BIOCHEMISTRY FREE FOR ALL



Kevin Ahern, Indira Rajagopal, & Taralyn Tan
Oregon State University



Oregon State University

Biochemistry Free For All

Kevin Ahern, Indira Rajagopal, & Taralyn Tan

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.

The LibreTexts name, logo, and book covers are trademarked by LibreTexts, Inc. (a 501(c)(3) not-for-profit organization) and protected against unauthorized use.

This text was compiled on 08/19/2026

TABLE OF CONTENTS

Licensing

1: In The Beginning

1.1: Introduction - Basic Biology

1.2: Introduction - Basic Chemistry

1.3: Introduction - Water and Buffers

2: Structure and Function

2.1: Prelude to Structure and Function

2.4: Structure and Function- Proteins II

2.5: Structure and Function- Protein Function II

2.6: Structure and Function - Nucleic Acids

2.7: Structure and Function- Carbohydrates

2.8: Structure and Function - Lipids and Membranes

2.2: Structure & Function - Amino Acids

2.3: Structure & Function- Proteins I

3: Membranes

3.1: Basic Concepts in Membranes

3.2: Transport in Membranes

3.3: Other Considerations in Membranes

4: Catalysis

4.1: Basic Principles of Catalysis

4.2: Control of Enzymatic Activity

4.3: Mechanisms of Catalysis

4.4: Blood Clotting

5: Energy

5.1: Basics of Energy

5.3: Energy - Photophosphorylation

5.2: Electron Transport and Oxidative Phosphorylation

6: Metabolism

6.1: Metabolism - Sugars

6.2: Citric Acid Cycle & Related Pathways

6.3: Fats and Fatty Acids

6.4: Other Lipids

6.5: Amino Acids and the Urea Cycle

6.6: Nucleotides

7: Information Processing

7.1: Prelude to Information Processing

7.2: Genes and Genomes

7.3: DNA Replication

7.4: DNA Repair

7.5: Transcription

7.6: RNA Processing

7.7: Translation

7.8: Gene Expression

7.9: Signaling

8: Basic Techniques

8.1: Cell Lysis

8.2: Fractionation and Chromatography Techniques

8.3: Electrophoresis

8.4: Detection, identification and quantitation of specific nucleic acids and proteins

8.5: Transcriptomics

8.6: Isolating Genes

8.7: Polymerase Chain Reaction (PCR)

8.8: Reverse Transcription

8.9: FRET

8.10: Genome Editing (CRISPR)

8.11: Protein Cleavage

8.12: Membrane Dynamics (FRAP)

9: Chapter 10

9.1: Section 1-

9.2: Section 2-

9.3: Section 3-

9.4: Section 4-

9.5: Section 5-

9.6: Section 6-

10: Chapter 11

10.1: Section 1-

10.2: Section 2-

10.3: Section 3-

10.4: Section 4-

10.5: Section 5-

10.6: Section 6-

11: Point by Point

11.1: In the Beginning

11.2: Structure and Function
11.3: Membranes
11.4: Catalysis
11.5: Energy
11.6: Metabolism
11.7: Information Processing
11.8: Techniques

Index

Glossary

Detailed Licensing

This page titled Biochemistry Free For All (Ahern, Rajagopal, and Tan) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

Licensing

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

This page titled Licensing is shared under a license and was authored, remixed, and/or curated by LibreTexts.

1: IN THE BEGINNING

1.1: Introduction - Basic Biology
1.2: Introduction - Basic Chemistry
1.3: Introduction - Water and Buffers

This page titled 1: In The Beginning is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

1.1: Introduction - Basic Biology

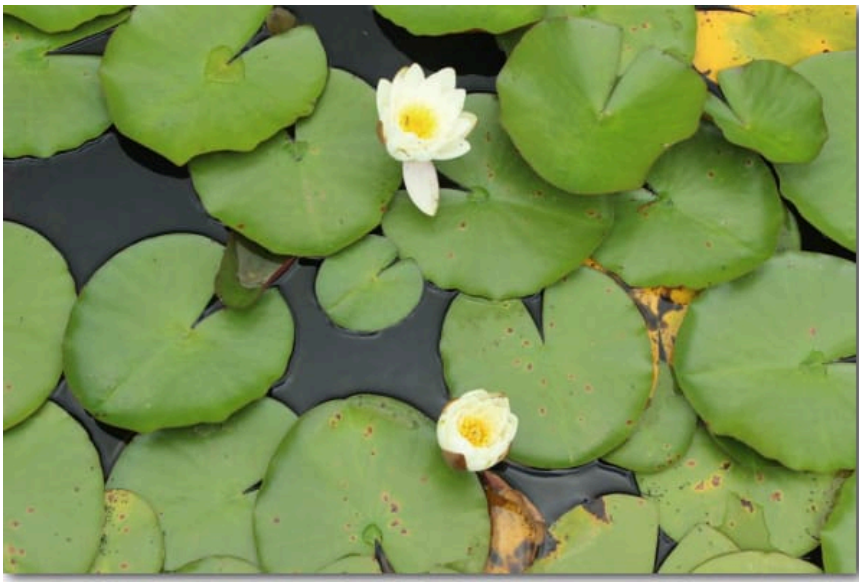


Figure 1.1 - Tardigrade Wikipedia

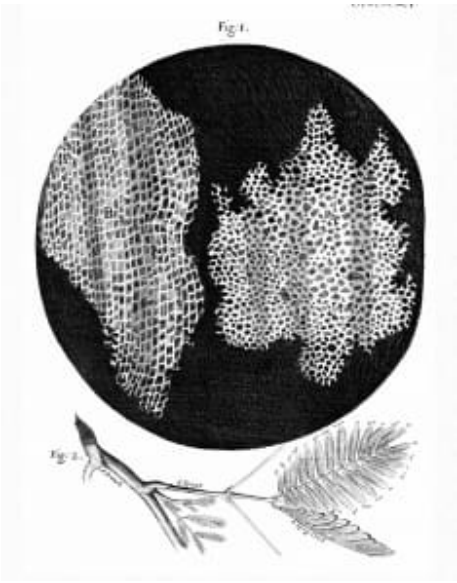


Figure 1.2 Slices of cork as seen by Hooke

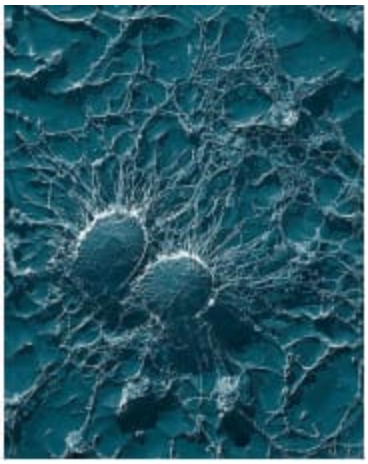


Figure 1.17 - Nerve cell anatomy Image by Pehr Jacobson

YouTube Lectures by Kevin HERE & HERE

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

This page titled 1.1: Introduction - Basic Biology is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

1.2: Introduction - Basic Chemistry

Source: BiochemFFA_1_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

"Organic chemistry is the chemistry of carbon compounds. Biochemistry is the chemistry of carbon compounds that crawl" -Michael Adams.

To understand biochemistry, one must possess at least a basic understanding of organic and general chemistry. In this brief section, we will provide a rapid review of the simple concepts necessary to understand cellular chemistry. Chemistry is chemistry, whether in a cell or outside it, but biological chemistry is a particular subset of organic chemistry that often involves enormous macromolecules, and that happens in the aqueous environment of the cell.



Figure 1.18 shows the various organic functional groups common in biochemistry. You will encounter these functional groups as you study the biosynthetic and breakdown pathways that build and recycle the chemical compounds of which cells are made. In addition to knowing the names and structures of these groups, students need a basic understanding of covalent and ionic bonds.

