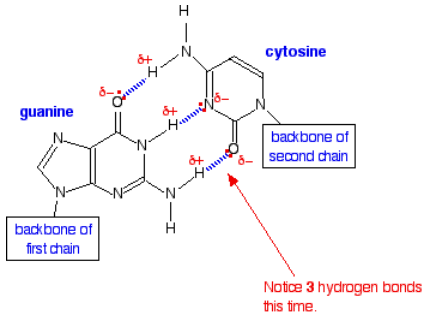
The A-T base pair:

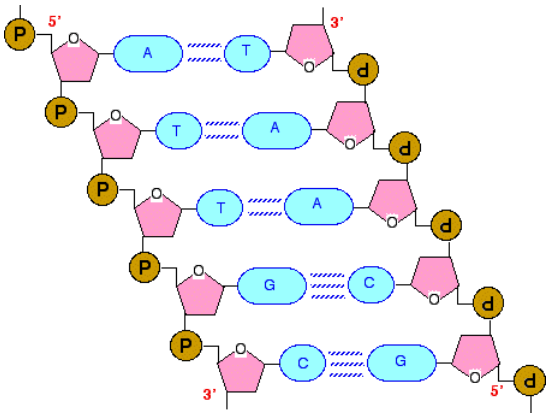


The G-C base pair:

A final structure for DNA showing the important bits

Notice that the two chains run in opposite directions, and the right-hand chain is essentially upside-down with 3' and 5'.





If you followed the left-hand chain to its very end at the top, you would have a phosphate group attached to the 5' carbon in the deoxyribose ring. If you followed it all the way to the other end, you would have an -OH group attached to the 3' carbon.

In the second chain, the top end has a 3' carbon, and the bottom end a 5'. This 5' and 3' notation becomes important when we start talking about the genetic code and genes. The genetic code in genes is always written in the 5' to 3' direction along a chain.

Contributors

Jim Clark (Chemguide.co.uk)

This page titled 12.1: The Structure of DNA is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Jim Clark.

12.2: DNA Replication

According to the central dogma of molecular genetics, DNA is the genetically active component of the chromosomes of a cell. That is, DNA in the cell nucleus contains all the information necessary to control synthesis of the proteins, enzymes, and other molecules which are needed as that cell grows, carries on metabolism, and eventually reproduces. Thus when a cell divides, its DNA must pass on genetic information to both daughter cells. It must somehow be able to divide into duplicate copies. This process is called **replication**.

Given the complementary double strands of DNA, it is relatively easy to see how DNA as a molecule is well structured for replication. Each strand serves as a template for a new strand. Thus, after DNA is replicated, each new DNA double helix will have one strand from the original DNA molecule, and one newly synthesized molecule. This is referred to as semiconservative replication.

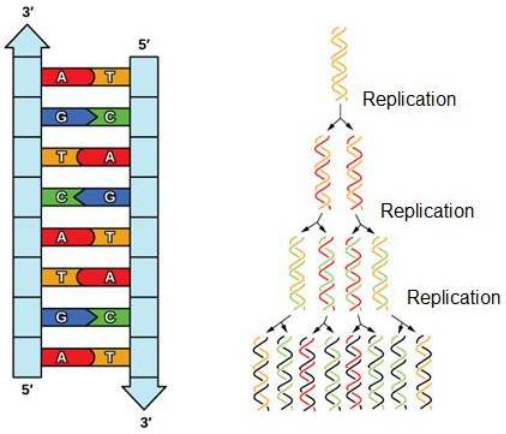


Figure 12.2.1 DNA has an anti-parallel double helix structure, the nucleotide bases are hydrogen bonded together and each strand complements the other. DNA is replicated in a semi-conservative manner, each strand is used as the template for the newly made strand.

Because eukaryotic genomes are quite complex, DNA replication is a very complicated process that involves several enzymes and other proteins. It occurs in three main stages: initiation, elongation, and termination.

Key Points

- During initiation, proteins bind to the origin of replication while helicase unwinds the DNA helix and two replication forks are formed at the origin of replication.
- During elongation, a primer sequence is added with complementary RNA nucleotides, which are then replaced by DNA nucleotides.
- During elongation the leading strand is made continuously, while the lagging strand is made in pieces called Okazaki fragments.
- During termination, primers are removed and replaced with new DNA nucleotides and the backbone is sealed by DNA ligase
- origin of replication: a particular sequence in a genome at which replication is initiated
- leading strand: the template strand of the DNA double helix that is oriented so that the replication fork moves along it in the 3′ to 5′ direction
- lagging strand: the strand of the template DNA double helix that is oriented so that the replication fork moves along it in a 5′ to 3′ manner

Initiation of Replication

The replication begins at specific nucleotide sequences called **origins of replication** along the length of the DNA. The circular *E. coli* chromosome has just one of these sites; the linear eukaryotic chromosomes, in contrast, have multiple sites on every chromosome. Once this site is identified, proteins generally called "initiators" have the capacity to bind DNA at or very near the DNA sequences that mark the origins of replication. The interaction of the initiator proteins with the DNA helps to destabilize the double helix and also to recruit other proteins, including an enzyme called a **helicase**. In this case the energy required to destabilize

the DNA double helix seems to come from the formation of new associations between DNA and the initiator proteins. The DNA helicase, in contrast, once loaded onto the origin, couples the exergonic hydrolysis of ATP to the unwinding of the DNA double helix. Additional proteins must be recruited to the partially unwound initiation complex. These include, but are not limited to, enzymes called **primase** and **DNA polymerase**. While the initiators are lost soon after the initiation of replication, the rest of the proteins work in concert to execute the process of DNA replication. This complex of enzymes function at Y-shaped structures in the DNA called **replication forks** (see figure below). For any replication event two replication forks may be formed at each origin of replication, extending in both directions.

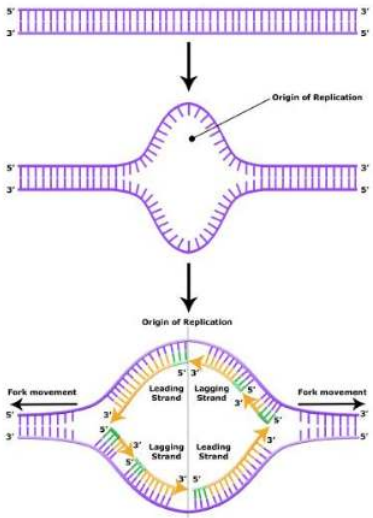


Figure 12.2.2**:** At the origin of replication, a replication bubble forms. The replication bubble is composed of two replication forks, each traveling in opposite directions along the DNA. The replication forks include all of the enzymes required for replication to occur.

Elongation of Replication

During elongation, an enzyme called DNA polymerase adds DNA nucleotides to the 3′ end of the newly synthesized polynucleotide strand. The template strand specifies which of the four DNA nucleotides (A, T, C, or G) is added at each position along the new chain. Only the nucleotide complementary to the template nucleotide at that position is added to the new strand.

DNA polymerase contains a groove that allows it to bind to a single-stranded template DNA and travel one nucleotide at at time. For example, when DNA polymerase meets an adenosine nucleotide on the template strand, it adds a thymidine to the 3′ end of the newly synthesized strand, and then moves to the next nucleotide on the template strand. This process will continue until the DNA polymerase reaches the end of the template strand.

DNA polymerase cannot initiate new strand synthesis; it only adds new nucleotides at the 3′ end of an existing strand. All newly synthesized polynucleotide strands must be initiated by a specialized RNA polymerase called primase. Primase initiates polynucleotide synthesis and by creating a short RNA polynucleotide strand complementary to template DNA strand. This short stretch of RNA nucleotides is called the primer. Once RNA primer has been synthesized at the template DNA, primase exits, and DNA polymerase extends the new strand with nucleotides complementary to the template DNA.

Eventually, the RNA nucleotides in the primer are removed and replaced with DNA nucleotides. Once DNA replication is finished, the daughter molecules are made entirely of continuous DNA nucleotides, with no RNA portions.

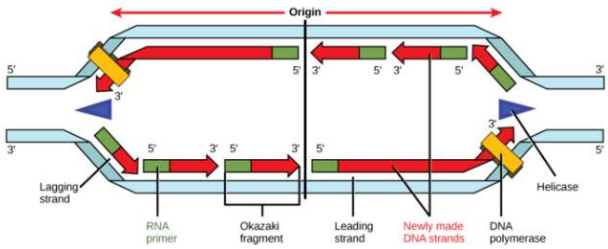


Figure 12.2.3: *Elongation of replication*

The Leading and Lagging Strands

DNA polymerases can only make DNA in the 5' to 3' direction, and this poses a problem during replication. A DNA double helix is always anti-parallel; in other words, one strand runs in the 5' to 3' direction, while the other runs in the 3' to 5' direction. This makes it necessary for the two new strands, which are also antiparallel to their templates, to be made in slightly different ways.

One new strand, which runs 5' to 3' towards the replication fork, is the easy one. This strand is made continuously, because the DNA polymerase is moving in the same direction as the replication fork. This continuously synthesized strand is called the **leading strand**.

The other new strand, which runs 5' to 3' away from the fork, is trickier. This strand is made in fragments because, as the fork moves forward, the DNA polymerase (which is moving away from the fork) must come off and reattach on the newly exposed DNA. This tricky strand, which is made in fragments, is called the **lagging strand**.

The small fragments are called **Okazaki fragments**, named for the Japanese scientist who discovered them. The leading strand can be extended from one primer alone, whereas the lagging strand needs a new primer for each of the short Okazaki fragments.

Termination of Replication

Eukaryotic chromosomes have multiple origins of replication, which initiate replication almost simultaneously. Each origin of replication forms a bubble of duplicated DNA on either side of the origin of replication. Eventually, the leading strand of one replication bubble reaches the lagging strand of another bubble, and the lagging strand will reach the 5′ end of the previous Okazaki fragment in the same bubble.

DNA polymerase halts when it reaches a section of DNA template that has already been replicated. However, DNA polymerase cannot catalyze the formation of a phosphodiester bond between the two segments of the new DNA strand, and it drops off. These unattached sections of the sugar-phosphate backbone in an otherwise full-replicated DNA strand are called nicks.

Once all the template nucleotides have been replicated, the replication process is not yet over. RNA primers need to be replaced with DNA, and nicks in the sugar-phosphate backbone need to be connected.

The group of cellular enzymes that remove RNA primers include the proteins FEN1 (flap endonulcease 1) and RNase H. The enzymes FEN1 and RNase H remove RNA primers at the start of each leading strand and at the start of each Okazaki fragment, leaving gaps of unreplicated template DNA. Once the primers are removed, a free-floating DNA polymerase lands at the 3′ end of the preceding DNA fragment and extends the DNA over the gap. However, this creates new nicks (unconnected sugar-phosphate backbone).

In the final stage of DNA replication, the enzyme **DNA ligase** joins the sugar-phosphate backbones at each nick site. After ligase has connected all nicks, the new strand is one long continuous DNA strand, and the daughter DNA molecule is complete.

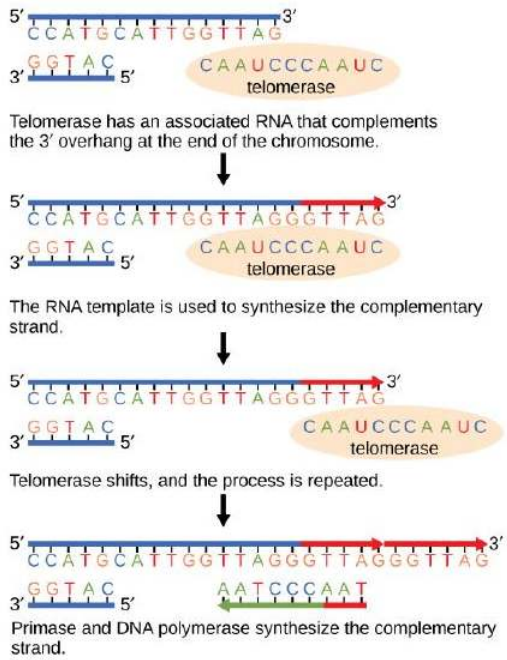


Figure 12.2.4: Termination of replication

References

1. ↑ Mariana Ruiz. Wikimedia Commons. 24 January 2007. Retrieved: 11 August 2009.
2. ↑ Nelson, D.L., Cox, M.M. *Lehninger Principles of Biochemistry(5thed).* New York: W.H. Freeman and Company, 2008. pp. 985-991.
3. ↑ Nelson, D.L., Cox, M.M. *Lehninger Principles of Biochemistry(5thed).* New York: W.H. Freeman and Company, 2008. pp. 982-991.

Contributors

- Ed Vitz (Kutztown University), John W. Moore (UW-Madison), Justin Shorb (Hope College), Xavier Prat-Resina (University of Minnesota Rochester), Tim Wendorff, and Adam Hahn.

This page titled 12.2: DNA Replication is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Ed Vitz, John W. Moore, Justin Shorb, Xavier Prat-Resina, Tim Wendorff, & Adam Hahn.

12.3: DNA Repair

Maintaining the Integrity of the Cell's Information: DNA Repair

All DNA suffers damage over time, from exposure to ultraviolet and other radiation, as well as from various chemicals in the environment. Even chemical reactions naturally occurring within cells can give rise to compounds that can damage DNA. As you already know, even minor changes in DNA sequence, such as point mutations can sometimes have far-reaching consequences. Likewise, unrepaired damage caused by radiation, environmental chemicals or even normal cellular chemistry can interfere with the accurate transmission of information in DNA. Maintaining the integrity of the cell's "blueprint" is of vital importance and this is reflected in the numerous mechanisms that exist to repair mistakes and damage in DNA.

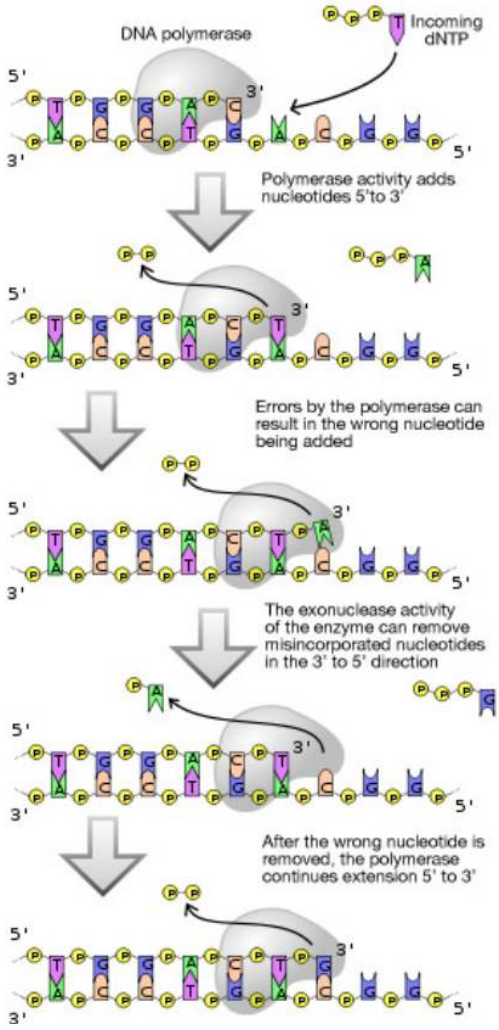


Figure 12.3.1: *Example of DNA correction*

Post-Replicative Mismatch Repair

We earlier discussed proof-reading by DNA polymerase during replication. Does proofreading eliminate all errors made during replication? No, While proof-reading significantly reduces the error rate, not all mistakes are fixed on the fly by DNA polymerases. What mechanisms exist to correct the replication errors that are missed by the proof-reading function of DNA polymerase?

Errors that slip by proofreading during replication can be corrected by a mechanism called **mismatch repair**. While the error rate of DNA replication is about one in 10^7 nucleotides in the absence of mismatch repair, this is further reduced a hundred-fold to one in 10^9 nucleotides when mismatch repair is functional.

What are the tasks that a mismatch repair system faces? It must:

- Scan newly made DNA to see if there are any mispaired bases (e.g., a G paired to a T)
- Identify and cut out the region of the mismatch.
- Correctly fill in the gap created by the excision of the mismatch region.

Importantly, the mismatch repair system must have a means to distinguish the newly made DNA strand from the template strand, if replication errors are to be fixed correctly. In other words, when the mismatch repair system encounters an A-G mispair, for example, it must know whether the A should be removed and replaced with a C or if the G should be removed and replaced with a T.

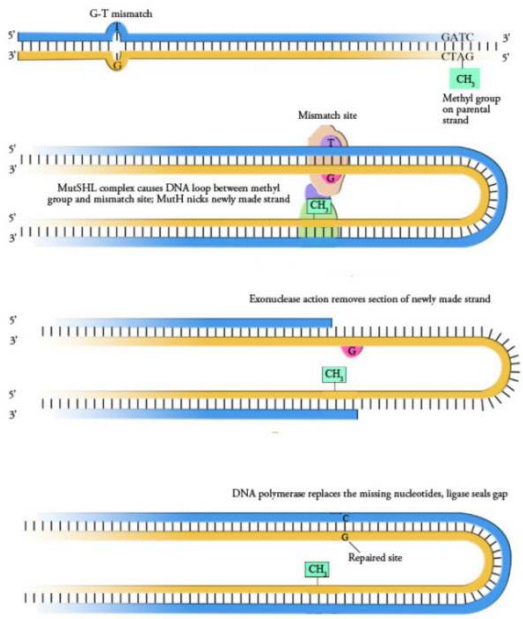


Figure 12.3.2: *Mismatch repair*

Mismatch repair has been well studied in bacteria, and the proteins involved have been identified. Eukaryotes have a mismatch repair system that repairs not only single base mismatches but also insertions and deletions. In bacteria, mismatch repair proteins are encoded by a group of genes collectively known as the **mut genes.** Some of the most important components of the mismatch repair machinery are the proteins **MutS, L** and **H**. ***MutS acts to recognize the mismatch, while MutL and MutH are recruited to the mismatch site by the binding of Mut S, to help cut out the region containing the mismatch. A DNA polymerase and ligase fill in the gap and join the ends, respectively.***

But how does the mismatch repair system distinguish between the original and the new strands of DNA? In bacteria, the existence of a system that methylates the DNA at GATC sequences is the solution to this problem. *E.coli* has an enzyme that adds methyl groups on the to adenines in GATC sequences. Newly replicated DNA lacks this methylation and thus, can be distinguished from the template strand, which is methylated. In Figure 12.3.2, the template strand shown in yellow is methylated at GATC sequences. The mismatch repair proteins selectively replace the strand lacking methylation, shown in blue in the figure, thus ensuring that it is mistakes in the newly made strand that are removed and replaced. Because methylation is the criterion that enables the mismatch repair system to choose the strand that is repaired, the bacterial mismatch repair system is described as being methyl-directed.

Eukaryotic cells do not use this mechanism to distinguish the new strand from the template, and it is not yet understood how the mismatch repair system in eukaryotes "knows" which strand to repair.

Systems to Repair Damage to DNA

In the preceding section we discussed mistakes made when DNA is copied, where the wrong base is inserted during synthesis of the new strand. But even DNA that is not being replicated can get damaged or mutated. These sorts of damage are not associated with DNA replication, rather they can occur at any time.

What causes damage to DNA? Some major causes of DNA damage are:

- Radiation (e.g., UV rays in sunlight, in tanning booths)
- Exposure to damaging chemicals (such as benzopyrene in car exhaust and cigarette smoke)

- Chemical reactions within the cell (such as the deamination of cytosine to give uracil).

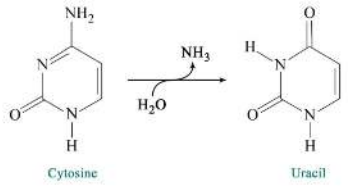


Figure 12.3.3: *Deamination of cytosine*

This means the DNA in your cells is vulnerable to damage simply from normal sorts of actions, such as walking outdoors, being in traffic, or from the chemical transformations occurring in every cell as part of its everyday activities. (Naturally, the damage is much worse in situations where exposure to radiation or damaging chemicals is greater, such as when people repeatedly use tanning beds or smoke).

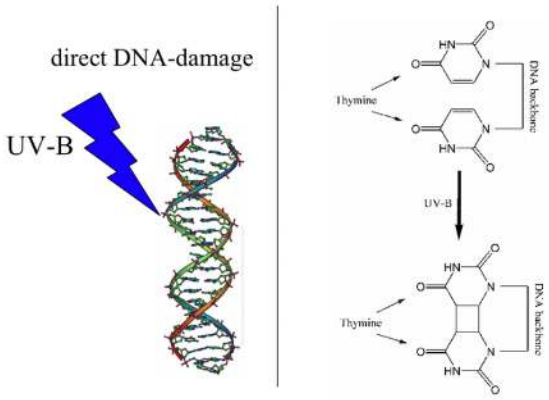


Figure 12.3.4: *Thymine dimer formation*

What kinds of damage do these agents cause? Radiation can cause different kinds of damage to DNA. Sometimes, as with much of the damage done by UV rays, two adjacent pyrimidine bases in the DNA will be cross-linked to form pyrimidine dimers (note that we are talking about two neighboring pyrimidine bases on the same strand of DNA). This is illustrated in the figure on the previous page where two adjacent thymines on a single DNA strand are cross-linked to form a thymine dimer. Radiation can also cause breaks in the DNA backbone.

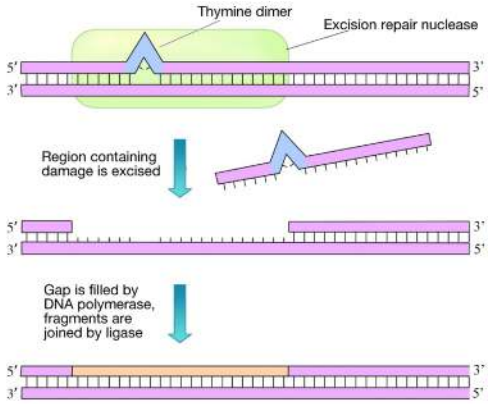


Figure 12.3.5: *Thymine dimer removal*

Chemicals like benzopyrene can attach themselves to bases, forming bulky DNA adducts in which large chemical groups are linked to bases in the DNA. The formation of chemical adducts can physically distort the DNA helix, making it hard for DNA and RNA polymerase to copy those regions of DNA.

Chemical reactions occurring within cells can cause cytosine in DNA to be deaminated to uracil, as shown in Figure 12.3.3.

Other sorts of damage in this category include the formation of oxidized bases like 8-oxo-guanine. These do not actually change the physical structure of the DNA helix, but they can cause problems because uracil and 8-oxo-guanine pair with different bases

than the original cytosine or guanine, leading to mutations on the next round of replication.

How do cells repair such damage? Cells have several ways to remove the sorts of damage described above, with excision repair being a common strategy. Excision repair is a general term for the cutting out and re-synthesis of the damaged region of the DNA. There are a couple of varieties of excision repair:

Nucleotide Excision Repair (NER)

This system fixes damage by chemicals as well as UV damage. As shown in the figure on the previous page, in nucleotide excision repair, the damage is recognized and a cut is made on either side of the damaged region by an enzyme called an excinuclease (shown in green). A short portion of the DNA strand containing the damage is then removed and a DNA polymerase fills in the gap with the appropriate nucleotides. The newly made DNA is joined to the rest of the DNA backbone by the enzyme DNA ligase. In E. coli, NER is carried out by a group of proteins encoded by the **uvrABC** genes. As you can see, NER is similar, in principle, to mismatch repair. However, in NER, the distortion of the helix, caused by the DNA damage, clearly indicates which strand of the DNA needs to be removed and replaced.

Base Excision Repair (BER)

BER deals with situations like the deamination of **cytosine to uraci**l. As noted earlier, cytosines in DNA sometimes undergo deamination to form the base uracil.

Because cytosines pair with guanines and uracils pair with adenine, the conversion of cytosine to uracil in the DNA would lead to the insertion of an A in the newly replicated strand instead of the G that should have gone in across from a C. To prevent this from happening, uracils are removed from DNA by base excision repair.

In base excision repair, a single base is first removed from the DNA, followed by removal of a region of the DNA surrounding the missing base. The gap is then repaired.

The removal of uracil from DNA is accomplished by the enzyme **uracil DNA glycosylase**, which breaks the bond between the uracil and the sugar in the nucleotide.

The removal of the uracil base creates a gap called an **apyrimidinic site (AP site)**. The presence of the AP site triggers the activity of an AP endonuclease that cuts the DNA backbone.

A short region of the DNA surrounding the site of the original uracil is then removed and replaced.

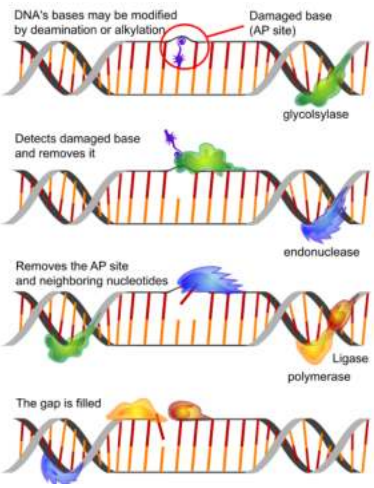


Figure 12.3.6: *Base excision repair*

Contributors

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

12.3: DNA Repair is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

12.4: RNA

RNA is typically single stranded and is made of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and a phosphate group. The subtle structural difference between the sugars gives DNA added stability, making DNA more suitable for storage of genetic information, whereas the relative instability of RNA makes it more suitable for its more short-term functions.

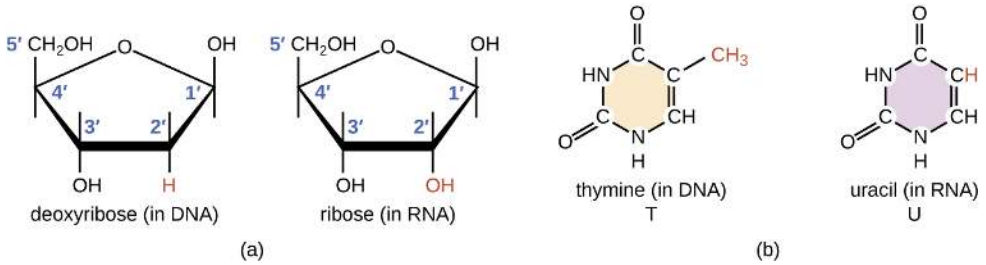


Figure 12.4.1: (a) Ribonucleotides contain the pentose sugar ribose instead of the deoxyribose found in deoxyribonucleotides. (b) RNA contains the pyrimidine uracil in place of thymine found in DNA.

The RNA-specific pyrimidine uracil forms a complementary base pair with adenine and is used instead of the thymine used in DNA. Even though RNA is single stranded, most types of RNA molecules show extensive intramolecular base pairing between complementary sequences within the RNA strand, creating a predictable three-dimensional (3D) structure essential for their function (Figure 12.3.2).

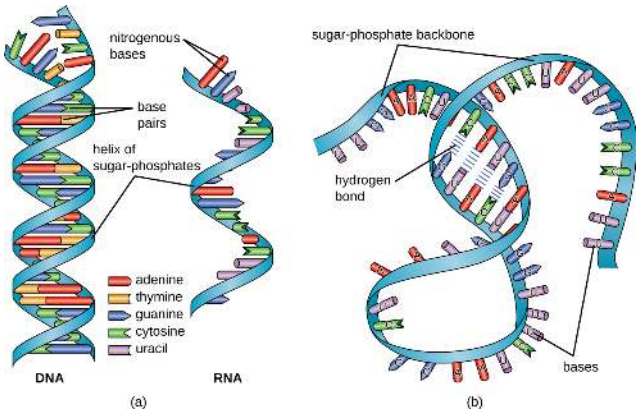


Figure 12.4.2: (a) DNA is typically double stranded, whereas RNA is typically single stranded. (b) Although it is single stranded, RNA can fold upon itself, with the folds stabilized by short areas of complementary base pairing within the molecule, forming a three-dimensional structure.

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

12.4.1 Types of RNA

- There are three main types of RNA, all involved in protein synthesis.
- Messenger RNA (**mRNA**) serves as the intermediary between DNA and the synthesis of protein products during translation.
- Ribosomal RNA (**rRNA**) is a type of stable RNA that is a major constituent of ribosomes. It ensures the proper alignment of the mRNA and the ribosomes during protein synthesis and catalyzes the formation of the peptide bonds between two aligned amino acids during protein synthesis.
- Transfer RNA (**tRNA**) is a small type of stable RNA that carries an amino acid to the corresponding site of protein synthesis in the ribosome. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to be inserted in the polypeptide chain being synthesized.
- Although RNA is not used for long-term genetic information in cells, many viruses do use RNA as their genetic material.

Messenger RNA

Messenger RNA (mRNA) is synthesized from a gene segment of DNA which ultimately contains the information on the primary sequence of amino acids in a protein to be synthesized. The genetic code as translated is for m-RNA not DNA. The **messenger RNA carries the code** into the cytoplasm where protein synthesis occurs.

Ribosomal RNA

In the cytoplasm, **ribsomal RNA (rRNA)** and protein combine to form a nucleoprotein called a ribosome. **The ribosome serves as the site and carries the enzymes necessary for protein synthesis.** The ribosome is made from two sub units, 50S and 30S. There are about equal parts rRNA and protein. The far left graphic shows the complete ribosome with three tRNA attached.

The ribosome attaches itself to m-RNA and provides the stabilizing structure to hold all substances in position as the protein is synthesized. Several ribosomes may be attached to a single RNA at any time. In upper right corner is the 30S sub unit with mRNA and tRNA attached.

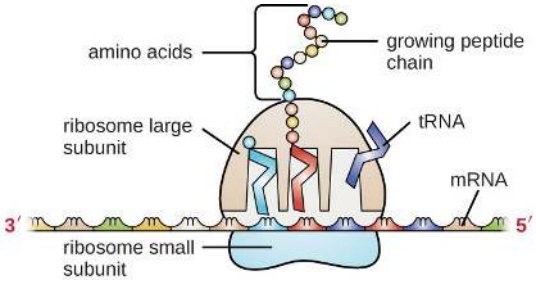


Figure 12.4.1.1: A generalized illustration of how mRNA and tRNA are used in protein synthesis within a cell.

Transfer RNA

Transfer RNA (tRNA) contains about 75 nucleotides, three of which are called anticodons, and one amino acid. **The tRNA reads the code and carries the amino acid to be incorporated into the developing protein.**

There are at least 20 different tRNA's - one for each amino acid. The basic structure of a tRNA is shown in the left graphic. Part of the tRNA doubles back upon itself to form several double helical sections. On one end, the amino acid, phenylalanine, is attached. On the opposite end, a specific base triplet, called the **anticodon**, is used to actually "read" the codons on the mRNA.

The tRNA "reads" the mRNA codon by using its own anticodon. The actual "reading" is done by matching the base pairs through hydrogen bonding following the base pairing principle. Each codon is "read" by various tRNA's until the appropriate match of the anticodon with the codon occurs. In this example, the tRNA anticodon (AAG) reads the codon (UUC) on the mRNA. The UUC codon codes for phenylalanine which is attached to the tRNA. Remember that the codons read from the mRNA make up the genetic code as read by humans.