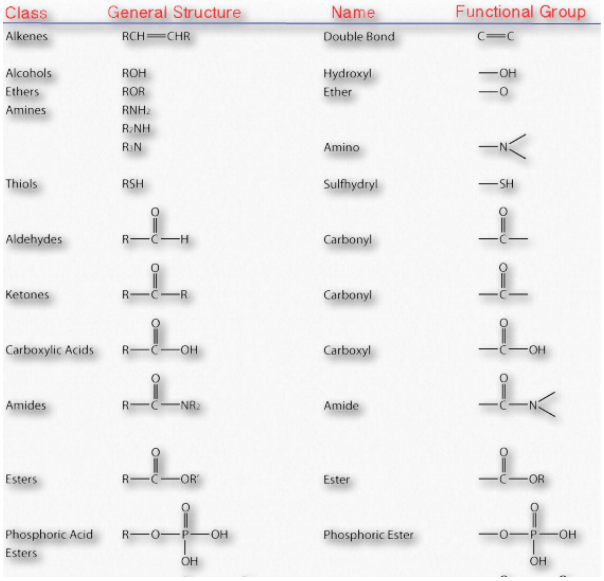


Figure 1.18 - Important functional groups in biochemistry Image by Aleia Kim

Covalent bonds, as you know, are the result of sharing of electrons between two atoms. Ionic bonds, by contrast, are formed when one atom donates an electron to another, such as in the formation of sodium chloride. Single covalent bonds can rotate freely, but double bonds cannot. Single bonds around a carbon atom are arranged in a tetrahedron with bond angles of 109.5° relative to each other, with the carbon at the center (Figure 1.19). Double bonded carbons create a planar structure with bond angles typically of about 120°.

Electronegativity

Electronegativity is a measure of the affinity a nucleus has for outer shell electrons (Table 1.2). High electronegativity corresponds to high affinity. Electrons in a covalent bond are held closer to the nucleus with a greater electronegativity compared to a nucleus with lower electronegativity.

Table 1.2

Electronegativities of Various Atoms

Atom	Electronegativity
Oxygen	3.5
Nitrogen	3.0
Sulfur	2.6
Carbon	2.5
Phosphorous	2.2
Hydrogen	2.1

Table 1.2 Image by Aleia Kim

For example, in a molecule of water, with hydrogen covalently bonded to oxygen, the electrons are “pulled” toward the oxygen, which is more electronegative. Because of this, there is a slightly greater negative charge near the oxygen atom of water, compared to the hydrogen (which, correspondingly has a slightly higher positive charge). This unequal charge distribution sets up a dipole, with one side being somewhat negative and the other somewhat positive. Because of this, the molecule is described as polar.

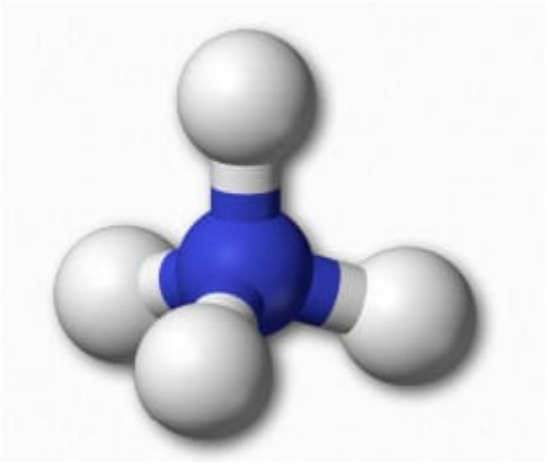


Figure 1.19 - Tetrahedral structure Wikipedia

Hydrogen bonds between water molecules are the result of the attraction of the partial positive and partial negative charges

on different water molecules (Figure 1.20). Hydrogen bonds can also form between hydrogens with a partial positive charge and other strongly electronegative atoms, like nitrogen, with a partial negative charge. It is important to remember that hydrogen bonds are interactions between molecules (or parts of molecules) and are not bonds between atoms, like covalent or ionic bonds. Bonds between hydrogen and carbon do not form significant partial charges because the electronegativities of the two atoms are similar. Consequently, molecules containing many carbon-hydrogen bonds will not form hydrogen bonds and therefore, do not mix well with water. Such molecules are called hydrophobic. Other compounds with the ability to make hydrogen bonds are polar and can dissolve in water. They are called hydrophilic. Molecules possessing both characteristics are called amphiphilic.

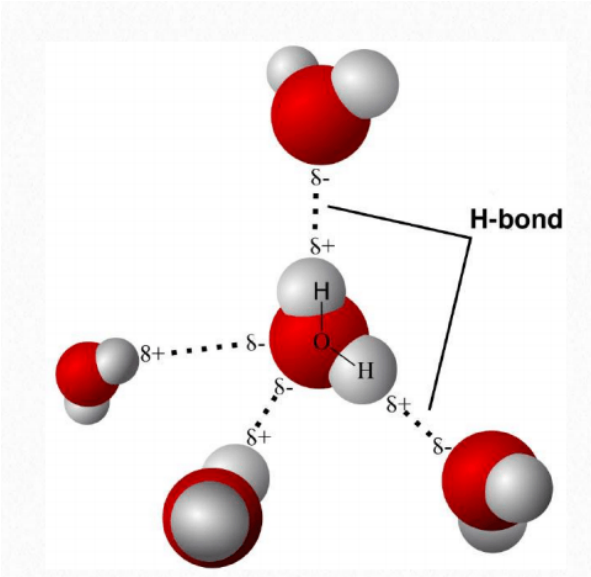


Figure 1.20 Hydrogen bonds (dotted lines) between water molecules Wikipedia

Weak interactions

Hydrogen bonds are one kind of electrostatic (i.e., based on charge) interaction between dipoles. Other forms of electrostatic interactions that are important in biochemistry include weak interactions between a polar molecule and a transient dipole, or between two temporary dipoles. These temporary dipoles result from the movement of electrons in a molecule. As electrons move around, the place where they are, at a given time, becomes temporarily more negatively charged and could now attract a temporary positive charge on another molecule. Since electrons don't stay put, these dipoles are very short-lived. Thus, the attraction that depends on these dipoles fluctuates and is very weak. Weak interactions like these are sometimes called van der Waals forces. Many molecular interactions in cells depend on weak interactions. Although the individual hydrogen bonds or other dipole-dipole interactions are weak, because of their large numbers, they can result in quite strong interactions between molecules.

Oxidation/reduction

Oxidation involves loss of electrons and reduction results in gain of electrons. For every biological oxidation, there is a corresponding reduction - one molecule loses electrons to another molecule. Oxidation reactions tend to release energy and are a source of bioenergy for chemotrophic cells.

Ionization

Ionization of biomolecules, by contrast does not involve oxidation/reduction. In ionization, a hydrogen ion (H+) leaves behind its electron as it exits (leaving behind a negative charge) or joins a group (adding a positive charge). Biological ionizations typically involve carboxyl groups or amines, though phosphates or sulfates can also be ionized. A carboxyl group can have two ionization states - a charge of -1 corresponds to the carboxyl without its proton and a charge of zero corresponds to the charge of the carboxyl with its proton on. An amine also has two ionization states. A charge of zero corresponds to a nitrogen with three covalent bonds (usually in the form of C-NH2) and a charge of +1 corresponds to a nitrogen making four covalent bonds (usually X-NH3 +).

Stereochemistry

A carbon has the ability to make four single bonds (forming a tetrahedral structure) and if it bonds to four different chemical groups, their atoms can be arranged around the carbon in two different ways, giving rise to stereochemical “handedness” (Figure 1.21). Each carbon with such a property is referred to as an asymmetric center. The property of handedness only occurs when a carbon has four different groups bonded to it. Enzymes have very specific 3-D structures, so for biological molecules that can exist in different stereoisomeric forms, an enzyme that synthesizes it would make only one of the possible isomers. By contrast, the same molecules made chemically (not using enzymes) end up with equal amounts of both isomers, called a **racemic mix**.

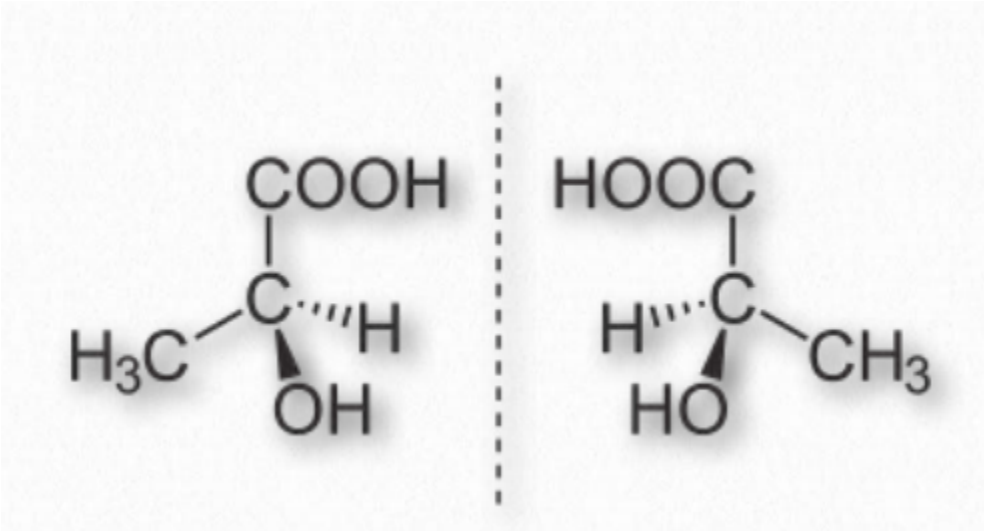


Figure 1.21 - Mirror images of lactic acid

Gibbs free energy

The Gibbs free energy calculation allows us to determine whether a reaction will be spontaneous, by taking into consideration two factors, change in enthalpy (ΔH) and change in entropy (ΔS). The free energy content of a system is given by the Gibbs free energy (G) and is equal to the enthalpy (H) for a process minus the absolute temperature (T) times the entropy (S)

$$G = H = TS \tag{1.2.1}$$

For a process, the change in the Gibbs free energy ΔG is given by

$$\Delta G = \Delta H - T\Delta S \tag{1.2.2}$$

A negative ΔG corresponds to release of free energy. Reactions that release energy are **exergonic**, whereas those that absorb energy are called **endergonic**.

The biological standard Gibbs free energy change (ΔG°’) corresponds to the ΔG for a process under standard conditions of temperature, pressure, and at pH = 7. For a reaction

$$aA + bB \rightleftharpoons cC + dD, \tag{1.2.3}$$

the equilibrium constant, K_{eq} is equal to

$$K_{eq} = \frac{[C]^c_{eq}[D]^d_{eq}}{[A]^a_{eq}[B]^b_{eq}} \tag{1.2.4}$$

where a, b, c, and d are integers in the balanced equation. Large values of K_{eq} correspond to favorable reactions (more C and D produced than A and B) and small values of K_{eq} mean the opposite. At equilibrium,

$$\Delta G^{o\prime} = -RT \ln K_{eq} \tag{1.2.5}$$

If a process has a $\Delta G = Z$ and a second process has a $\Delta G = Y$, then if the two processes are linked, ΔG and $\Delta G^{o\prime}$ values for the overall reaction will be the sum of the individual ΔG and ΔG°’ values.

$$\Delta G_{total} = \Delta G_1 + \Delta G_2 = Z + Y \tag{1.2.6}$$

$$\Delta G^{o\prime}_{total} = \Delta G^{o\prime}_1 + \Delta G^{o\prime}_2 \tag{1.2.7}$$

Catalysis

Catalysis is an increase in the rate of a reaction induced by a substance that is, itself, unchanged by the reaction. Because catalysts remain unchanged at the end of a reaction, a single catalyst molecule can be reused for many reaction cycles. Proteins that catalyze reactions in cells are called enzymes, while ribozymes are RNA molecules that act as catalysts.

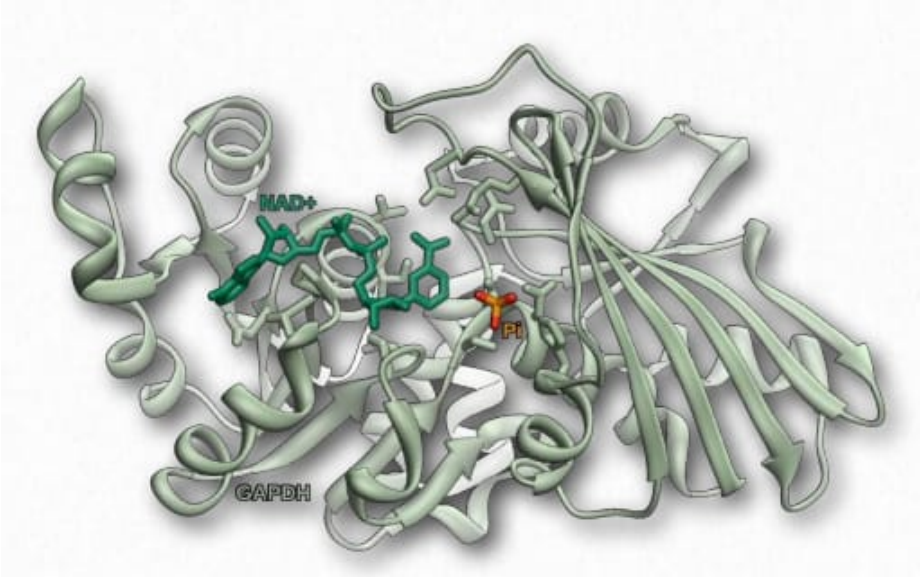


Figure 1.22 - Glyceraldehyde-3-phosphate dehydrogenase in the midst of catalysis Wikipedia

Contributors

This page titled 1.2: Introduction - Basic Chemistry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

1.3: Introduction - Water and Buffers

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



When it comes to water, we're literally drowning in it, as water is by far the most abundant component of every cell. To understand life, we begin the discussion with the basics of water, because everything that happens in cells, even reactions buried deep inside enzymes, away from water, is influenced by water's chemistry.

The water molecule has wide 'V' shape (the HO-H angle is 104°) with uneven sharing of electrons between the oxygen and the hydrogen atoms (Figure 1.23). Oxygen, with its higher electronegativity, holds electrons closer to itself than the hydrogens do. The hydrogens, as a result, are described as having a partial positive charge (typically designated as δ+) and the oxygen has a partial negative charge (written as δ-). Thus, water is a polar molecule because charges are distributed around it unevenly, not symmetrically.

Water as a solvent

Water (Figure 1.23) is described as a solvent because of its ability to solvate (dissolve) many, but not all, molecules. Molecules that are ionic or polar dissolve readily in water, but non-polar substances dissolve poorly in water, if at all. Oil, for example, which is non-polar, separates from water when mixed with it. On the other hand, sodium chloride, which ionizes, and ethanol, which is polar, are able to form hydrogen bonds, so both dissolve in water. Ethanol's solubility in water is crucial for brewers, winemakers, and distillers – but for this property, there would be no wine, beer or spirits. As explained in an earlier section, we use the term hydrophilic to describe substances that interact well with water and dissolve in it and the term hydrophobic to refer to materials that are non-polar and do not dissolve in water. Table 1.3 illustrates some polar and non-polar substances. A third term, amphiphilic, refers to compounds that have both properties. Soaps, for example are amphiphilic, containing a long, non-polar aliphatic tail and a head that ionizes.

Hydrophilic vs Hydrophobic Compounds

Hydrophobic	Hydrophilic
Nonpolar hydrocarbons (hexane)	Ionic compounds (NaCl)
Lipids (fats and cholesterol)	Polar organic compounds (alcohols, ketones or carbonyls)
	Weak acids (phosphates, amino acids)
	Sugars/carbohydrates

Table 1.3 Image by Aleia Kim

Solubility

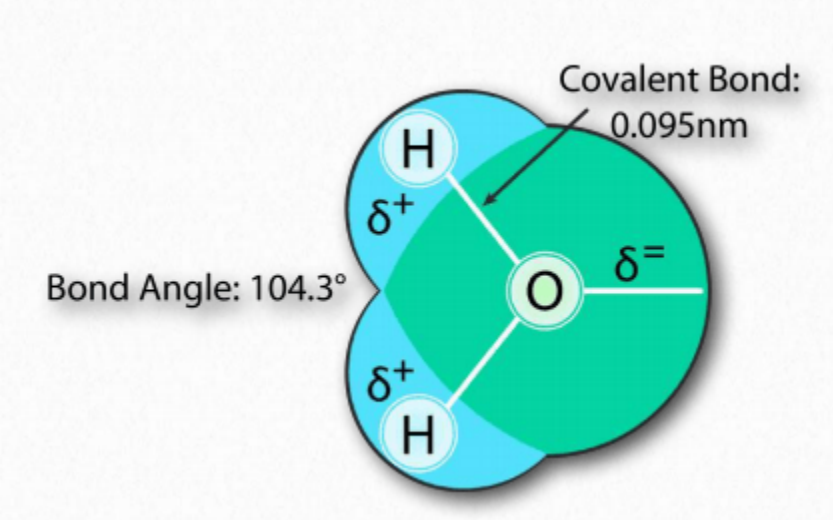


Figure 1.23 - Arrangement of atoms in water Image by Aleia Kim

The solubility of materials in water is based in free energy changes, as measured by ΔG. Remember, from chemistry, that H is the enthalpy (heat at constant pressure) and S is entropy. Given this,

$$\Delta G = \Delta H - T\Delta S \tag{1.3.1}$$

where T is the temperature in Kelvin. For a process to be favorable, the ΔG for it must be less than zero.