References

1. 30S Ribosome: Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., Noller, H. F. (2001) "Crystal structure of the ribosome at 5.5 Å resolution" Science 292:883.
2. Sussman, J.L. et al. (1978) "Crystal Structure of Yeast Phenylalanine Transfer RNA I. Crystallographic Refinement" J. Mol. Biol. 123 607.

Contributors

Charles Ophardt (Professor Emeritus, Elmhurst College); Virtual Chembook

12.4.1 Types of RNA is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

12.4.2. RNA - Transcription

The copying of DNA to mRNA is relatively straightforward, with one nucleotide being added to the mRNA strand for every complementary nucleotide read in the DNA strand. The translation to protein is more complex because groups of three mRNA nucleotides correspond to one amino acid of the protein sequence. However, as we shall see in the next module, the translation to protein is still systematic, such that nucleotides 1 to 3 correspond to amino acid 1, nucleotides 4 to 6 correspond to amino acid 2, and so on.

Transcription: from DNA to mRNA

Both prokaryotes and eukaryotes perform fundamentally the same process of transcription, with the important difference of the membrane-bound nucleus in eukaryotes. With the genes bound in the nucleus, transcription occurs in the nucleus of the cell and the mRNA transcript must be transported to the cytoplasm. The prokaryotes, which include bacteria and archaea, lack membrane-bound nuclei and other organelles, and transcription occurs in the cytoplasm of the cell. In both prokaryotes and eukaryotes, transcription occurs in three main stages: initiation, elongation, and termination.

Initiation

Transcription requires the DNA double helix to partially unwind in the region of mRNA synthesis. The region of unwinding is called a **transcription bubble**. The DNA sequence onto which the proteins and enzymes involved in transcription bind to initiate the process is called a promoter. In most cases, promoters exist upstream of the genes they regulate. The specific sequence of a promoter is very important because it determines whether the corresponding gene is transcribed all of the time, some of the time, or hardly at all (Figure 12.3.2.1).

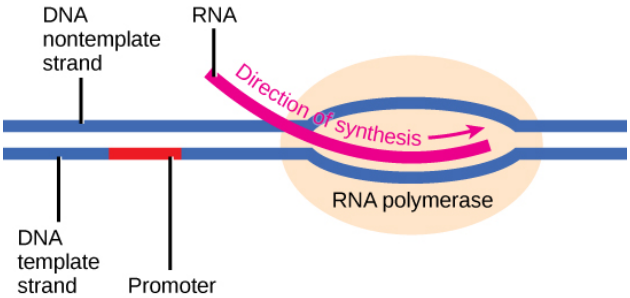


Figure 12.4.2.1: The initiation of transcription begins when DNA is unwound, forming a transcription bubble. Enzymes and other proteins involved in transcription bind at the promoter.

Elongation

Transcription always proceeds from one of the two DNA strands, which is called the template strand. The mRNA product is complementary to the template strand and is almost identical to the other DNA strand, called the nontemplate strand, with the exception that RNA contains a uracil (U) in place of the thymine (T) found in DNA. During elongation, an enzyme called **RNA polymerase** proceeds along the DNA template adding nucleotides by base pairing with the DNA template in a manner similar to DNA replication, with the difference that an RNA strand is being synthesized that does not remain bound to the DNA template. As elongation proceeds, the DNA is continuously unwound ahead of the core enzyme and rewound behind it (Figure 12.3.2.2).

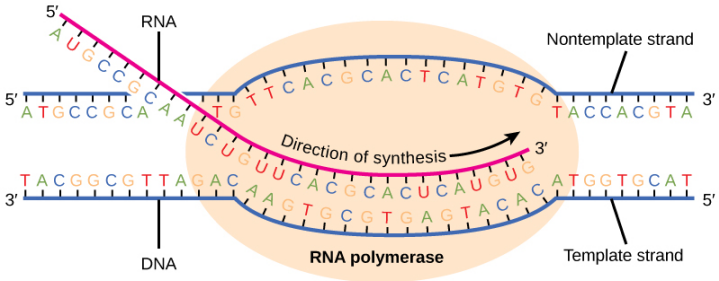


Figure 12.4.2.2: During elongation, RNA polymerase tracks along the DNA template, synthesizes mRNA in the 5' to 3' direction, and unwinds then rewinds the DNA as it is read.

Termination

Once a gene is transcribed, the prokaryotic polymerase needs to be instructed to dissociate from the DNA template and liberate the newly made mRNA. Depending on the gene being transcribed, there are two kinds of termination signals, but both involve repeated nucleotide sequences in the DNA template that result in RNA polymerase stalling, leaving the DNA template, and freeing the mRNA transcript.

On termination, the process of transcription is complete. In a prokaryotic cell, by the time termination occurs, the transcript would already have been used to partially synthesize numerous copies of the encoded protein because these processes can occur concurrently using multiple ribosomes (polyribosomes) (Figure 9.3.4). In contrast, the presence of a nucleus in eukaryotic cells precludes simultaneous transcription and translation.

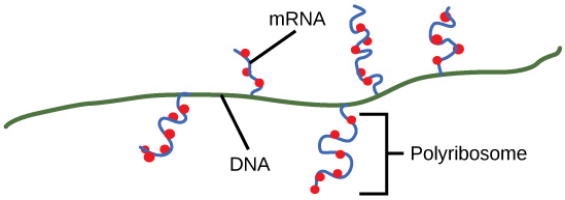


Figure 12.4.2.3: **Multiple polymerases can transcribe a single bacterial gene** while numerous ribosomes concurrently translate the mRNA transcripts into polypeptides. In this way, a specific protein can rapidly reach a high concentration in the bacterial cell.

Eukaryotic RNA Processing

The newly transcribed eukaryotic mRNAs must undergo several processing steps before they can be transferred from the nucleus to the cytoplasm and translated into a protein. The additional steps involved in eukaryotic mRNA maturation create a molecule that is much more stable than a prokaryotic mRNA. For example, eukaryotic mRNAs last for several hours, whereas the typical prokaryotic mRNA lasts no more than five seconds.

The mRNA transcript is first coated in RNA-stabilizing proteins to prevent it from degrading while it is processed and exported out of the nucleus. This occurs while the pre-mRNA still is being synthesized by adding a special nucleotide "**cap**" to the 5' end of the growing transcript. In addition to preventing degradation, factors involved in protein synthesis recognize the cap to help initiate translation by ribosomes.

Once elongation is complete, an enzyme then adds a string of approximately 200 adenine residues to the 3' end, called the **poly-A tai**l. This modification further protects the pre-mRNA from degradation and signals to cellular factors that the transcript needs to be exported to the cytoplasm.

Eukaryotic genes are composed of protein-coding sequences called exons (*ex*-on signifies that they are *ex*pressed) and *int*ervening sequences called introns (*int*-ron denotes their *int*ervening role). Introns are removed from the pre-mRNA during processing. Intron sequences in mRNA do not encode functional proteins. It is essential that all of a pre-mRNA's introns be completely and precisely removed before protein synthesis so that the exons join together to code for the correct amino acids. If the process errs by even a single nucleotide, the sequence of the rejoined exons would be shifted, and the resulting protein would be nonfunctional. The process of removing introns and reconnecting exons is called splicing (Figure 12.3.2.4). Introns are removed and degraded while the pre-mRNA is still in the nucleus.

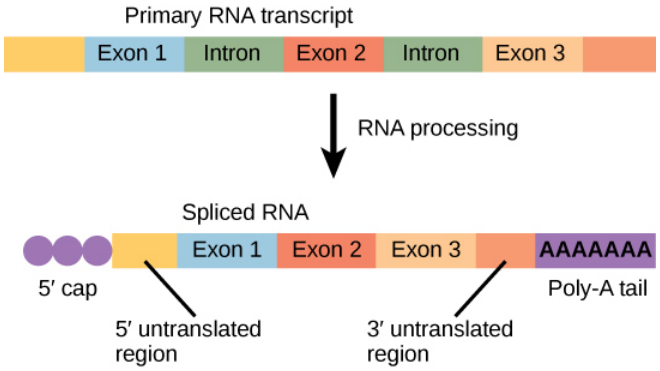


Figure 12.4.2.4: Eukaryotic mRNA contains introns that must be spliced out. A 5' cap and 3' tail are also added.

Summary

In prokaryotes, mRNA synthesis is initiated at a promoter sequence on the DNA template. Elongation synthesizes new mRNA. Termination liberates the mRNA and occurs by mechanisms that stall the RNA polymerase and cause it to fall off the DNA template. On the other hand, newly transcribed eukaryotic mRNAs are modified with a cap and a poly-A tail. These structures protect the mature mRNA from degradation and help export it from the nucleus. Eukaryotic mRNAs also undergo splicing, in which introns are removed and exons are reconnected with single-nucleotide accuracy. The matured mRNAs are exported from the nucleus to the cytoplasm.

Contributors

Charles Ophardt, Professor Emeritus, Elmhurst College; Virtual Chembook

Template:ContribOpenStaxConcepts

12.4.2. RNA - Transcription is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

12.4.3. Regulation of Transcription

We know that promoters indicate where transcription begins, but what determines that a given gene will be transcribed? In addition to the promoter sequences required for transcription initiation, genes have additional regulatory sequences (sequences of DNA on the same DNA molecule as the gene) that control when a gene is transcribed. **Regulatory sequences** are bound tightly and specifically by transcriptional regulators, proteins that can recognize DNA sequences and bind to them. The binding of such proteins to the DNA can regulate transcription by preventing or increasing transcription from a particular promoter.

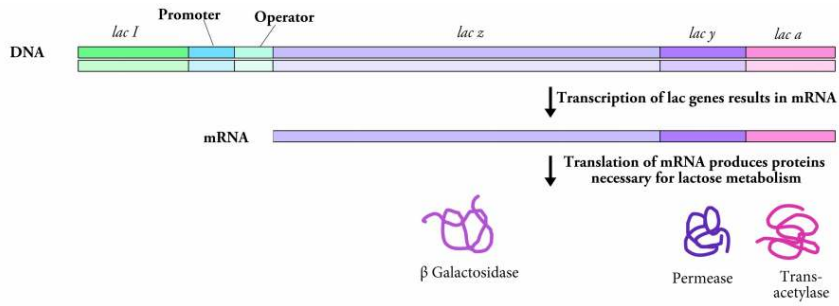


Figure 12.4.3.1: *The genes lac z, lac y, and lac a are all under the control of a single promotor in the lac operon*

Regulation in Prokaryotes

In bacteria, genes are often clustered in groups, such that genes that need to be expressed at the same time are next to each other and all of them are controlled as a single unit by the same promoter. The lac operon, shown in Figure 12.3.3.2. the three genes of the lac operon, lac z, lac y and lac a are controlled by a single promoter.

Bacterial cells generally prefer to use glucose for their energy needs, but if glucose is unavailable, and lactose is present, the bacteria will take up lactose and break it down for energy. Since the proteins for taking up and breaking down lactose are only needed when glucose is absent and lactose is available, the bacterial cells need a way to express the genes of the lac operon only under those conditions. At times when lactose is absent, the cells do not need to express these genes.

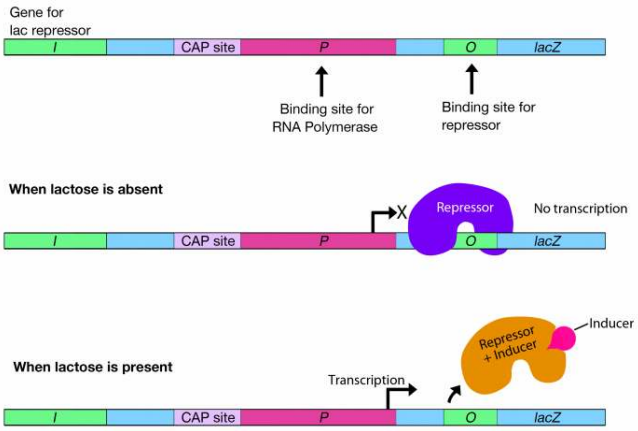


Figure 12.4.3.2: *Lac operon regulation*

How do bacteria achieve this? Transcription of the lac cluster of genes is primarily controlled by a repressor protein that binds to a region of the DNA just downstream of the -10 sequence of the lac promoter. **Recall that the promoter is where the RNA polymerase must bind to begin transcription**. The place where the repressor is bound is called the operator (labeled O in the figure). When the repressor is bound at this position, it physically blocks the RNA polymerase from transcribing the genes. In order for transcription to occur, the repressor must be removed from the operator to clear the path for RNA polymerase. How is the repressor removed?

When the sugar lactose is present, it binds to the repressor, changing its conformation so that it no longer binds to the operator. When the repressor is no longer bound at the operator, the "road-block" in front of the RNA polymerase is removed, permitting the transcription of the genes of the lac operon.

Because the binding of the lactose induces the expression of the genes in the lac operon, lactose is called an inducer. (Technically, the inducer is allolactose, a molecule made from lactose by the cell, but the principle is the same.)

What makes this an especially effective control system is that the genes of the lac operon encode proteins that break down lactose. Turning on these genes requires lactose to be present. Once the lactose is broken down, the repressor binds to the operator once more and the lac genes are no longer expressed. This allows the genes to be expressed only when they are needed.

But how do glucose levels affect the expression of the lac genes? We noted earlier that if glucose was present, lactose would not be used. A second level of control is exerted by a protein called CAP that binds to a site adjacent to the promoter and recruits RNA polymerase to bind the lac promoter. When glucose is depleted, there is an increase in levels of cAMP which binds to CAP. The CAP cAMP complex then binds the CAP site, as shown in Figure 5.4.3. The combination of CAP binding and the lac repressor dissociating from the operator when lactose levels are high ensures high levels of transcription of the lac operon just when it is most needed. The CAP protein binding may be thought of as a green light for the RNA polymerase, while the removal of repressor is like the lifting of a barricade in front of it. When both conditions are met, the RNA polymerase transcribes the downstream genes.

Regulation in Eukaryotes

Transcription in eukaryotes is also regulated by the binding of proteins to specific DNA sequences, but with some differences from the simple schemes outlined above. For most eukaryotic genes, general transcription factors and RNA polymerase (i.e., the basal transcription complex) are necessary, but not sufficient, for high levels of transcription.

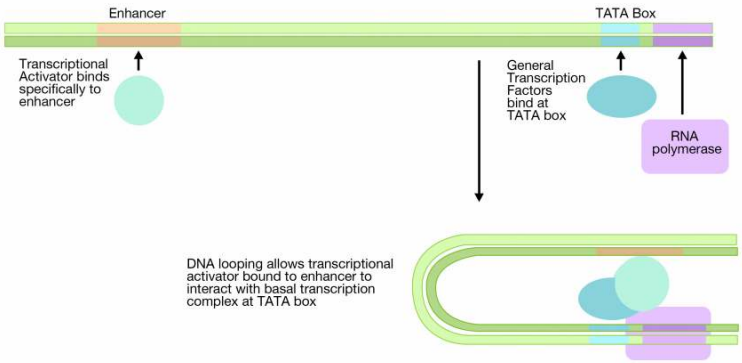


Figure 12.4.3.3: *Enhancer Mechanism*

In eukaryotes, additional regulatory sequences called **enhancers** and the proteins that bind to the enhancers are needed to achieve high levels of transcription. Enhancers are DNA sequences that regulate the transcription of genes. Unlike prokaryotic regulatory sequences, enhancers don't need to be next to the gene they control. Often they are many kilobases away on the DNA. As the name suggests, enhancers can enhance (increase) transcription of a particular gene.

How can a DNA sequence far from the gene being transcribed affect the level of its transcription?