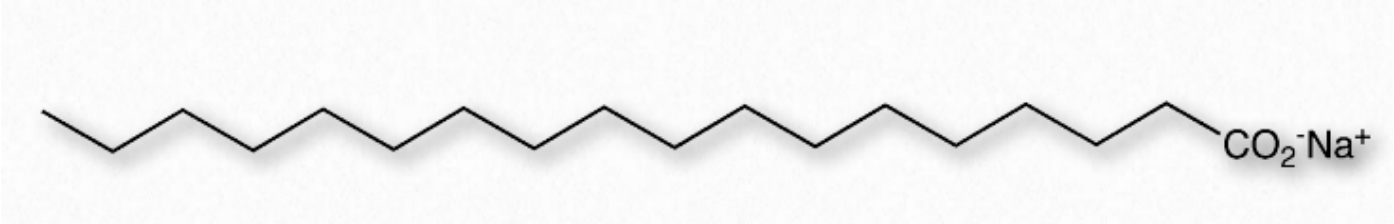


Figure 1.24 - Structure of a Soap

From the equation, lowered ΔG values will be favored with decreases in enthalpy and/or increases in entropy. Let us first consider why non-polar materials do not dissolve in water. We could imagine a situation where the process of dissolving involves the "surrounding" of each molecule of the nonpolar solute in water, just like each sodium and each chloride ion gets surrounded by water molecules as salt dissolves.

Water organization

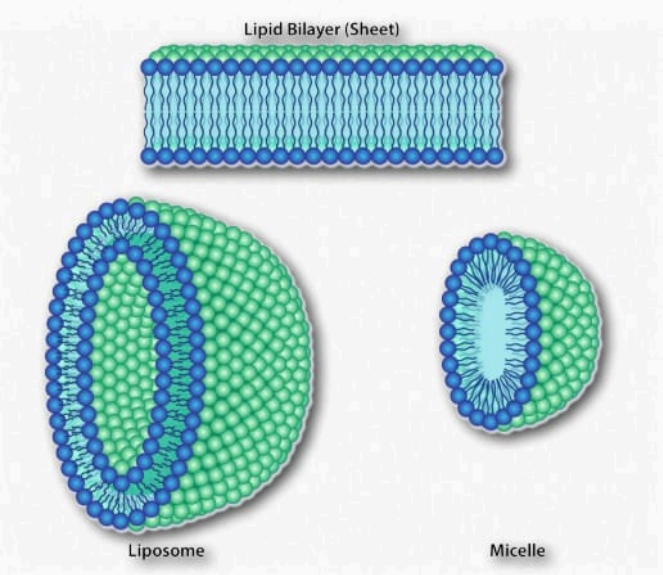


Figure 1.25 - Structures formed by amphiphilic substances in water. Image by Aleia Kim

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

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

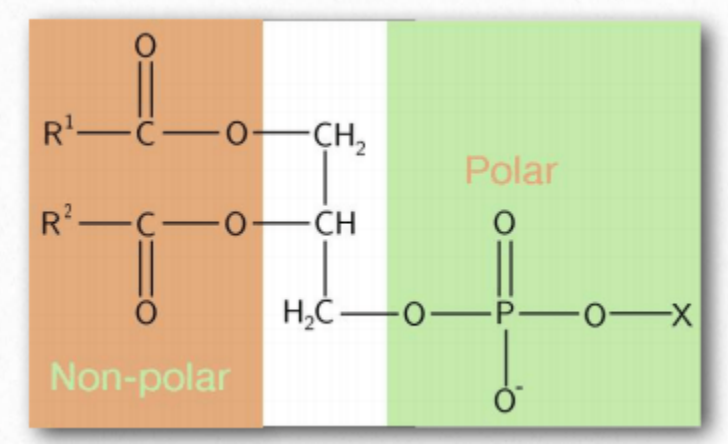


Figure 1.26 - A phospholipid - an amphiphilic substance

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



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

Amphiphilic substances

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

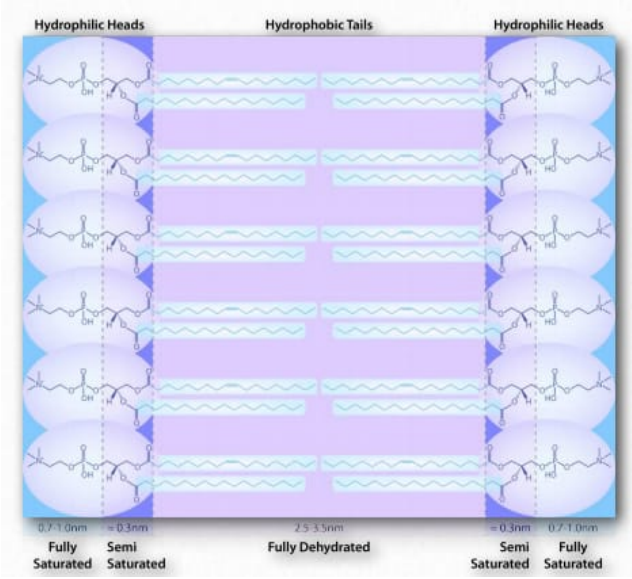


Figure 1.28 - Environment of a lipid bilayer. Water is concentrated away from the hydrophobic center, being saturated on the outside, semi-saturated near the head-tail junction and fully dehydrated in the middle. Image by Aleia Kim

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

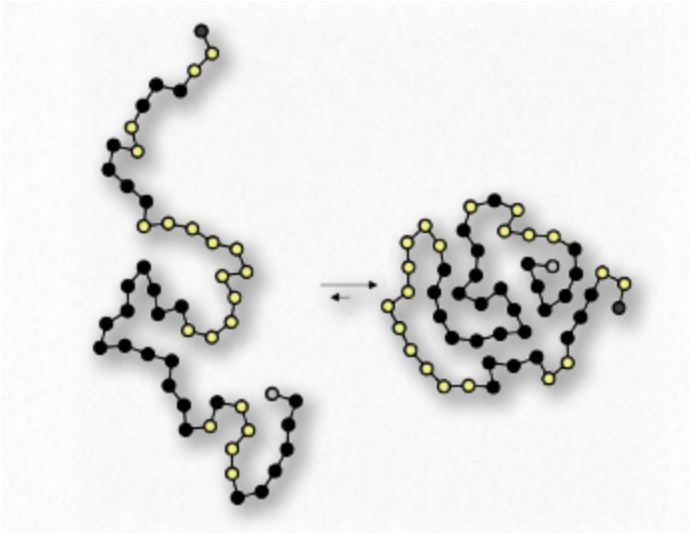


Figure 1.29 - Protein folding arranges hydrophobic amino acids (black dots) inside the protein

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

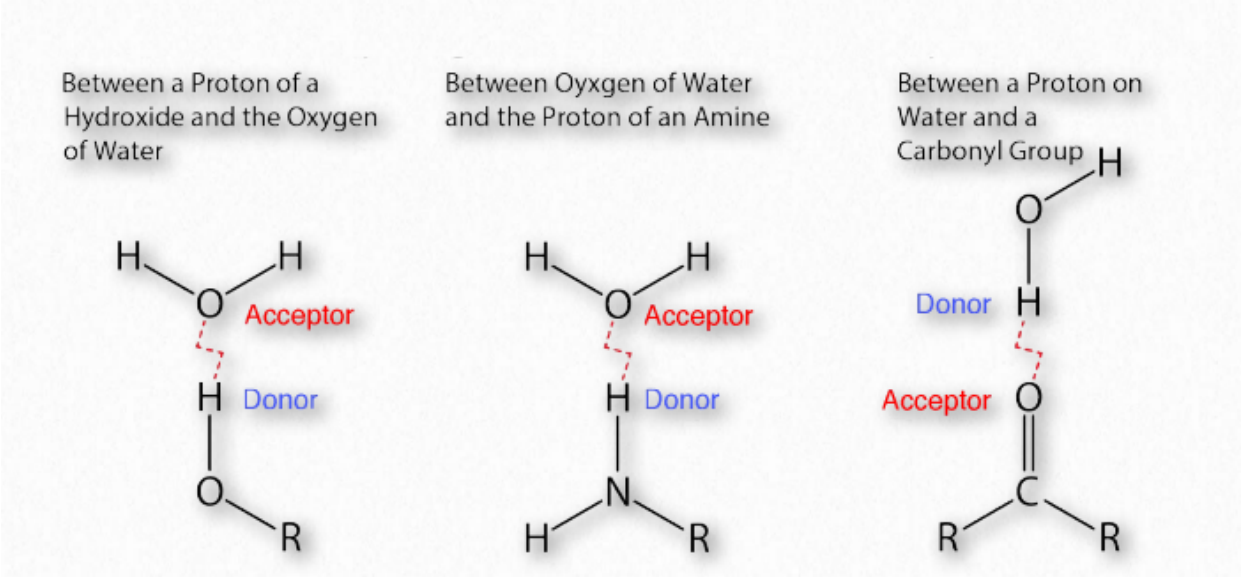


Figure 1.30 - Common hydrogen bonds in biochemistry Image by Aleia Kim

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

Hydrogen bonds

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

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

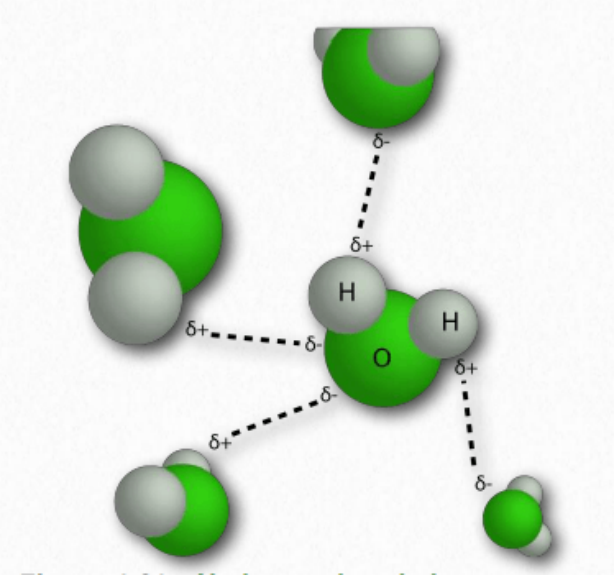


Figure 1.31 - Hydrogen bonds between water molecules Image by Pehr Jacobson

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

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

Partial Charges

The difference in electronegativity between hydrogen and the molecule to which it is covalently bound give rise to partial charges as described above. These tiny charges (δ+ and δ-) result in formation of hydrogen bonds, which occur when the partial positive charge of a hydrogen atom is attracted to the partial negative of another molecule. In water, that means the hydrogen of one water molecule is attracted to the oxygen of another (Figure 1.31). Since water is an asymmetrical molecule, it means also that the charges are asymmetrical. Such an uneven distribution is what makes a dipole. Dipolar molecules are important for interactions with other dipolar molecules and for dissolving ionic substances (Figure 1.32).

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

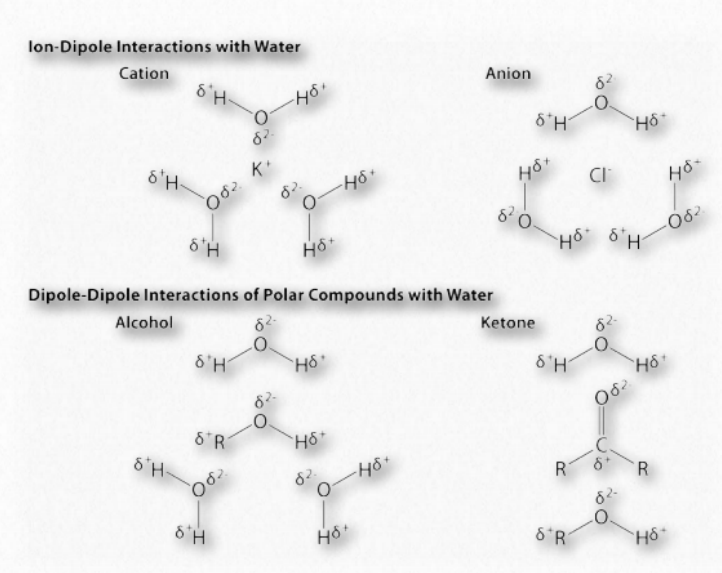


Figure 1.32 - Example dipole interactions in biochemistry Image by Aleia Kim

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

Bond Energies

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

Table 1.4 Image by Aleia Kim

Benefits of weak interactions

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

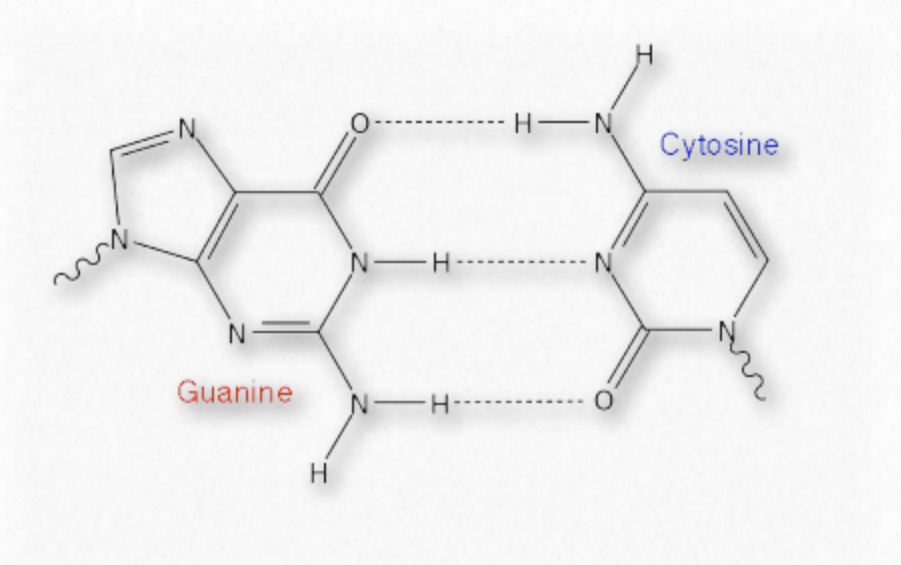


Figure 1.33 - Hydrogen bonds in a base pair of DNA Image by Aleia Kim

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

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

Weak Acid pKa Values

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

Image by Aleia Kim Table 1.5

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

The importance of buffers

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

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

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

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

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

$$pH + pOH = 14$$

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

Acids vs bases

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

Clearing Confusion - Students are often puzzled and expect that

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

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

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

Weak Acids

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

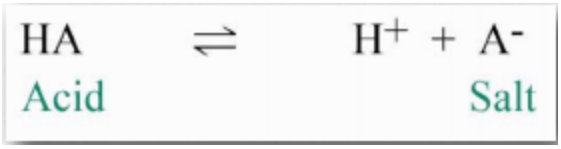


Figure 1.34 - Dissociation of a weak acid Image by Aleia Kim

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

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

Salt/Acid Ratio as a Function of pK_as

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

Image by Aleia Kim Table 1.6

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

UPS

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

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

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

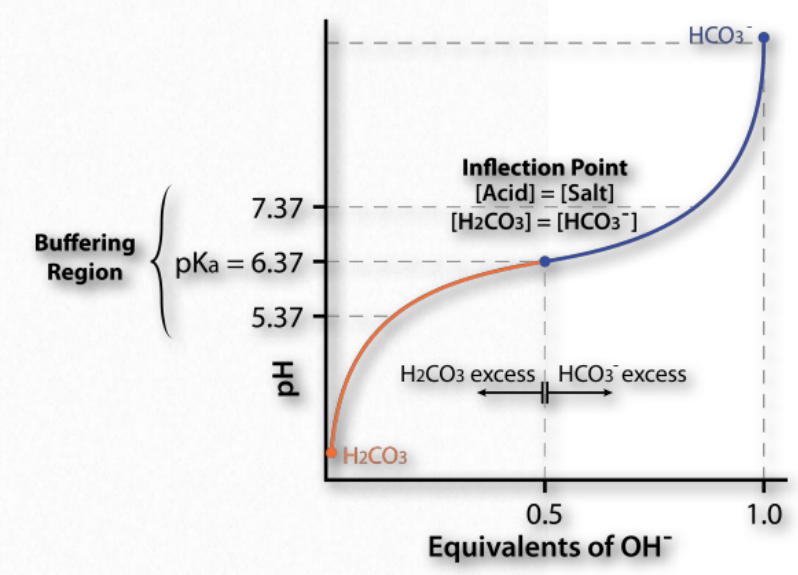


Figure 1.35 - Titration curve for carbonic acid Image by Aleia Kim

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

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

Henderson-Hasselbalch

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

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

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

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

just as

$$pH = -Log\ [H^+].$$

The Ka is the acid dissociation constant and is a measure of the strength of an acid. For a general acid, HA, which dissociates as

$$HA \rightleftarrows H^+ + A^-,\ K_a = [H^+][A^-]/[HA]$$

Thus, the stronger the acid, the more protons that will dissociate from it when added to water and the larger the value its Ka will have. Large values of Ka translate to lower values of pKa. As a result, the lower the pKa value is for a given acid, the stronger the weak acid is.