Enhancers work by binding proteins (**transcriptional activators**) that can, in turn, interact with the proteins bound at the promoter. The enhancer region of the DNA, with its associated transcriptional activator(s) can come in contact with the basal transcription complex that is bound at a distant TATA box by looping of the DNA. This allows the protein bound at the enhancer to make contact with the proteins in the basal transcription complex.

One way that the transcriptional activator bound to the enhancer increases the transcription from a distant promoter is that it increases the frequency and efficiency with which the basal transcription complex is formed at the promoter.

Another mechanism by which proteins bound at the enhancer can affect transcription is by recruiting to the promoter other proteins that can modify the structure of the chromatin in that region. As we noted earlier, in eukaryotes, DNA is packaged with proteins to form chromatin. When the DNA is tightly associated with these proteins, it is difficult to access for transcription. So proteins that can make the DNA more accessible to the transcription machinery can also play a role in the extent to which transcription occurs.

In addition to enhancers, there are also negative regulatory sequences called **silencers**. Such regulatory sequences bind to transcriptional repressor proteins. Transcriptional activators and repressors are modular proteins- they have a part that binds DNA

and a part that activates or represses transcription by interacting with the basal transcription complex.

Contributors

Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

12.4.3. Regulation of Transcription is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

12.5: The Genetic Code

Introduction

You can think of the sequences of bases in the coding strand of DNA or in messenger RNA as coded instructions for building protein chains out of amino acids. There are 20 amino acids used in making proteins, but only four different bases to be used to code for them.

Obviously one base can't code for one amino acid. That would leave 16 amino acids with no codes.

If you took two bases to code for each amino acid, that would still only give you 16 possible codes (TT, TC, TA, TG, CT, CC, CA and so on) - still not enough.

However, if you took three bases per amino acid, that gives you 64 codes (TTT, TTC, TTA, TTG, TCT, TCC and so on). That's enough to code for everything with lots to spare. You will find a full table of these below.

A three base sequence in DNA or RNA is known as a codon.

The code in DNA

The codes in the coding strand of DNA and in messenger RNA aren't, of course, identical, because in RNA the base uracil (U) is used instead of thymine (T).

The table shows how the various combinations of three bases in the coding strand of DNA are used to code for individual amino acids - shown by their three letter abbreviation.

first base in codon	second base in codon: T	C	A	G	third base in codon
T	TTT Phe	TCT Ser	TAT Tyr	TGT Cys	T
	TTC Phe	TCC Ser	TAC Tyr	TGC Cys	C
	TTA Leu	TCA Ser	TAA stop	TGA stop	A
	TTG Leu	TCG Ser	TAG stop	TGG Trp	G
C	CTT Leu	CCT Pro	CAT His	CGT Arg	T
	CTC Leu	CCC Pro	CAC His	CGC Arg	C
	CTA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CTG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	ATT Ile	ACT Thr	AAT Asn	AGT Ser	T
	ATC Ile	ACC Thr	AAC Asn	AGC Ser	C
	ATA Ile	ACA Thr	AAA Lys	AGA Arg	A
	ATG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GTT Val	GCT Ala	GAT Asp	GGT Gly	T
	GTC Val	GCC Ala	GAC Asp	GGC Gly	C
	GTA Val	GCA Ala	GAA Glu	GGA Gly	A
	GTG Val	GCG Ala	GAG Glu	GGG Gly	G

The table is arranged in such a way that it is easy to find any particular combination you want. It is fairly obvious how it works and, in any case, it doesn't take very long just to scan through the table to find what you want.

The colors are to stress the fact that most of the amino acids have more than one code. Look, for example, at leucine in the first column. There are six different codons all of which will eventually produce a leucine (Leu) in the protein chain. There are also six for serine (Ser).

In fact there are only two amino acids which have only one sequence of bases to code for them - methionine (Met) and tryptophan (Trp). You have probably noticed that three codons don't have an amino acid written beside them, but say "stop" instead. For obvious reasons these are known as stop codons. We'll leave talking about those until we have looked at the way the code works in messenger RNA.

The code in messenger RNA

You will remember that when DNA is transcribed into messenger RNA, the sequence of bases remains exactly the same, except that each thymine (T) is replaced by uracil (U). That gives you the table:

first base in codon	second base in codon: U	C	A	G	third base in codon
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA stop	UGA stop	A
	UUG Leu	UCG Ser	UAG stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

In many ways, this is the more useful table. Messenger RNA is directly involved in the production of the protein chains (see the next page in this sequence). The DNA coding chain is one stage removed from this because it must first be transcribed into a messenger RNA chain.

Start and stop codons

The **stop codons** in the RNA table (UAA, UAG and UGA) serve as a signal that the end of the chain has been reached during protein synthesis.

The codon that marks the start of a protein chain is AUG. If you check the table, that's the amino acid, methionine (Met). That ought to mean that every protein chain must start with methionine. That's not quite true because in some cases the methionine can get chopped off the chain after synthesis is complete.

Contributors

Jim Clark (Chemguide.co.uk)

This page titled 12.5: The Genetic Code is shared under a CC BY-NC 4.0 license and was authored, remixed, and/or curated by Jim Clark.

12.6: Translation

The Mechanism of Protein Synthesis (Translation)

Just as with mRNA synthesis, protein synthesis can be divided into three phases: initiation, elongation, and termination. The process of translation is similar in bacteria, archaea and eukaryotes.

Translation Initiation

In general, protein synthesis begins with the formation of an initiation complex. The small subunit binds to a site upstream (on the 5' side) of the start of the mRNA. It proceeds to scan the mRNA in the **5'-->3'** direction until it encounters the **START codon (AUG)**. The large subunit attaches and the initiator tRNA, which carries methionine (Met), binds to the P site on the ribosome.

The small ribosomal subunit will bind to the mRNA at the **ribosomal binding site**. Soon after, the methionine-tRNA will bind to the AUG start codon (through complementary binding with its anticodon). This complex is then joined by large ribosomal subunit. This initiation complex then recruits the second tRNA and thus translation begins.The small subunit It proceeds to scan the mRNA in the 5'-->3' direction until it encounters the START codon (AUG). The large subunit attaches and the initiator tRNA, which carries methionine (Met), binds to the P site on the ribosome.

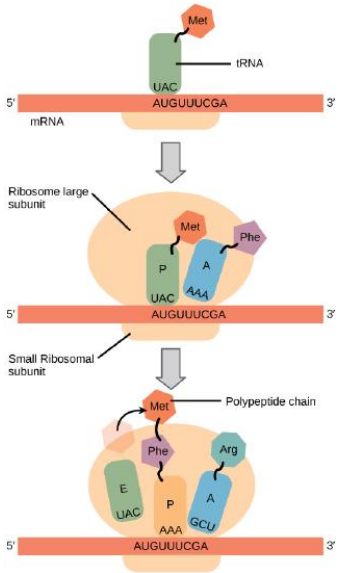


Figure 12.6.1: *Translation begins with when tRNA anticodon recognizes a codon on the mRNA. The large ribosomal subunits joins the small subunits, and a second tRNA is recruited. As the mRNA moves relative to the ribosome, the polypeptide chain ofis formed. Entry of a release factor into the A site terminates translation and the component dissociate.*

Bacterial vs Eukaryotic initiation

Protein synthesis begins at an AUG (met) codon- but proteins may have many methionines, and mRNAs may have many AUGs. How does the ribosome know where to begin?

In prokayrotic mRNA, a sequence upstream of the first AUG codon, called the **Shine-Dalgarno sequence** (AGGAGG), base-pairs with a rRNA molecule within the small subunit of bacterial and archeal ribosomes. This interaction anchors the 30S ribosomal subunit at a precise location on the mRNA template.

Instead of binding at the Shine-Dalgarno sequence, the **eukaryotic initiation complex** (a number of proteins in addition to the small subunit) recognizes the 7-methylguanosine cap at the 5' end of the mRNA. Once at the cap, the initiation complex tracks along the mRNA in the 5' to 3' direction, searching for the AUG start codon. Many eukaryotic mRNAs are translated from the first AUG, but this is not always the case. The nucleotides around the AUG affect the probability that it will be chosen as the start codon, and the consensus sequence varies between species. The "helper" proteins of the initiation complex fall off once the large subunit is loaded.

Note that in both cases the selection of an AUG establishes the reading frame (one of a possible 3) for the entire protein. A very important difference between these modes of start-site selection is that a single prokaryotic transcript can potentially encode several sequential proteins, as the ribosome can scan the entire length of the message for Shine-Dagarno sequences. Often the several proteins involved in a single process (subunits of a holoenzyme. or sequential steps in a metabolic pathway) are encoded on a single message. In contrast, in eukaryotic nuclear genes, each transcript only encodes a single protein (as always, there are exceptions).

Translation Elongation

During translation elongation, the mRNA template provides specificity. As the ribosome moves along the mRNA, each mRNA codon comes into 'view', and specific binding with the corresponding charged tRNA anticodon is ensured. If mRNA were not present in the elongation complex, the ribosome would bind tRNAs nonspecifically.

The large ribosomal subunit consists of three compartments: the A site binds incoming charged tRNAs (tRNAs with their attached specific amino acids), the P site binds charged tRNAs carrying amino acids that have formed bonds with the growing polypeptide chain but have not yet dissociated from their corresponding tRNA, and the E site which releases dissociated tRNAs so they can be recharged with another free amino acid.

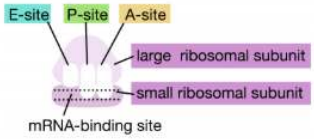


Figure 12.6.2: Ribosomal subunit

A tRNA bound to its amino acid (known as an aminoacyl-tRNA) that is able to base pair with the next codon on the mRNA arrives at the A site. The preceding amino acid (Met at the start of translation) is covalently linked to the incoming amino acid with a peptide bond. The initiator tRNA moves to the E site and the ribosome moves one codon downstream. This shifts the more most recent tRNA from the A site to the P site, opening up the A site for the arrival of a new aminoacyl-tRNA. Polypeptide synthesis repeats, the tRNA residing in the E site is released from the complex, the tRNAs in the P site and A site shift over and the next amino acid is added to the growing polypeptide chain. This cycle repeats until a stop codon is reached.

Ribosomal steps are induced by conformational changes that advance the ribosome by three bases in the 3' direction. The energy for each step of the ribosome is donated by an elongation factor that hydrolyzes GTP. Peptide bonds form between the amino group of the amino acid attached to the A-site tRNA and the carboxyl group of the amino acid attached to the P-site tRNA. The formation of each peptide bond is catalyzed by **peptidyl transferase**, a catalytic RNA (surprise! not a protein) that is integrated into the 50S ribosomal subunit. The energy for each peptide bond formation is derived from GTP hydrolysis, which is catalyzed by a separate elongation factor. The amino acid bound to the P-site tRNA is linked to the growing polypeptide chain. As the ribosome steps across the mRNA, the former P-site tRNA enters the E site, detaches from the amino acid, and is expelled (it will be recharged by tRNA synthetase later). The ribosome moves along the mRNA, one codon at a time, catalyzing each process that occurs in the three sites. With each step, a charged tRNA enters the complex, the polypeptide becomes one amino acid longer, and an uncharged tRNA departs.

This and subsequent steps in the synthesis of the polypeptide are called the elongation phase of translation. Once the first two amino acids are linked , the first tRNA dissociates, and moves out of the P-site and into the E, or Exit site. The second tRNA then moves into the P-site, vacating the A-site for the tRNA corresponding to the next codon.

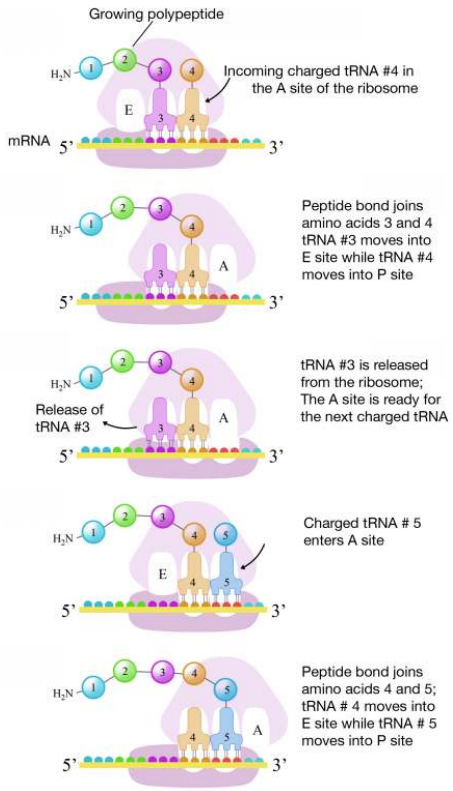


Figure 12.6.3: *Elongation of translation*

Termination

Translation ends when the ribosome reaches a STOP codon (UAA, UAG or UGA). There are no tRNA molecules with anticodons for stop codons, instead protein release factors recognize these codons when they arrive at the A site. Binding of a release protein causes the polypeptide (protein) to be released from the ribosome. The ribosome subunits dissociate (split) from each other and can be reassembled later for another round of protein synthesis.

Post-translational Protein Modification

After translation individual amino acids may be chemically modified. These modifications add chemical variation not present in the genetically encoded amino acids, and new properties that are rooted in the chemistries of the functional groups that are being added. Common modifications include phosphate groups, methyl, acetate, and amide groups. Some proteins, typically targeted to membranes, will be lipidated - a lipid will be added. Other proteins will be glycosylated - a sugar will be added. Another common post-translational modification is cleavage or linking of parts of the protein itself. Signal-peptides may be cleaved, parts may be excised from the middle of the protein, or new covalent linkages may be made between cysteine or other amino acid side chains. Nearly all modifications will be catalyzed by enzymes and all change the functional behavior of the protein.

Summary

An mRNA is used to synthesize proteins by the process of translation. The genetic code is the correspondence between the three-nucleotide mRNA codon and an amino acid. The genetic code is “translated” by the tRNA molecules, which associate a specific codon with a specific amino acid. The genetic code is degenerate because 64 triplet codons in mRNA specify only 20 amino acids and three stop codons. This means that more than one codon corresponds to an amino acid. Almost every species on the planet uses the same genetic code; the "deviant codes" are not radically different, but change the meaning of one or two codons. More impressive exceptions are species that encode 21 or 22 amino acids, rather than the usual 20.

The players in translation include the mRNA template, ribosomes, tRNAs, and various enzymatic factors. The small ribosomal subunit binds to the mRNA template. Translation begins at the initiating AUG on the mRNA (this also establishes the reading frame). The formation of bonds occurs between sequential amino acids specified by the mRNA template according to the genetic

code. The ribosome accepts charged tRNAs, and as it steps along the mRNA, it catalyzes bonding between the new amino acid and the end of the growing polypeptide. The entire mRNA is translated in three-nucleotide “steps” of the ribosome. When a stop codon is encountered, a release factor binds and dissociates the components and frees the new protein.

This page titled 12.6: Translation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern & Indira Rajagopal.

12.7: Exercises

Multiple Choice

1. Which of the following is the name of the three-base sequence in the mRNA that binds to a tRNA molecule?

A. P site
B. codon
C. anticodon
D. CCA binding site

---

2. Which component is the last to join the initiation complex during the initiation of translation?

A. the mRNA molecule
B. the small ribosomal subunit
C. the large ribosomal subunit
D. the initiator tRNA

---

3. During elongation in translation, to which ribosomal site does an incoming charged tRNA molecule bind?

A. A site
B. P site
C. E site
D. B site

---

4. Which of the following is the amino acid that appears at the N-terminus of all newly translated prokaryotic and eukaryotic polypeptides?