Constant pKa

Please note that pKa is a constant for a given acid. The pKa for carbonic acid is 6.37. By comparison, the pKa for formic acid is 3.75. Formic acid is therefore a stronger acid than acetic acid. A stronger acid will have more protons dissociated at a given pH than a weaker acid.

Now, how does this translate into stabilizing pH? Figure 1.35 shows a titration curve. In this curve, the titration begins with the conditions at the lower left (very low pH). At this pH, the H_2CO_3 form predominates, but as more and more OH- is added (moving to the 45 Why do we care about pH? Because biological molecules can, in some cases, be exquisitely sensitive to changes in it. As the pH of a solution changes, the charges of molecules in the solution can change, as you will see. Changing charges on biological molecules, especially proteins, can drastically affect how they work and even whether they work at all right), the pH goes up, the amount of HCO3- goes up and (correspondingly), the amount of H_2CO_3 goes down. Notice that the curve “flattens” near the pKa (6.37).

Buffering region

Flattening of the curve tells us is that the pH is not changing much (not going up as fast) as it did earlier when the same amount of hydroxide was added. The system is resisting a change in pH (not stopping the change, but slowing it) in the region of about one pH unit above and one pH unit below the pKa. Thus, the buffering region of the carbonic acid/ bicarbonate buffer is from about 5.37 to 7.37. It is maximally strong at a pH of 6.37.

Now it starts to become apparent how the buffer works. HA can donate protons when extras are needed (such as when OH- is added to the solution by the addition of NaOH). Similarly, A- can accept protons when extra H+ are added to the solution (adding HCl, for example). The maximum ability to donate or accept protons comes when

$$[A^-] = [HA]$$

This is consistent with the Henderson Hasselbalch equation and the titration curve. When $[A^-] = [HA]$, pH = 6.37 + Log(1). Since Log(1) = 0, pH = 6.37 = pKa for carbonic acid. Thus for any buffer, the buffer will have maximum strength and display flattening of its titration curve when $[A^-] = [HA]$ and when pH = pKa. If a buffer has more than one pKa (Figure 1.36), then each pKa region will display the behavior.

Buffered vs non-buffered

To understand how well a buffer protects against changes in pH, consider the effect of adding .01 moles of HCl to 1.0 liter of pure water (no volume change) at pH 7, compared to adding it to 1.0 liter of a 1M acetate buffer at pH 4.76. Since HCl completely dissociates, in 0.01M (10^{-2} M) HCl you will have 0.01M H+. For the pure water, the pH drops from 7.0 down to 2.0 (pH = -log(0.01M)).

By contrast, the acetate buffer’s pH after adding the same amount of HCl is 4.74. Thus, the pure water solution sees its pH fall from 7 to 2 (5 pH units), whereas the buffered solution saw its pH drop from 4.76 to 4.74 (0.02 pH units). Clearly, the buffer minimizes the impact of the added protons compared to the pure water.

Buffer capacity

It is important to note that buffers have capacities limited by their concentration. Let’s imagine that in the previous paragraph, we

had added the 0.01 moles HCl to an acetate buffer that had a concentration of 0.01M and equal amounts of Ac- and HAc. When we try to do the math in parallel to the previous calculation, we see that there are 0.01M protons, but only 0.005M A- to absorb them. We could imagine that 0.005M of the protons would be absorbed, but that would still leave 0.005M of protons unbuffered. Thus, the pH of this solution would be approximately

$$pH = -\log(0.005M) = 2.30$$

Exceeding buffer capacity dropped the pH significantly compared to adding the same amount of protons to a 1M acetate buffer. Consequently, when considering buffers, it is important to recognize that their concentration sets their limits. Another limit is the pH range in which one hopes to control proton concentration.

Multiple ionizable groups

Now, what happens if a molecule has two (or more) ionizable groups? It turns out, not surprisingly, that each group will have its own pKa and, as a consequence, will have multiple regions of buffering.

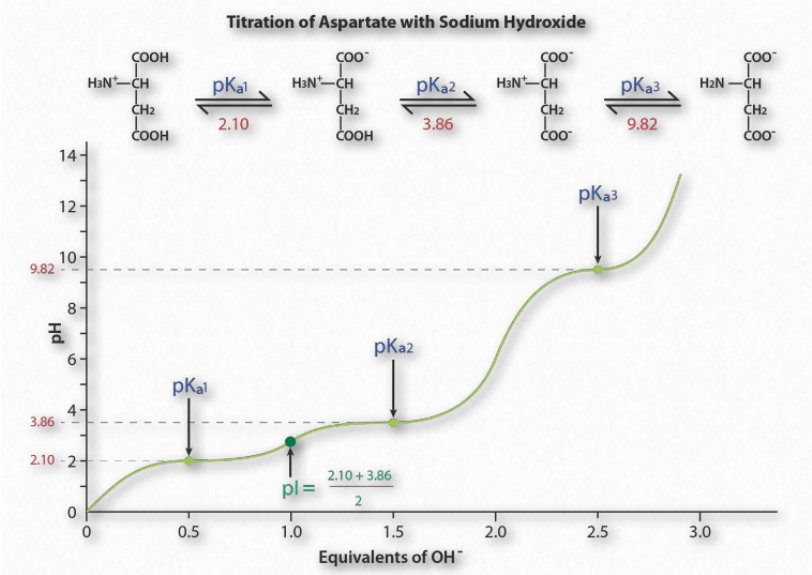


Figure 1.36 - Titration of an acidic amino acid Image by Aleia Kim

Figure 1.36 shows the titration curve for the amino acid aspartic acid. Note that in- stead of a single flattening of the curve, as was seen for acetic acid, aspartic acid's titration curve displays three such regions. These are individual buffering regions, each centered on the respective pKa values for the carboxyl group and the amine group.

Aspartic acid has four possible charges: +1 (α-carboxyl group, α-amino group, and Rgroup carboxyl each has a proton), 0 (α-carboxyl group missing proton, α- amino group has a proton, R-group carboxyl has a proton), -1 (α-carboxyl group and R-group carboxyl each lack a proton, α-amino group retains a proton), -2 (α-carboxyl, R-group carboxyl, and α-amino groups all lack extra proton).

Prediction

How does one predict the charge for an amino acid at a given pH? A good rule of thumb for estimating charge is that if the pH is more than one unit below the pKa for a group (carboxyl or amino), the proton is on. If the pH is more than one unit above the pKa for the group, the proton is off. If the pH is NOT more than one or less than one pH unit from the pKa, this simple assumption will not work.

Further, it is important to recognize that these rules of thumb are estimates only. The pI (pH at which the charge of a molecule is zero) is an exact value calculated as the average of the two pKa values on either side of the zero region. It is calculated at the average of the two pKa values around the point where the charge of the molecule is zero. For aspartic acid, this corresponds to $pK_{a1}$and pK_{a2}.