A. tryptophan
B. methionine
C. selenocysteine
D. glycine

---

5. When the ribosome reaches a nonsense codon, which of the following occurs?

A. a methionine is incorporated
B. the polypeptide is released
C. a peptide bond forms
D. the A site binds to a charged tRNA

---

6. Write the anticodon on tRNA that would pair with each mRNA codon.

1. 5′-UUU-3′
2. 5′-CAU-3′
3. 5′-AGC-3′
4. 5′-CCG-3

7. Write the codon on mRNA that would pair with each tRNA anticodon.

1. 5′-UUG-3′
2. 5′-GAA-3′

3. 5′-UCC-3′
4. 5′-CAC-3′

8. Which of the following types of RNA codes for a protein?

A. dsRNA
B. mRNA
C. rRNA
D. tRNA

---

9. A nucleic acid is purified from a mixture. The molecules are relatively small, contain uracil, and most are covalently bound to an amino acid. Which of the following was purified?

A. DNA
B. mRNA
C. rRNA
D. tRNA

---

10. Which of the following types of RNA is known for its catalytic abilities?

A. dsRNA
B. mRNA
C. rRNA
D. tRNA

---

11. Ribosomes are composed of rRNA and what other component?

1. protein
2. polypeptides
3. DNA
4. mRNA

12. Which of the following may use RNA as its genome?

1. a bacterium
2. an archaeon
3. a virus
4. a eukaryote

---

13. A nucleotide of DNA may contain _____.

1. ribose, uracil, and a phosphate group
2. deoxyribose, uracil, and a phosphate group
3. deoxyribose, thymine, and a phosphate group
4. ribose, thymine, and a phosphate group

14. The building blocks of nucleic acids are _____.

1. sugars
2. nitrogenous bases
3. peptides
4. nucleotides

15. The given nucleotide could be translated into a peptide of 7 amino acids, please provide the sequence of the peptide using one letter abbreviation. (7 points)

AUCGAGCUAUGCUGAUGACUGUGACAUACUAAAUCUCA

16. Which of the following polymerases do not have proof reading ability?

a. DNA polymerase I

b. DNA polymerase II

c. DNA polymerase III

d. DNA polymerase a (alpha)

e. DNA polymerase d (delta)

17. Which of teh following may detect G:T mismatch in bacterial genome?

a. exonuclease I

b. MutS

c. DNA polymerase III

d. Uracil DNA glycosidase

e. uvrABC excinuclease

18. Which of the following is the primary target of ciprofloxacin?

a. DNA topoisomerase II (DNA gyrase)

b. DNA topoisomerase I

c. Helicase

d. DnaA

e. DnaB

19. Which of the following is not a function of reverse transcriptase?

a. Synthesis of DNA complementary to RNA

b. Digestion of RNA

c. Synthesis of second strand of DNA

d. Incorporation of DNA fragment into host genome.

Free Response

What are the four types of RNA and how do they function?

Short Answer

1. Why does translation terminate when the ribosome reaches a stop codon? What happens?

2. How does the process of translation differ between prokaryotes and eukaryotes?

3. What is meant by the genetic code being nearly universal?

4. Below is an antisense DNA sequence. Translate the mRNA molecule synthesized using the genetic code, recording the resulting amino acid sequence, indicating the N and C termini.

Antisense DNA strand: 3’-T A C T G A C T G A C G A T C-5’

5.What are the roles of mRNA and tRNA in protein synthesis?

6. What is the initiation codon?

7. What are the termination codons and how are they recognized?

8.What are the differences between DNA nucleotides and RNA nucleotides?

9. How is the information stored within the base sequence of DNA used to determine a cell's properties?

10. How do complementary base pairs contribute to intramolecular base pairing within an RNA molecule?

11. If an antisense RNA has the sequence 5′AUUCGAAUGC3′, what is the sequence of the mRNA to which it will bind? Be sure to label the 5′ and 3′ ends of the molecule you draw.

12. Why does double-stranded RNA (dsRNA) stimulate RNA interference?

13. A portion of an mRNA molecule has the sequence 5′-AUGCCACGAGUUGAC-3′. What amino acid sequence does this code for? Help: genetic code table

14. What are the benefits and drawbacks to translating a single mRNA multiple times?

15. What are the processing steps for messenger RNAs?

16. How does the structure of RNA differ from the structure of DNA?

17. Describe how DNA is replicated in eukaryotes (sq)

18. In base excision repair, which enzymes are specific for a particular kind of damage and which are used for all repair by this pathway?

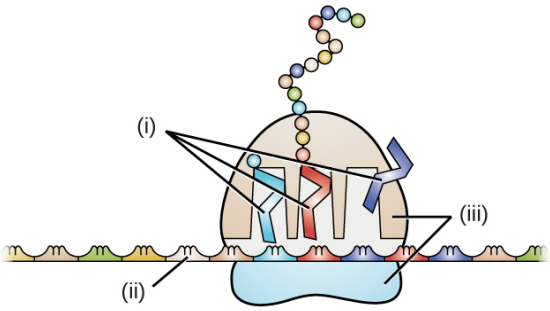
19. Discuss the different types of mutations in DNA.

20. Explain DNA repair mechanisms.

21. What are the three types of RNA and how do they function?

Critical Thinking

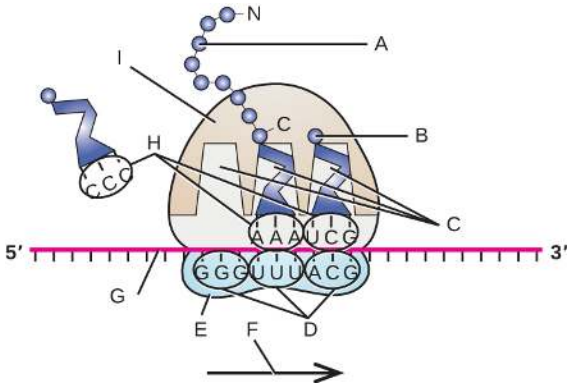
Identify the location of mRNA, rRNA, and tRNA in the figure.



Why does it make sense that tRNA and rRNA molecules are more stable than mRNA molecules?

Critical Thinking

Label the following in the figure: ribosomal E, P, and A sites; mRNA; codons; anticodons; growing polypeptide; incoming amino acid; direction of translocation; small ribosomal unit; large ribosomal unit.



Quiz: Read and translate the codons on mRNA into the appropriate amino acids. G U A C G A A A A	Answer

Quiz: Read and translate the codon on mRNA into the appropriate amino acid.
AGA. What is the anticodon on the appropriate tRNA?

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

13: Integrated chapter (HIV)

Human immunodeficiency virus (HIV)

Human immunodeficiency virus (HIV) is a retrovirus, which is a class of viruses that carry genetic information in RNA. There are two types of HIV, **HIV-1** and **HIV-2**, with HIV-1 being the most predominant, it is commonly called just HIV. Both types of HIV damage a person's body by destroying specific blood cells, called **CD4+ T cells**, which are crucial to helping the body fight diseases in the immune system.

Many people infected with HIV eventually develop **acquired immune deficiency syndrome (AIDS).** This may not occur until many years after the virus first enters the body. The virus infects and destroys helper T cells of the human immune system. This can lead to immune deficiency, which is when the infection with the virus progressively deteriorates the immune system and is considered deficient when it no longer works to help fight infection and disease.

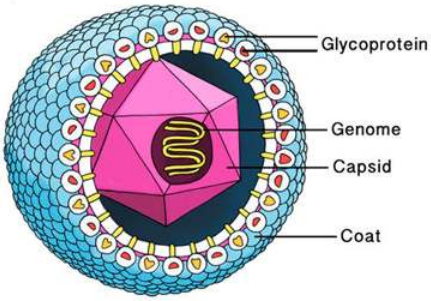


Figure 13.1: *Structure of HIV virion*

HIV Transmission

HIV is considered to be a sexually transmitted infection (STI) because it is the most common mode of transmission. It is transmitted, or spread, through direct contact of **mucous membranes** or body fluids such as blood, semen, or breast milk. It can also be transmitted through an infected mother's blood to her baby during late pregnancy or birth or through breast milk after birth. In the past, HIV was also transmitted through blood transfusions. Because donated blood is now screened for HIV, the virus is no longer transmitted this way. HIV is not spread through saliva, touching or in swimming pools.

Replication cycle

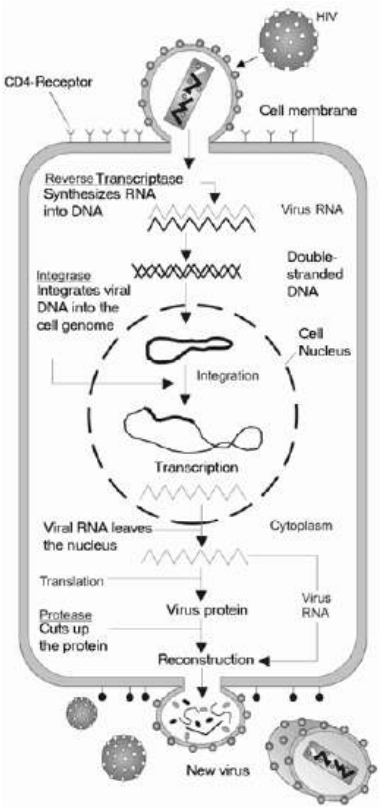


Figure 13.2: *Steps in the HIV Replication Cycle*

First, the viral particle attaches to the CD4 receptor and other associated receptors on the host cell membrane. The viral envelope then fuses with the cell membrane, and the viral capsid moves into the cell.

1. Once the viral capsid enters the cell, **reverse transcriptase** frees the single-stranded RNA from the viral proteins and copies it into a complementary strand of DNA. This process of reverse transcription is error-prone and it is during this step that mutations may occur. Such mutations may cause drug resistance.
2. The reverse transcriptase then makes a complementary DNA strand to form a double-stranded viral DNA (*vDNA*).
3. The *vDNA* is then moved into the cell nucleus. The integration of the *viral DNA* into the host cell's genome is carried out by another viral enzyme called **integrase**. This integrated viral DNA may then lie dormant, during the latent stage of the HIV infection. Clinical latency for HIV can vary between two weeks and 20 years.
4. To actively produce viruses, certain cellular transcription factors need to be present. These transcription factors are plentiful in activated T cells. This means that those cells most likely to be killed by HIV are those currently fighting infection. The virus DNA is transcribed to mRNA which then leads to new virus protein and genome production.
5. Viral particles are assembled inside the cell and then exit the cell by budding. The virus gets its viral envelope from the cell’s plasma membrane. The cycle begins again when the new particles infect another cell.

Prevention of HIV Transmission

There is currently no approved vaccine for HIV or AIDS, although vaccine trials are ongoing. Instead, prevention of HIV transmission depends on adopting safe behaviors and/or the administration of antiretroviral drugs. HIV transmission through intravenous drug use can be reduced through harm-reduction strategies such as needle-exchange programs or the substitution of prescription drugs for illegal drugs. In cases of unanticipated exposure to infected blood, such as a needle-stick injury or sexual assault by an HIV-positive perpetrator, risk of HIV infection can be substantially reduced by the administration of **antiretroviral medications** within two or three days of the incident.

13: Integrated chapter (HIV) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

13.1: Envelope glycoprotein GP120

In order to infect a human cell, an envelope glycoprotein found on the surface of HIV called **Gp120** must adsorbs to both a $CD4^+$ molecule and then a chemokine receptor found on the surface of only certain types of human cells.

Human cells possessing $CD4^+$molecules include:

- T4-helper lymphocytes (also called T4-cells and $CD4^+$ cells)
- monocytes
- macrophages
- dendritic cells

The interaction of viral envelope glycoprotein gp120 with host CD4+ and chemokine receptor molecules is electrostatic, and the interaction induces structural changes in gp120 and gp41 which initiate a fusion of the viral and cellular membranes.

This allows for a more stable two-pronged attachment, which facilates the N-terminal fusion peptide gp41 to penetrate the cell membrane. Repeat sequences in gp41, known as HR1 and HR2, then interact, causing the collapse of the extracellular portion of gp41 into a hairpin. This loop structure brings the virus and cell membranes close together, allowing fusion of the membranes and subsequent entry of the viral capsid.

After HIV has bound to the target cell, the HIV RNA and various enzymes (including *reverse transcriptase, integrase, ribonuclease, and protease*) are injected into the cell. Because HIV attachment is critical for the HIV replication cycle, understanding the specific mechanisms through which HIV attachment occurs has implications for potential treatments of HIV.

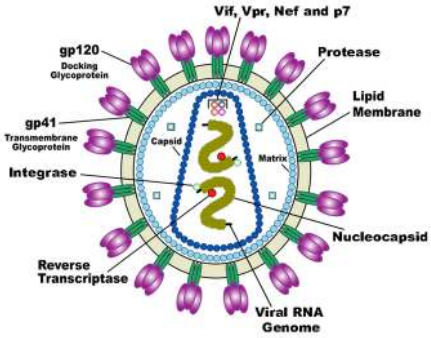


Figure 13.1.1: The attachment and fusion of HIV virons to host cells are crucial to allowing HIV infection to occur. Shown in purple is gp120 and in green gp41, two proteins crucial in viral docking to host cells.

Additional Reading

1. During adsorption, an envelope glycoprotein on the surface of HIV called Gp120 must adsorbs to both a $CD4^+$ receptor and then a chemokine receptor found on the surface of only certain types of certain human cells such as T4-lymphocytes, monocytes, macrophages, and dendritic cells.
2. Following adsorption, glycoprotein Gp41 enabling the viral envelope to fuse with the host cell membrane, allowing the nucleocapsid of the virus enters the host cell's cytoplasm.
3. During uncoating, the single-stranded RNA genomes within the capsid of the virus are released into the cytoplasm and HIV now uses the enzyme reverse transcriptase to make a single-stranded DNA copy of its single-stranded RNA genome. The reverse transcriptase then makes a complementary DNA strand to form a double-stranded viral DNA intermediate.
4. A viral enzyme called integrase then binds to the double-stranded viral DNA intermediate, transports it through the pores of the host cell’s nuclear membrane, and inserts into one of the host cell's chromosomes to form a provirus.
5. Following activation of the provirus, molecules of mostly polycistronic mRNA are transcribed off of the proviral DNA strand, go through the nuclear pores into the rough endoplasmic reticulum where it is translated by host cell's ribosomes HIV structural proteins, enzymes, glycoproteins, and regulatory proteins.
6. Polyproteins translated from polycistronic mRNAs must be cleaved into function proteins by HIV protease enzymes.
7. The two HIV envelope glycoproteins Gp41 and Gp120 are transported to the plasma membrane of the host cell where Gp41 anchors the Gp120 to the membrane of the infected cell. HIV obtains its envelope from the plasma membrane by budding.

8. Most maturation occurs either during the budding of the virion from the host cell or after its release from the cell.

Contributors

- Dr. Gary Kaiser (COMMUNITY COLLEGE OF BALTIMORE COUNTY, CATONSVILLE CAMPUS)
- Contributed by Boundless, General Microbiology at Boundless

13.1: Envelope glycoprotein GP120 is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

13.2: HIV-1 protease (PR)

Proteases are enzymes that have the ability to cut proteins into peptides. In order for maturation of HIV to occur, a HIV enzyme termed a protease has to cleave a long HIV-encoded gag-pol polyprotein to produce reverse transcriptase and integrase (coded by the HIV *pol* gene) and gag polyprotein (coded by the HIV *gag* gene). The HIV protease then cleaves the gag polyprotein into capsid protein p17, matrix protein p24, and nucleocapsid protein p7, as well as proteins p6, p2, and p1 whose functions are not yet fully understood. Proteases also cleave the env-polyprotein (coded by the HIV *env* gene) into the envelope glycoproteins gp120 and gp41. This allows the completion of the assembly step in the viral life cycle where the proteins and the viral RNA come together to form virion particles ready to exit the cell.

HIV-1 protease (PR) Inhibitor

Protease inhibitors are short peptide-like molecules that are competitive inhibitors of the enzyme. Instead of -NH-CO- peptide link, they contain -(CH_2-CH(OH)-). When such a peptide gets into the enzyme active site, the protease is unable to cut the linkage and gets inactivated. This leads to a lack of cleavage of the polypeptide chains of two crucial viral proteins, Gag and Pol, which are essential structural and enzymatic proteins of HIV. Their absence blocks the formation of mature virion particles.

Mutations in the enzyme active site and other sites, which cause conformational changes, can cause resistance. Quite often one mutation can lead to resistance to many different drugs simultaneously since they all share the same mode of action. This is called **cross-resistance**. It is one of the major drawbacks of protease inhibitors therapy.

Mechanism of HIV-1 protease

The mechanism of HIV-1 protease is still yet to be fully understood. The main way the mechanism are studied is through the use of mimicry substrates and simulations. HIV-1 protease has been studied intensely using various inhibitors, observing partial steps of the process. Since the main target of these inhibitors is to bind to the **Asp-25 of the catalytic triad**, each inhibitor would vary in its mechanism to accomplish this.

Figure 13.2.1: *HIV-1 Protease with Asp-25 of the catalytic triad. A single active site lies between the identical subunits and has the characteristic (Asp25, Thr26 and Gly27).*

The first part of the mechanism begins with the substrate binding onto the protease. ***Figure 13.2.1*** accents the key amino acids in HIV-1 protease that assists in substrate binding. It is predicted that a substrate first binds via a hydrogen bond to **aspartic acid 30** on one chain. Once this initial bond is made, the binding is then further stabilized by bondage to the glycine rich region in the flap of the same monomer. A salt bridge is then formed from the substrate to glutamic acid 35 of the other monomer. This completes binding of the substrate to the protease. At this point, waters molecules that are found at the tips of the flaps at isoleucine 50 on each monomer dissociates from the protease. The release of the water molecule results in a structural conformation change of the protease, changing it from semi-open to closed, tightening the space between the protease and substrate.

Once in the tightened state, aspartic acid 25 and 25' hydrogen binds to their adjacent glycine, and then becomes supported by the following threonine. Originally, there is a water molecule bound between the aspartic acids. One of the aspartic acid exists in a deprotonated state and the other one is protonated. The water molecule stabilizes the aspartates in this form.When the substrate binds to the protease, it causes conformational changes that brings the substrate to the position of the water molecule, and the water molecule acts as a nucleophile to the substrate. The oxygen of the water attacks the carbonyl group of the substrate peptide bond that is by the active site as the nitrogen picks up the hydrogen of the protonated aspartic acid. What results is an hydroxl group is

added to the carbonyl group as an amine is formed on other side of the peptide bond, leaving a hydrogen atom behind to stablize the two aspartates. This is proposed to occur in a concerted fashion.

Contributors

Curation and Revision. **Provided by**: Boundless.com. **License**: ***CC BY-SA: Attribution-ShareAlike***

13.2: HIV-1 protease (PR) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

13.3: HIV vaccine

A vaccine that can prevent infection would teach the immune system to respond to HIV by making antibodies that can bind to the virus and stop it from infecting cells. There are two approaches to an HIV vaccine:

1. Preventive HIV vaccine approach, is given to people who do not have HIV. The purpose of this vaccine to protect individuals from being infected with the virus.

2. Therapeutic HIV vaccine approach, is for the individuals who already have HIV. The goal of the vaccine is to strengthen a person's immune response to HIV.

Today, there is no licensed HIV vaccine available on the market, but several research projects are ongoing and scientific communities throughout the world are trying to find an effective ways to fight against the disease. Some HIV-infected individuals, but certainly not all, naturally produce **broadly neutralizing antibodies** which keep the virus suppressed and these people do not tend to see any clinical signs of HIV illness and remain asymptomatic for decades.

A neutralizing antibody defends a cell from an antigen or infectious body by inhibiting or neutralizing any effect it has biologically. The antibody response is crucial for preventing many viral infections and may also contribute to the resolution of an infection. When a vertebrate is infected with a virus, antibodies are produced against many epitopes of multiple virus proteins. A subset of these antibodies can block viral infection by a process called neutralization. This usually involves the formation of a virus-antibody complex.

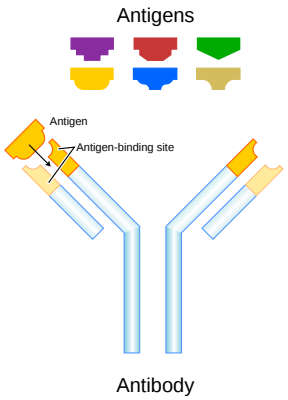


Figure13.3.1: Antibody neutralizing an antigen and preventing its biological effect.

This virus-antibody complex can prevent viral infections in many ways. It may interfere with virion binding to receptors, block uptake into cells, prevent uncoating of the genomes in endosomes, or cause aggregation of virus particles. Many enveloped viruses are lysed when antiviral antibodies and serum complement disrupt membranes. Antibodies can also neutralize viral infectivity by binding to cell surface receptors.

Neutralizing antibodies have shown potential in the treatment of retroviral infections. Medical professionals and researchers have shown how the encoding of genes which influence the production of this particular type of antibody could help in the treatment of infections that attack the immune system. Experts in the field have used HIV treatment as an example of infections these antibodies can treat. Recently, potent and broadly neutralizing human antibodies against influenza have been reported, and have suggested possible strategies to generate an improved vaccine that would confer long-lasting immunity. Another disease which has been linked to the production of neutralizing antibodies is multiple sclerosis.

In diagnostic immunology and virology laboratories, the evaluation of neutralizing antibodies, which destroy the infectivity of viruses, can be measured by the neutralization method. In this procedure, patient serum is mixed with a suspension of infectious virus particles of the same type as those suspected of causing disease in the patient. A control suspension of virus is mixed with normal serum and is then inoculated into an appropriate cell culture. If the patient serum contains antibody to the virus, the antibody will bind to the virus particles and prevent them from invading the cells in culture, thereby neutralizing the infectivity of the virus. This technique is labor-intensive, demanding, and time consuming. It application is restricted to laboratories that perform routine viral cultures and related diagnosis.

Summary

- Human immunodeficiency virus (HIV) is a sexually transmitted virus that infects and destroys helper T cells of the immune system. The virus can also be transmitted through contaminated blood or breast milk. HIV infection is diagnosed on the basis of a blood test for antibodies to the virus.
- AIDS stands for acquired immunodeficiency syndrome, is a disease that develops in people with untreated HIV infections, typically several years after their initial infection with the virus. AIDS is diagnosed when the immune system has been weakened to the point that it can no longer fight off opportunistic diseases that do not normally occur in healthy individuals.
- The development of new anti-retroviral drugs to treat HIV infection has changed the disease from fatal to a chronic disease. The drugs keep the virus at low levels, reducing the risk of transmission as well as infection progressing to AIDS.

13.3: HIV vaccine is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

13.4: Exercises

1. What cells are affected by HIV?
2. What happens to the number of HIV copies and the helper T cells over time in an infected individual?
3. Explain how HIV infection changed from a fatal to a chronic disease.
4. Explain why it is important to treat people exposed to or infected with HIV as early as possible with antiretroviral medications.

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

9: Glycolysis and Gluconeogenesis

In a well-fed animal, most cells can store a small amount of glucose as glycogen. All cells break glycogen down as needed to retrieve nutrient energy as G-6-P. Glycogen hydrolysis, or ***glycogenolysis***, produces G-1-P that is converted to G-6-P, as we saw at the top of *Stage* 1 of glycolysis. But, glycogen in most cells is quickly used up between meals. Therefore, most cells depend on a different, external source of glucose other than diet. Those sources are liver and to a lesser extent, kidney cells, that can store large amounts of glycogen after meals. In continual feeders (for examples cows and other ruminants), glycogenolysis is ongoing. In *intermittent feeders* (like us), liver glycogenolysis can supply glucose to the blood for 6-8 hours between meals, to be distributed as needed to all cells of the body. Thus, you can expect to use up liver and kidney glycogen reserves after a good night's sleep, a period of intense exercise, or any prolonged period of low carbohydrate intake (fasting or starvation). Under these circumstances, animals use **gluconeogenesis** (literally, *new glucose synthesis*) in liver and kidney cells to provide systemic glucose to nourish other cells. In healthy individuals, the hormones insulin and glucagon regulate blood *glucose homeostasis*, protecting against *hypoglycemia* (low blood sugar) and *hyperglycemia* (high blood sugar) respectively. The gluconeogenic pathway produces glucose from carbohydrate and non-carbohydrate precursor substrates. These precursors include pyruvate, lactate, glycerol and *gluconeogenic amino acids*.

A general summary of the several stages involved is shown in ***Figure 9.1.*** Initially, the storage fuels or foodstuffs (fats, carbohydrates, and proteins) are hydrolyzed into smaller components (fatty acids and glycerol, glucose and other simple sugars, and amino acids). In the next stage, these simple fuels are degraded further to two-carbon fragments that are delivered as the $CH_3C{=}O$ group (ethanoyl, or acetyl) in the form of the thioester of coenzyme A, $CH_3COSCoA$. This section is concerned mainly with the pathway by which glucose is metabolized by the process known as **glycolysis**.

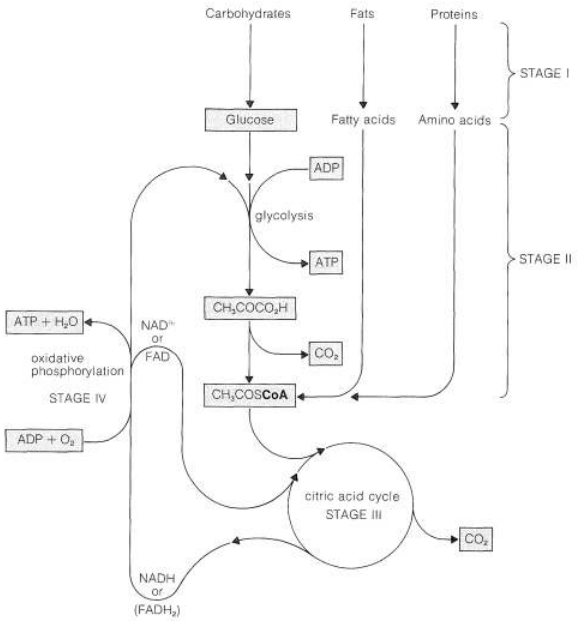


Figure 9.1: *Perspective of the metabolic scheme whereby carbohydrates, fats, and proteins in foodstuffs are oxidized to* CO_2, *showing the link between glycolysis, the citric acid cycle, and oxidative phosphorylation.*

Contributors

John D. Robert and Marjorie C. Caserio (1977) *Basic Principles of Organic Chemistry, second edition.* W. A. Benjamin, Inc. , Menlo Park, CA. ISBN 0-8053-8329-8. This content is copyrighted under the following conditions, "You are granted permission for individual, educational, research and non-commercial reproduction, distribution, display and performance of this work in any format."

9: Glycolysis and Gluconeogenesis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.1: Glycolysis - Reaction and Regulation

Glycolysis is the first step in the breakdown of glucose to extract energy for cellular metabolism. Nearly all living organisms carry out glycolysis as part of their metabolism. The process does not use oxygen and is therefore anaerobic. **Glycolysis takes place in the cytoplasm** of both prokaryotic and eukaryotic cells. Glucose enters heterotrophic cells in two ways. One method is through secondary **active transport** in which the transport takes place against the glucose concentration gradient. The other mechanism uses a group of integral proteins called GLUT proteins, also known as **glucose transporter proteins**. These transporters assist in the facilitated diffusion of glucose.

First Half of Glycolysis (Energy-Requiring Steps)

Step 1. The first step in glycolysis (Figure 9.1.1) is catalyzed by hexokinase, an enzyme with broad specificity that catalyzes the phosphorylation of six-carbon sugars. Hexokinase phosphorylates glucose using ATP as the source of the phosphate, producing glucose-6-phosphate, a more reactive form of glucose. This reaction prevents the phosphorylated glucose molecule from continuing to interact with the GLUT proteins, and it can no longer leave the cell because the negatively charged phosphate will not allow it to cross the hydrophobic interior of the plasma membrane.

Step 2. In the second step of glycolysis, an isomerase converts glucose-6-phosphate into one of its isomers, fructose-6-phosphate. An isomerase is an enzyme that catalyzes the conversion of a molecule into one of its isomers. (This change from phosphoglucose to phosphofructose allows the eventual split of the sugar into two three-carbon molecules.).

Step 3. The third step is the phosphorylation of fructose-6-phosphate, catalyzed by the enzyme phosphofructokinase. A second ATP molecule donates a high-energy phosphate to fructose-6-phosphate, producing fructose-1,6-bisphosphate. In this pathway, **phosphofructokinase is a rate-limiting enzyme**. It is active when the concentration of ADP is high; it is less active when ADP levels are low and the concentration of ATP is high. Thus, if there is "sufficient" ATP in the system, the pathway slows down. This is a type of end product inhibition, since ATP is the end product of glucose catabolism.

Step 4. The newly added high-energy phosphates further destabilize fructose-1,6-bisphosphate. The fourth step in glycolysis employs an enzyme, aldolase, to cleave 1,6-bisphosphate into two three-carbon isomers: dihydroxyacetone-phosphate and glyceraldehyde-3-phosphate.

Step 5. In the fifth step, an isomerase transforms the dihydroxyacetone-phosphate into its isomer, glyceraldehyde-3-phosphate. Thus, the pathway will continue with two molecules of a single isomer. At this point in the pathway, there is a net investment of energy from two ATP molecules in the breakdown of one glucose molecule.

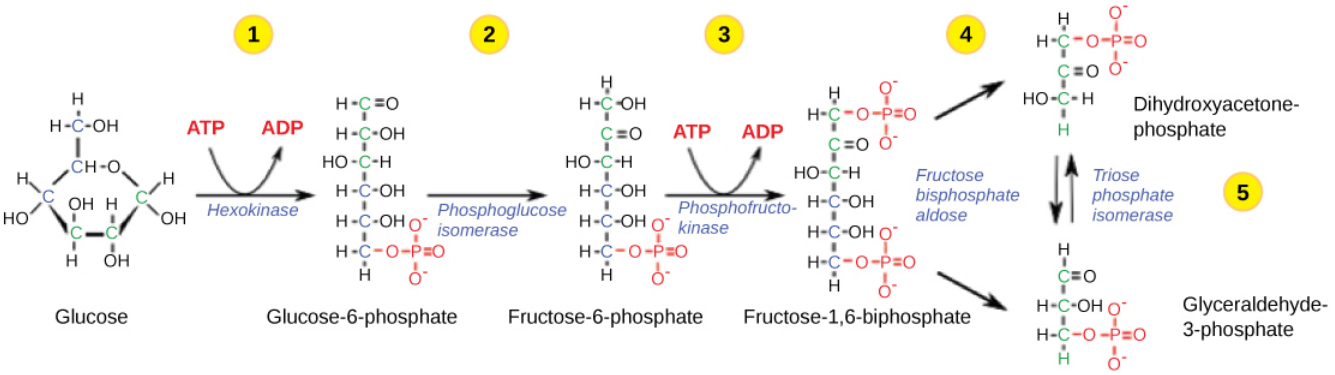


Figure 9.1.1: The first half of glycolysis uses two ATP molecules in the phosphorylation of glucose, which is then split into two three-carbon molecules.

Second Half of Glycolysis (Energy-Releasing Steps)

So far, glycolysis has cost the cell two ATP molecules and produced two small, three-carbon sugar molecules. Both of these molecules will proceed through the second half of the pathway, and sufficient energy will be extracted to pay back the two ATP molecules used as an initial investment and produce a profit for the cell of two additional ATP molecules and two even higher-energy NADH molecules.