References

1. http://www.lpi.usra.edu/lunar/missions/apollo/ apollo_12/experiments/surveyor/
2. Arunan, Elangannan; Desiraju, Gautam R.; Klein, Roger A.; Sadlej, Joanna; Scheiner, Steve; Alkorta, Ibon; Clary, David C.; Crabtree, Robert H.; Dannenberg, Joseph J.; Hobza, Pavel; Kjaergaard, Henrik G.; Legon, Anthony C.; Mennucci, Benedetta; Nesbitt, David J. (2011). "Definition of the hydrogen bond". Pure Appl. Chem. 83 (8): 1637–1641. doi:10.1351/PAC-REC-10-01-02

This page titled 1.3: Introduction - Water and Buffers is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2: STRUCTURE AND FUNCTION

- Topic hierarchy

2.1: Prelude to Structure and Function
2.4: Structure and Function- Proteins II
2.5: Structure and Function- Protein Function II
2.6: Structure and Function - Nucleic Acids
2.7: Structure and Function- Carbohydrates
2.8: Structure and Function - Lipids and Membranes
2.2: Structure & Function - Amino Acids
2.3: Structure & Function- Proteins I

This page titled 2: Structure and Function is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.1: Prelude to Structure and Function

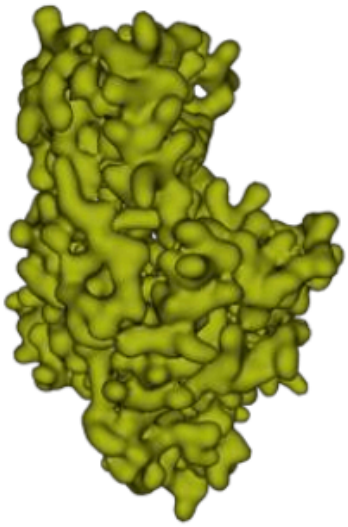
Source: BiochemFFA_2_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

"*The man who does not read good books has no advantage over the man who cannot read them.*" Mark Twain



In this chapter, we will examine the structures of the major classes of biomolecules, with an eye to understanding how these structures relate to function.

As noted earlier, water is the most abundant molecule in cells, and provides the aqueous environment in which cellular chemistry happens. Dissolved in this water are inorganic ions like sodium, potassium and calcium. But the distinctiveness of biochemistry derives from the vast numbers of complex, large, carbon compounds, that are made by living cells. You have probably learned that the major classes of biological molecules are proteins, nucleic acids, carbohydrates and lipids. The first three of these major groups are macromolecules that are built as long polymers made up of smaller subunits or monomers, like strings of beads. The lipids, while not chains of monomers, also have smaller subunits that are assembled in various ways to make the lipid components of cells, including membranes. The chemical properties and three dimensional conformations of these molecules determine all the molecular interactions upon which life depends. Whether building structures within cells, transferring information, or catalyzing reactions, the activities of biomolecules are governed by their structures. The properties and shapes of macromolecules, in turn, depend on the subunits of which they are built.



Interactive 2.1: *The enzyme Hexokinase: as for all enzymes, the activity of hexokinase depends on its structure. Protein Database (PDB)*

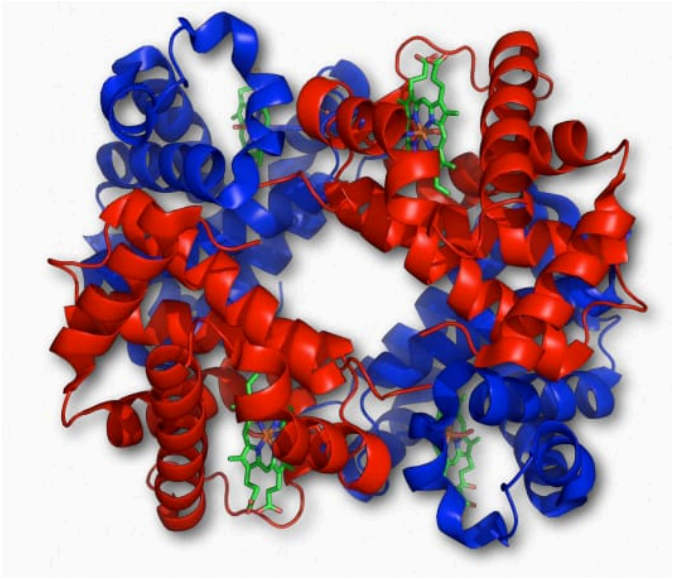
We will next examine the major groups of biological macromolecules: proteins, polysaccharides, nucleic acids, and lipids. The building blocks of the first three, respectively, are amino acids, monosaccharides (sugars), and nucleotides. Acetyl-CoA is the most common building block of lipids.

This page titled 2.1: Prelude to Structure and Function is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.4: Structure and Function- Proteins II

Source: BiochemFFA_2_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

In this section, we hope to bring to life the connection between structure and function of proteins. So far, we have described notable features of the four elements (primary, secondary, tertiary, and quaternary) of protein structure and discussed example proteins/motifs exhibiting them. In this section, we will examine from a functional perspective a few proteins/domains whose function relies on secondary, tertiary, or quaternary structure. It is, of course a bit of a narrow focus to ascribe protein function to any one component of structure, but our hope is by presenting these examples, we can bring to life the way in which a protein's secondary, tertiary, and quarternary structure lead to the functions it has.



Hemoglobin Wikipedia

Fibrous proteins - secondary structure

Proteins whose cellular or extracellular roles have a strong structural component are composed primarily of primary and second structure, with little folding of the chains. Thus, they have very little tertiary structure and are fibrous in nature. Proteins exhibiting these traits are commonly insoluble in water and are referred to as fibrous proteins (also called scleroproteins). The examples described in this category are found exclusively in animals where they serve roles in flesh, connective tissues and hardened external structures, such as hair. They also contain the three common fibrous protein structures α -helices (keratins), β-strands/ sheets (fibroin & elastin) and triple helices (collagen). The fibrous proteins have some commonality of amino acid sequence. Each possesses an abundance of repeating sequences of amino acids with small, non-reactive side groups. Many contain short repeats of sequences, often with glycine.

Keratins



Figure 2.56 - The horns of an impala are composed of keratin Wikipedia

The keratins are a family of related animal proteins that take numerous forms. α-keratins are structural components of the outer layer of human skin and are integral to hair, nails, claws, feathers, beaks, scales, and hooves. Keratins provide strength to tissues, such as the tongue, and over 50 different keratins are encoded in the human genome. At a cellular level, keratins comprise the intermediate filaments of the cytoskeleton. α- keratins primarily contain α-helices, but can also have β-strand/sheet structures. Individual α-helices are often intertwined to form coils of coiled structures and these strands can also be further joined together by disulfide bonds, increasing structural strength considerably. This is particularly relevant for α-keratin in hair, which contains about 14% cysteine. The odor of burned hair and that of the chemicals used to curl/uncurl hair (breaking/re-making disulfide bonds) arise from their sulfurous components. β-keratins are comprised of β-sheets, as their name implies.

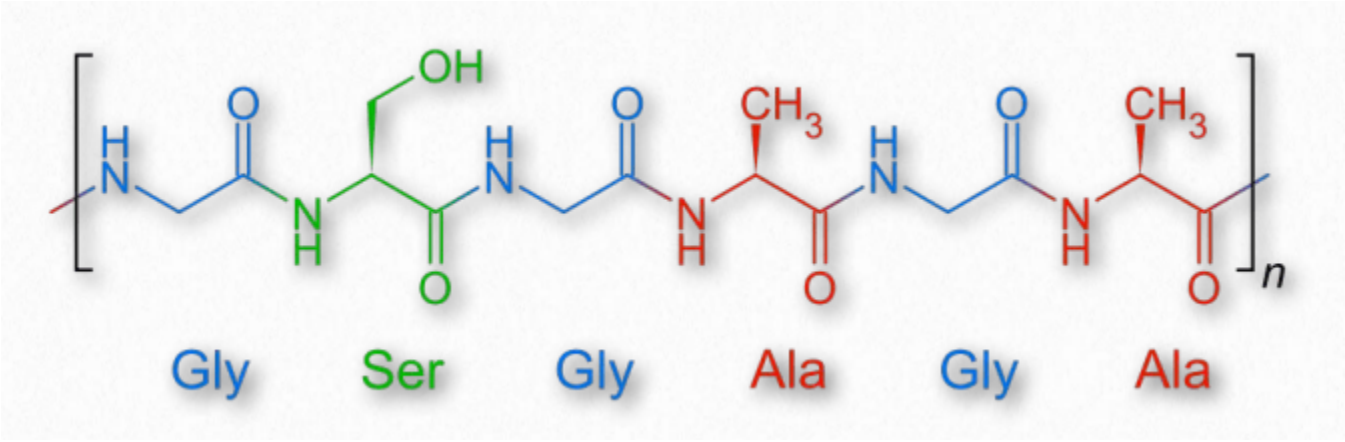


Figure 2.57 - The repeating amino acid sequence of fibroin

Fibroin

An insoluble fibrous protein that is a component of the silk of spiders and the larvae of moths and other insects, fibroin is comprised of antiparallel β-strands tightly packed together to form β- sheets. The primary structure of fibroin is a short repeating sequence with glycine at every other residue (Figure 2.57). The small R-groups of the glycine and alanine in the repeating sequence allows for the tight packing characteristic of the fibers of silk. Wikipedia link HERE Elastin As suggested by its name, elastin is a protein with elastic characteristics that functions in many tissues of the body to allow them to resume their shapes after expanding or contracting. The protein is rich in glycine and proline and can comprise over 50% of the weight of dry, defatted arteries.