Step 6. The sixth step in glycolysis (Figure 9.1.2) oxidizes the sugar (glyceraldehyde-3-phosphate), extracting high-energy electrons, which are picked up by the electron carrier NAD^+, producing NADH. The sugar is then phosphorylated by the addition of a second phosphate group, producing 1,3-bisphosphoglycerate. Note that the second phosphate group does not require another ATP molecule.

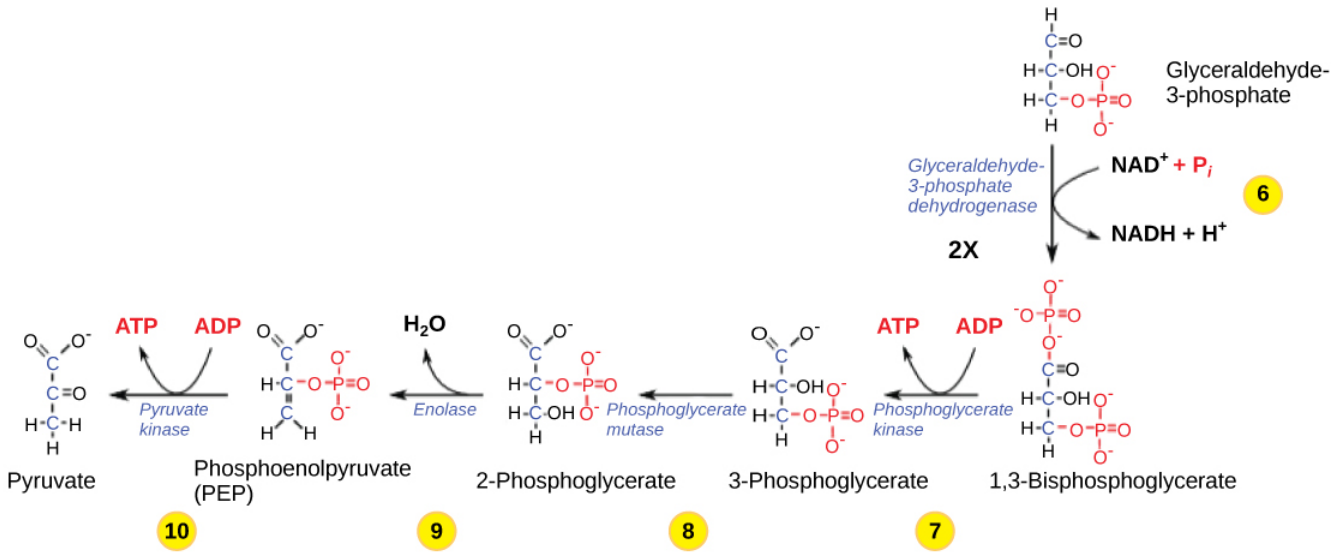


Figure 9.1.2: The second half of glycolysis involves phosphorylation without ATP investment (step 6) and produces two NADH and four ATP molecules per glucose.

Here again is a potential limiting factor for this pathway. The continuation of the reaction depends upon the availability of the oxidized form of the electron carrier, NAD^+. Thus, NADH must be continuously oxidized back into NAD^+ in order to keep this step going. If NAD^+ is not available, the second half of glycolysis slows down or stops. If oxygen is available in the system, the NADH will be oxidized readily, though indirectly, and the high-energy electrons from the hydrogen released in this process will be used to produce ATP. In an environment without oxygen, an alternate pathway (fermentation) can provide the oxidation of NADH to NAD^+.

Step 7. In the seventh step, catalyzed by phosphoglycerate kinase (an enzyme named for the reverse reaction), 1,3-bisphosphoglycerate donates a high-energy phosphate to ADP, forming one molecule of ATP. (This is an example of substrate-level phosphorylation.) A carbonyl group on the 1,3-bisphosphoglycerate is oxidized to a carboxyl group, and 3-phosphoglycerate is formed.

Step 8. In the eighth step, the remaining phosphate group in 3-phosphoglycerate moves from the third carbon to the second carbon, producing 2-phosphoglycerate (an isomer of 3-phosphoglycerate). The enzyme catalyzing this step is a mutase (isomerase).

Step 9. Enolase catalyzes the ninth step. This enzyme causes 2-phosphoglycerate to lose water from its structure; this is a dehydration reaction, resulting in the formation of a double bond that increases the potential energy in the remaining phosphate bond and produces phosphoenolpyruvate (PEP).

Step 10. The last step in glycolysis is catalyzed by the enzyme pyruvate kinase (the enzyme in this case is named for the reverse reaction of pyruvate's conversion into PEP) and results in the production of a second ATP molecule by substrate-level phosphorylation and the compound pyruvic acid (or its salt form, pyruvate). Many enzymes in enzymatic pathways are named for the reverse reactions, since the enzyme can catalyze both forward and reverse reactions (these may have been described initially by the reverse reaction that takes place in vitro, under non-physiological conditions).

*The **net reaction** in the transformation of glucose into pyruvate is:*

$$\text{Glucose} + 2\ \text{Pi} + 2\ \text{ADP} + 2\ \text{NAD}^+ \longrightarrow 2\ \text{Pyruvate} + 2\ \text{ATP} + 2\ \text{NADH} + 2\ \text{H}_2\text{O}$$

Thus, *two molecules of ATP are generated in the conversion of glucose into two molecules of pyruvate*.

Note that the energy released in the anaerobic conversion of glucose into two molecules of pyruvate is -21 kcal mol^{-1} (- 88 kJ mol^{-1}).

The Fates of Pyruvate

Pyruvic acid can be made from glucose through glycolysis, converted back to carbohydrates (such as glucose) via gluconeogenesis, or to fatty acids through acetyl-CoA. It can also be used to construct the amino acid alanine, and it can be converted into ethanol.

Pyruvic acid supplies energy to living cells through the citric acid cycle (also known as the Krebs cycle) when oxygen is present (aerobic respiration); when oxygen is lacking, it ferments to produce lactic acid. Pyruvate is an important chemical compound in biochemistry. It is the output of the anaerobic metabolism of glucose known as glycolysis. One molecule of glucose breaks down into two molecules of pyruvate, which are then used to provide further energy in one of two ways. Pyruvate is converted into acetyl- coenzyme A, which is the main input for a series of reactions known as the Krebs cycle.

The net reaction of converting pyruvate into acetyl CoA and CO_2 is:

$$\text{2 Pyruvate} + \text{2NAD}^+ + \text{2 CoA} \longrightarrow \text{2 Acetyl CoA} + \text{2 NADH} + \text{2CO}_2$$

Pyruvate is also converted to oxaloacetate by an anaplerotic reaction, which replenishes Krebs cycle intermediates; also, oxaloacetate is used for gluconeogenesis. These reactions are named after Hans Adolf Krebs, the biochemist awarded the 1953 Nobel Prize for physiology, jointly with Fritz Lipmann, for research into metabolic processes. The cycle is also known as the citric acid cycle or tri-carboxylic acid cycle, because citric acid is one of the intermediate compounds formed during the reactions.

If insufficient oxygen is available, the acid is broken down anaerobically, creating lactate in animals and ethanol in plants and microorganisms. Pyruvate from glycolysis is converted by fermentation to lactate using the enzyme lactate dehydrogenase and the coenzyme NADH in lactate fermentation. Alternatively it is converted to acetaldehyde and then to ethanol in alcoholic fermentation.

Pyruvate is a key intersection in the network of metabolic pathways. Pyruvate can be converted into carbohydrates via gluconeogenesis, to fatty acids or energy through acetyl-CoA, to the amino acid alanine, and to ethanol. Therefore, it unites several key metabolic processes.

Regulation

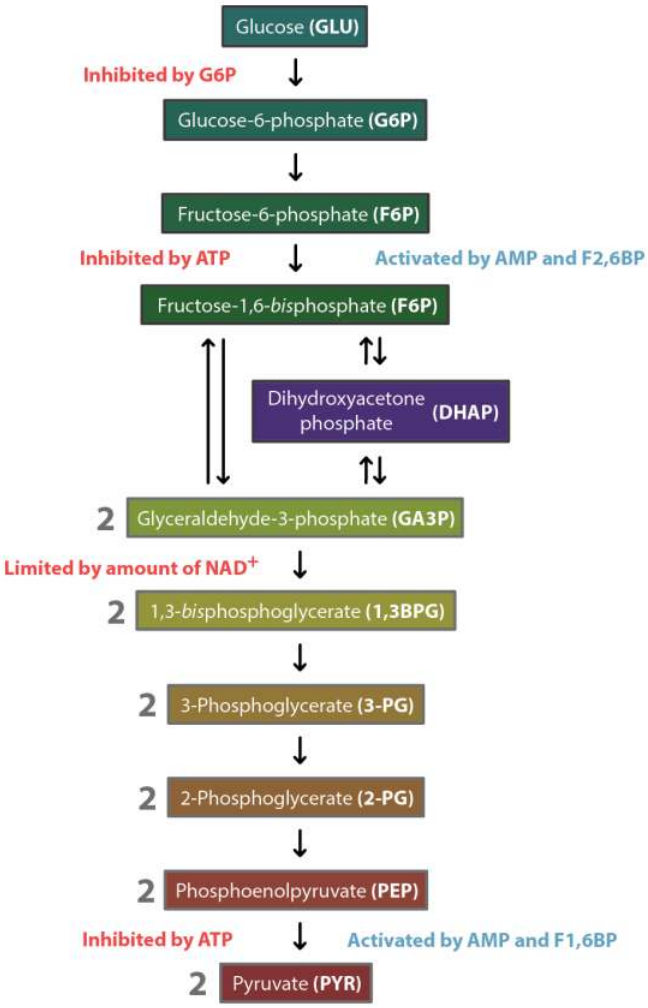


Figure 9.1.3**:** Glycolysis Regulation

Control of glycolysis is unusual for a metabolic pathway, in that regulation occurs at three enzymatic points:

$$\underbrace{\text{Glucose} \rightleftharpoons \text{G}_6\text{P}}_{\text{hexokinase}}$$

$$\underbrace{\text{F}_6\text{P} \rightleftharpoons \text{F}_1,6\,\text{BP}}_{\text{phosphofructokinase (PFK)}}$$

$$\underbrace{\text{PEP} \rightleftharpoons \text{pyruvate}.}_{\text{pyruvate kinase}}$$

Glycolysis is regulated in a reciprocal fashion compared to its corresponding anabolic pathway, gluconeogenesis. Reciprocal regulation occurs when the same molecule or treatment (phosphorylation, for example) has opposite effects on catabolic and anabolic pathways. Reciprocal regulation is important when anabolic and corresponding catabolic pathways are occurring in the same cellular location.

As an example, consider regulation of PFK. It is activated by several molecules, most importantly fructose-2,6- bisphosphate (F2,6BP). This molecule has an inhibitory effect on the corresponding gluconeogenesis enzyme, fructose-1,6-bisphosphatase (F1,6BPase).

You might wonder why pyruvate kinase, the last enzyme in the pathway, is regulated. The answer is simple. Pyruvate kinase catalyzes the most energetically rich reaction of glycolysis. The reaction is favored so strongly in the forward direction that cells must do a ‘two-step’ around it in the reverse direction when making glucose. In other words, it takes two enzymes, two reactions, and two triphosphates to go from pyruvate back to PEP in gluconeogenesis. When cells are needing to make glucose, they can’t be sidetracked by having the PEP they have made in gluconeogenesis be converted directly back to pyruvate by pyruvate kinase. Consequently, pyruvate kinase is inhibited during gluconeogenesis, lest a “futile cycle" occur.

Another interesting control mechanism called feedforward activation involves pyruvate kinase. Pyruvate kinase is activated allosterically by F1,6BP. This molecule is a product of the PFK reaction and a substrate for the aldolase reaction. It should be noted that the aldolase reaction is energetically unfavorable (high +ΔΔG°’), thus allowing F1,6BP to accumulate. When this happens, some of the excess F1,6BP activates pyruvate kinase, which jump-starts the conversion of PEP to pyruvate. The resulting drop in PEP levels has the effect of “pulling" on the reactions preceding pyruvate kinase. As a consequence, the concentrations of G3P and DHAP fall, helping to move the aldolase reaction forward.

Outcomes of Glycolysis

Glycolysis starts with one molecule of glucose and ends with two pyruvate (pyruvic acid) molecules, a total of four ATP molecules, and two molecules of NADH. Two ATP molecules were used in the first half of the pathway to prepare the six-carbon ring for cleavage, so the cell has a net gain of two ATP molecules and 2 NADH molecules for its use. If the cell cannot catabolize the pyruvate molecules further (via the citric acid cycle or Krebs cycle), it will harvest only two ATP molecules from one molecule of glucose.

Mature mammalian red blood cells do not have mitochondria and are not capable of aerobic respiration, the process in which organisms convert energy in the presence of oxygen. Instead, glycolysis is their sole source of ATP. Therefore, if glycolysis is interrupted, the red blood cells lose their ability to maintain their sodium-potassium pumps, which require ATP to function, and eventually, they die. For example, since the second half of glycolysis (which produces the energy molecules) slows or stops in the absence of NAD^+, when NAD^+ is unavailable, red blood cells will be unable to produce a sufficient amount of ATP in order to survive.

Additionally, the last step in glycolysis will not occur if pyruvate kinase, the enzyme that catalyzes the formation of pyruvate, is not available in sufficient quantities. In this situation, the entire glycolysis pathway will continue to proceed, but only two ATP molecules will be made in the second half (instead of the usual four ATP molecules). Thus, pyruvate kinase is a rate-limiting enzyme for glycolysis.

Contributors

- Darik Benson, (University California Davis)
- Dr. Kevin Ahern and Dr. Indira Rajagopal (Oregon State University)

9.1: Glycolysis - Reaction and Regulation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.2 Gluconeogenesis: Reaction and regulation

Introduction

The need for energy is important to sustain life. Organisms have evolved ways of producing substrates required for the catabolic reactions necessary to sustain life when desired substrates are unavailable. The main source of energy for eukaryotes is glucose. When glucose is unavailable, organisms are capable of metabolizing glucose from other non-carbohydrate precursors. The process that coverts pyruvate into glucose is called gluconeogenesis.

Gluconeogenesis Is Not a Reversal of Glycolysis

In glycolysis, glucose is converted into pyruvate; in **gluconeogenesis**, pyruvate is converted into glucose. However, ***gluconeogenesis*** *is not a reversal of glycolysis*. Several reactions must differ because the equilibrium of glycolysis lies far on the side of pyruvate formation. The actual ΔG for the formation of pyruvate from glucose is about -20 kcal mol^{-1} (-84 kJ mol^{-1}) under typical cellular conditions. Most of the decrease in free energy in glycolysis takes place in the three essentially irreversible steps catalyzed by hexokinase, phosphofructokinase, and pyruvate kinase.

$$\text{Glucose + ATP} \xrightarrow{\text{Hexokinase}} \text{Glucose 6- Phosphate + ADP}$$
$$\triangle G = -8.0 \text{ kcal mol}^{-1} (-33 \text{ kJ mol}^{-1})$$

$$\text{Fructose 6-phosphate + AT} \xrightarrow{\text{Phosphofructokinas}} \text{fructose 1,6-bisphosphate + ADP}$$
$$\triangle G = -5.3 \text{ kcal mol}^{-1} (-22 \text{ kJ mol}^{-1})$$

$$\textbf{Phosphoenolpyruvate + ADP} \xrightarrow{\text{Pyruvate kinase}} \textbf{pyruvate + ATP}$$
$$\triangle G = -4.0 \text{ kcal mol}^{-1} (-17 \text{ kJ mol}^{-1})$$

In **gluconeogenesis**, the following new steps bypass these virtually irreversible reactions of glycolysis:

1. Phosphoenolpyruvate is formed from pyruvate by way of oxaloacetate through the action of pyruvate carboxylase and phosphoenolpyruvate carboxykinase.

$$\textbf{Pyruvate + } CO_2 \textbf{ + ATP + } H_2O \xrightarrow{\text{Pyruvate carboxylase}} \textbf{oxaloacetate + ADP + } P_i \textbf{ + 2 } H^+$$

$$\textbf{Oxaloacetate + GTP} \xrightarrow{\text{Phosphoenolpyruvate carboxykinase}} \textbf{phosphoenolpyruvate + GDP + } CO_2$$

2. *Fructose 6-phosphate is formed from fructose 1,6-bisphosphate by hydrolysis of the phosphate ester at carbon 1*. Fructose 1,6-bisphosphatase catalyzes this exergonic hydrolysis.

$$\textbf{Fructose 1,6-bisphosphate + } H_2O \longrightarrow \textbf{fructose 6-phosphate + } P_i$$

3. *Glucose is formed by hydrolysis of glucose 6-phosphate* in a reaction catalyzed by glucose 6-phosphatase.

$$\textbf{Glucose 6-phosphate+ } H_2O \longrightarrow \textbf{Glucose + Pi}$$

Regulation

It is important for organisms to conserve energy, they have derived ways to regulate those metabolic pathways that require and release the most energy. In glycolysis and gluconeogenesis seven of the ten steps occur at or near equilibrium. In gluconeogenesis the conversion of pyruvate to glucose all occur very spontaneously which is why these processes are highly regulated. It is important for the organism to conserve as much energy as possible. When there is an excess of energy available, gluconeogenesis is inhibited. When energy is required, gluconeogenesis is activated.

1. The conversion of pyruvate to PEP is regulated by acetyl-CoA. Because acetyl-CoA is an important metabolite in the TCA cycle which produces a lot of energy, when concentrations of acetyl-CoA are high organisms use pyruvate carboxylase to channel pyruvate away from the TCA cycle. If the organism does not need more energy, then it is best to divert those metabolites towards storage or other necessary processes.

2. The conversion of fructose-1,6-bisphosphate to fructose-6-phosphate with the use of fructose-1,6-phosphatase is negatively regulated and inhibited by the molecules AMP and fructose-2,6-bP. These are reciprocal regulators to glycolysis' phosphofructokinase. Phosphofructosekinase is positively regulated by AMP and fructose-2,6-bP. Once again, when the energy levels produced are higher than needed, i.e. a large ATP to AMP ratio, the organism increases gluconeogenesis and decreases glycolysis. The opposite also applies when energy levels are lower than needed, i.e. a low ATP to AMP ratio, the organism increases glycolysis and decreases gluconeogenesis.
3. The conversion of glucose-6-P to glucose with use of glucose-6-phosphatase is controlled by substrate level regulation. The metabolite responsible for this type of regulation is glucose-6-P. As levels of glucose-6-P increase, glucose-6-phosphatase increases activity and more glucose is produced. Thus glycolysis is unable to proceed.

Importance

This metabolic pathway is important because the brain depends on glucose as its primary fuel and red blood cells use only glucose as a fuel. The daily glucose requirement of the brain in a typical adult human being is about 120 g, which accounts for most of the 160 g of glucose needed daily by the whole body. The amount of glucose present in body fluids is about 20 g, and that readily available from glycogen, a storage form of glucose, is approximately 190 g. Thus, the direct glucose reserves are sufficient to meet glucose needs for about a day. During a longer period of starvation, glucose must be formed from noncarbohydrate sources.

The major site of **gluconeogenesis** is the *liver*, with a small amount also taking place in the *kidney, brain, skeletal muscle, or heart muscle*.

References

Garrett, H., Reginald and Charles Grisham. Biochemistry. Boston: Twayne Publishers, 2008.

Raven, Peter. Biology. Boston: Twayne Publishers, 2005.

Contributors

Gerald Bergtrom

Professor Emeritus (Biosciences) at University of Wisconsin-Milwaukee

9.2 Gluconeogenesis: Reaction and regulation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

9.3: Exercises

Problems

1.What are the three stages of cellular respiration?

2.What is the purpose of glycolysis?

3.What is the output of glycolysis from a single glucose molecule?

4.How many molecules of ATP are "invested" in glycolysis? How many are produced?

5. Name the enzymes that are key regulatory sites in Glycolysis.

6. Draw the entire pathway for glycolysis including enzymes, reactants and products for each step.

8. How many enzymes are unique to Gluconeogenesis?

9. What is reciprocal regulation and why is it important to Glycolysis and Gluconeogenesis?

10. Where does the activity of glucose-6-phosphatase occur?

11. Why is it necessary for gluconeogenesis to incorporate other enzymes in its pathway that are different from glycolysis?

12. Why is it important for organisms to conserve energy?

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

Index

A

affinity chromatography
- 3.4.1. Affinity Chromatography

allosteric competitive inhibition
- 7.1: Control of Metabolism Through Enzyme Regulation

Alzheimer's disease
- 2.4: Protein Folding and Prions

amino acids
- 1.1: Nomenclature of Amino acids

B

Blaber
- 1.3.2. Stereochemistry of Amino Acids

C

Cell Disruption
- 3.2: Cell Disruption

citric acid cycle
- 9: Glycolysis and Gluconeogenesis

competitive inhibition
- 7.1: Control of Metabolism Through Enzyme Regulation

D

denaturation
- 2.5: Denaturation of proteins

DNA repair
- 12: The Flow of Genetic Information: from DNA to RNA and Proteins

E

Edman degradation
- 2.2: Protein Sequencing

enzyme regulation
- 7.1: Control of Metabolism Through Enzyme Regulation

F

Fractionation
- 3.3:Cell Fractionation and Centrifugation

G

gel electrophoresis
- 3.5: Electrophoresis

Gel Exclusion Chromatography
- 3.4.2. Gel Exclusion Chromatography

gel filtration chromatography
- 3.4.2. Gel Exclusion Chromatography

gluconeogenesis
- 9: Glycolysis and Gluconeogenesis

glycolysis
- 9: Glycolysis and Gluconeogenesis

H

HIV
- 13: Integrated chapter (HIV)
- 13.3: HIV vaccine

HIV gp120
- 13.3: HIV vaccine

HIV vaccine
- 13: Integrated chapter (HIV)
- 13.3: HIV vaccine

I

Ion Exchange Chromatography
- 3.4.3. Ion Exchange Chromatography

Isoelectric focusing
- 3.5: Electrophoresis

isoelectric point
- 1.4.2. Acid-Base Reactions of Amino Acids

K

Krebs cycle
- 9: Glycolysis and Gluconeogenesis

L

lagging strand
- 12.2: DNA Replication

leading strand
- 12.2: DNA Replication

Lysis
- 3.2: Cell Disruption

N

Noncompetitive Inhibition
- 7.1: Control of Metabolism Through Enzyme Regulation

northern blot
- 3.5.1: Blotting

O

oxidative phosphorylation
- 9: Glycolysis and Gluconeogenesis

P

pentose phosphate pathway
- 9: Glycolysis and Gluconeogenesis

peptide bond
- 1.2: Structure of Amino Acids

phosphorylation
- 11: Electron Transport Chain and Oxidative Phosphorylation

polysaccharide
- 8.5: Polysaccharides

protein folding
- 2.4: Protein Folding and Prions

R

Renaturation
- 2.5: Denaturation of proteins

Replication
- 12.2: DNA Replication

repressible system
- 7: Regulation of Enzyme Activity

Repressors
- 7: Regulation of Enzyme Activity

RNA
- 12.4: RNA

S

Sonication
- 3.2: Cell Disruption

Southern blot
- 3.5.1: Blotting

stop codons
- 12.6: Translation

T

Transcription
- 12: The Flow of Genetic Information: from DNA to RNA and Proteins

translation
- 12: The Flow of Genetic Information: from DNA to RNA and Proteins

W

Western blot
- 3.5.1: Blotting

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Chem 4320/5320: Biochemistry 1

Webpages: 95

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 77.9% (74 pages)
- Undeclared: 13.7% (13 pages)
- CC BY-NC 4.0: 4.2% (4 pages)
- CC BY-SA 4.0: 2.1% (2 pages)
- CC BY 4.0: 2.1% (2 pages)

By Page

- Chem 4320/5320: Biochemistry 1 - *CC BY-NC-SA 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Properties of the Twenty Common Amino Acids - *CC BY-NC-SA 4.0*
    - 1.1: Nomenclature of Amino acids - *CC BY-NC-SA 4.0*
    - 1.2: Structure of Amino Acids - *CC BY-NC-SA 4.0*
    - 1.3: Properties of Amino Acids - *CC BY-NC-SA 4.0*
      - 1.3.1. Charged Nature of Amino Acid - *CC BY-NC-SA 4.0*
      - 1.3.2. Stereochemistry of Amino Acids - *CC BY-NC-SA 4.0*
    - 1.4: Reactions of Amino Acids - *CC BY-NC-SA 4.0*
      - 1.4.1 Acid-base Chemistry of Amino Acids - *CC BY-NC-SA 4.0*
      - 1.4.2. Acid-Base Reactions of Amino Acids - *CC BY-NC 4.0*
  - 2: Proteins Structure: from Amino Acid Sequence to Three Dimensional Structure - *CC BY-NC-SA 4.0*
    - 2.1: The Structure of Proteins - *CC BY-NC 4.0*
    - 2.2: Protein Sequencing - *CC BY-SA 4.0*
    - 2.3: Protein Structural Determination - *CC BY-NC-SA 4.0*
    - 2.4: Protein Folding and Prions - *CC BY-NC-SA 4.0*
    - 2.5: Denaturation of proteins - *CC BY-NC-SA 4.0*
    - 2.6: Amino Acids and Proteins (Exercises) - *CC BY-NC-SA 4.0*
  - 3: Methods of Protein Purification and Characterization - *CC BY-NC-SA 4.0*
    - 3.1: Protein Purification - *CC BY-NC-SA 4.0*
    - 3.2: Cell Disruption - *CC BY-NC-SA 4.0*
    - 3.3:Cell Fractionation and Centrifugation - *CC BY-NC-SA 4.0*
    - 3.4: Chromatography - *CC BY-NC-SA 4.0*
      - 3.4.1. Affinity Chromatography - *CC BY-NC-SA 4.0*
      - 3.4.2. Gel Exclusion Chromatography - *CC BY-NC-SA 4.0*
      - 3.4.3. Ion Exchange Chromatography - *CC BY-NC-SA 4.0*
    - 3.5: Electrophoresis - *CC BY-NC-SA 4.0*
      - 3.5.1: Blotting - *CC BY-SA 4.0*
      - 3.5.2: Exercises - *CC BY-NC-SA 4.0*
  - 4: Overview of Hemoglobin and Myoglobin - *CC BY-NC-SA 4.0*
    - 4.1: Myoglobin, Hemoglobin, and their Ligands - *CC BY-NC-SA 4.0*
    - 4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin - *CC BY-NC-SA 4.0*
    - 4.3: Exercises - *CC BY-NC-SA 4.0*
  - 5: Michaelis-Menten Enzyme Kinetics, Inhibitors, pH optima; Bi-Substrate Reactions - *CC BY-NC-SA 4.0*
    - 5.1: Catalytic Efficiency of Enzymes - *CC BY-NC-SA 4.0*
    - 5.2: Enzyme Parameters - *CC BY-NC-SA 4.0*
    - 5.3: Michaelis-Menten Kinetics - *CC BY-NC-SA 4.0*
    - 5.4: Enzyme Inhibition - *CC BY-NC-SA 4.0*
    - 5.5: Temperature, pH, and enzyme concentration on the rate of a reaction - *CC BY-NC-SA 4.0*

- 5.6: Multi-Substrate Sequential Mechanisms - *CC BY-NC-SA 4.0*
- 5.7: Double displacement reaction - *CC BY 4.0*

- 6: Classification and Catalytic Strategies of Enzymes - *CC BY-NC-SA 4.0*
  - 6.1: Serine proteases - *CC BY-NC-SA 4.0*
  - 6.2: Transititon State Analogs and Catalytic Antibodies - *CC BY-NC-SA 4.0*
  - 6.3: Restriction Endonuclease - *CC BY-NC-SA 4.0*
- 7: Regulation of Enzyme Activity - *CC BY 4.0*
  - 7.1: Control of Metabolism Through Enzyme Regulation - *CC BY-NC-SA 4.0*
  - 7.2: Amino Acids, Proteins, and Enzymes (Summary) - *CC BY-NC-SA 4.0*
  - 7.3: Exercises - *CC BY-NC-SA 4.0*
- 8: Carbohydrate Structures, Stereochemistry, and Glycosides - *CC BY-NC-SA 4.0*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
  - 8.1: Carbohydrates Fundamentals - *CC BY-NC-SA 4.0*
  - 8.2: Monosaccharides - *CC BY-NC-SA 4.0*
  - 8.3: Disaccharides - *CC BY-NC-SA 4.0*
  - 8.4: Oligosaccharides - *CC BY-NC-SA 4.0*
  - 8.5: Polysaccharides - *CC BY-NC-SA 4.0*
  - 8.6: Exercises - *CC BY-NC-SA 4.0*
  - Back Matter - *Undeclared*
    - Index - *Undeclared*
- 9: Glycolysis and Gluconeogenesis - *CC BY-NC-SA 4.0*
  - 9.1: Glycolysis - Reaction and Regulation - *CC BY-NC-SA 4.0*
  - 9.2 Gluconeogenesis: Reaction and regulation - *CC BY-NC-SA 4.0*
  - 9.3: Exercises - *CC BY-NC-SA 4.0*
- 10: Pyruvate Dehydrogenase Links Glycolysis to Krebs Cycle - *CC BY-NC-SA 4.0*
  - 10.1: The Krebs Cycle (Citric Acid Cycle) - *CC BY-NC-SA 4.0*
  - 10.2: Exercises - *CC BY-NC-SA 4.0*
- 11: Electron Transport Chain and Oxidative Phosphorylation - *CC BY-NC-SA 4.0*
  - 11.1: ETC and Oxidative Phosphorylation - *Undeclared*
  - 11.1: Exercises - *CC BY-NC-SA 4.0*
- 12: The Flow of Genetic Information: from DNA to RNA and Proteins - *CC BY-NC-SA 4.0*
  - 12.1: The Structure of DNA - *CC BY-NC 4.0*
  - 12.2: DNA Replication - *CC BY-NC-SA 4.0*
  - 12.3: DNA Repair - *CC BY-NC-SA 4.0*
  - 12.4: RNA - *CC BY-NC-SA 4.0*
    - 12.4.1 Types of RNA - *CC BY-NC-SA 4.0*
    - 12.4.2. RNA - Transcription - *CC BY-NC-SA 4.0*
    - 12.4.3. Regulation of Transcription - *CC BY-NC-SA 4.0*
  - 12.5: The Genetic Code - *CC BY-NC 4.0*
  - 12.6: Translation - *CC BY-NC-SA 4.0*
  - 12.7: Exercises - *CC BY-NC-SA 4.0*
- 13: Integrated chapter (HIV) - *CC BY-NC-SA 4.0*
  - 13.1: Envelope glycoprotein GP120 - *CC BY-NC-SA 4.0*
  - 13.2: HIV-1 protease (PR) - *CC BY-NC-SA 4.0*
  - 13.3: HIV vaccine - *CC BY-NC-SA 4.0*
  - 13.4: Exercises - *CC BY-NC-SA 4.0*
- Back Matter - *CC BY-NC-SA 4.0*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*