Figure 2.58 - Weaving of a silk sari Wikipedia

Elastin

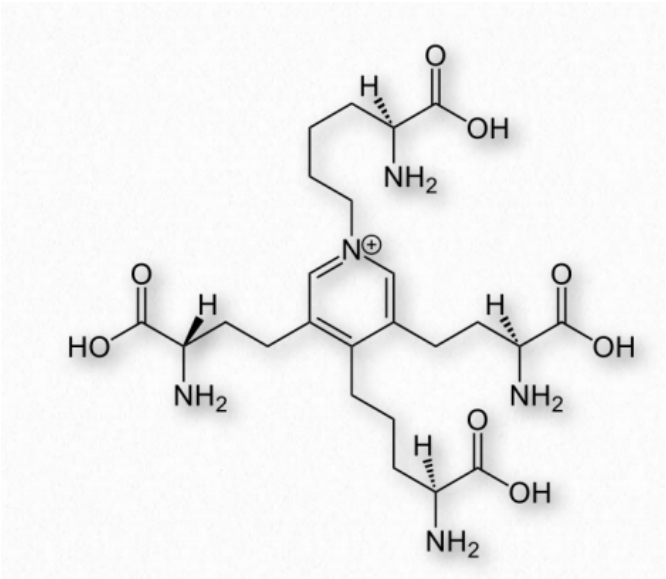


Figure 2.59 - Desmosine Wikipedia

is made by linking tropoelastin proteins together through lysine residues to make a durable complex crosslinked by desmosine. In arteries, elastin helps with pressure wave propagation for facilitating blood flow.

Collagen

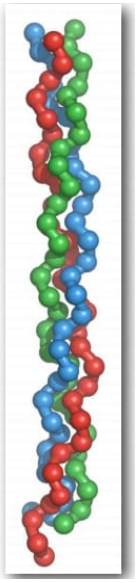


Figure 2.60 - Collagen's triple helix Wikipedia

Collagen is the most abundant protein in mammals, occupying up to a third of the total mass. There are at least 16 types of collagen. Its fibers are a major component of tendons and they are also found abundantly in skin. Collagen is also prominent in cornea, cartilage, bone, blood vessels and the gut.

Collagen's structure is an example of a helix of helices, being composed of three lefthanded helical chains that each are coiled together in a right-handed fashion to make the collagen fiber (Figure 2.60). Each helix is stretched out more than an α-helix, giving it an extended appearance. On the inside of the triple helical structure, only residues of glycine are found, since the side chains of other amino acids are too bulky. Collagen chains have the repeating structure glycinem-n where m is often proline and n is often hydroxyproline (Figure 2.61).

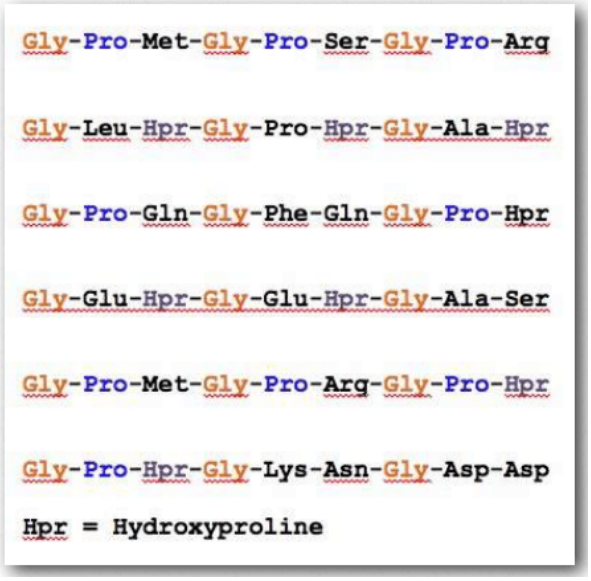


Figure 2.61 - Repeating sequences in collagen

Collagen is synthesized in a pre-procollagen form. Processing of the pre-procollagen in the endoplasmic reticulum results in glycosylation, removal of the 'pre' sequence, and hydroxylation of lysine and proline residues (see below). The hydroxides can form covalent cross-links with each other, strengthening the collagen fibers. As pro-collagen is exported out of the cell, proteases trim it, resulting in a final form of collagen called tropocollagen.

Hydroxylation

Hydroxylation of proline and lysine side chains occurs post-translationally in a reaction catalyzed by prolyl-4-hydroxylase and lysyl-hydroxylase (lysyl oxidase), respectively. The reaction requires vitamin C. Since hydroxylation of these residues is essential for formation of stable triple helices at body temperature, vitamin C deficiency results in weak, unstable collagen and, consequently, weakened connective tissues. It is the cause of the disease known as scurvy. Hydrolyzed collagen is used to make gelatin, which is important in the food industry. collagens. Wikipedia link HERE

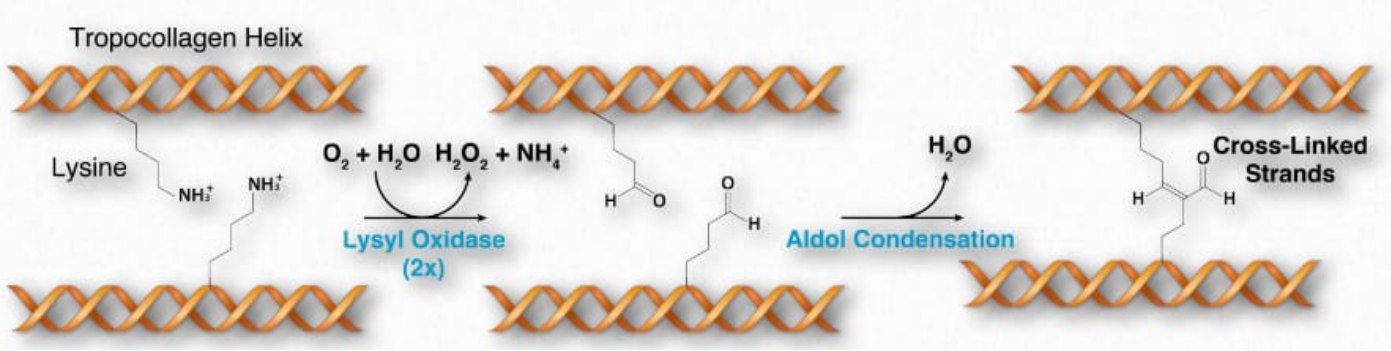


Figure 2.62 - Oxidation and cross-linking of lysine residues in tropocollagen. Only two strands of the triple helix are shown for simplicity Image by Aleia Kim

Lamins

Lamins are fibrous proteins that provide structure in the cell nucleus and play a role in transcription regulation. They are similar to proteins making up the intermediate filaments, but have extra amino acids in one coil of the protein. Lamins help to form the nuclear lamin in the interior of the nuclear envelope and play important roles in assembling and disassembling the latter in the process of mitosis. They also help to position nuclear pores. In the process of mitosis, disassembly of the nuclear envelope is promoted by phosphorylation of lamins by a protein called mitosis promoting factor and assembly is favored by reversing the reaction (dephosphorylation).

Structural domains - tertiary structure

Every globular protein relies on its tertiary structure to perform its function, so rather than trying to find representative proteins for tertiary structure (an almost impossible task!), we focus here on a few elements of tertiary structure that are common to many proteins. These are the structural domains and they differ from the structural motifs of supersecondary structure by being larger (25-500 amino acids), having a conserved amino acid sequence, and a history of evolving and functioning independently of the protein chains they are found in. Structural domains are fundamental units of tertiary structure and are found in more than one protein. A structural domain is selfstabilizing and often folds independently of the rest of the protein chain.

Leucine zipper

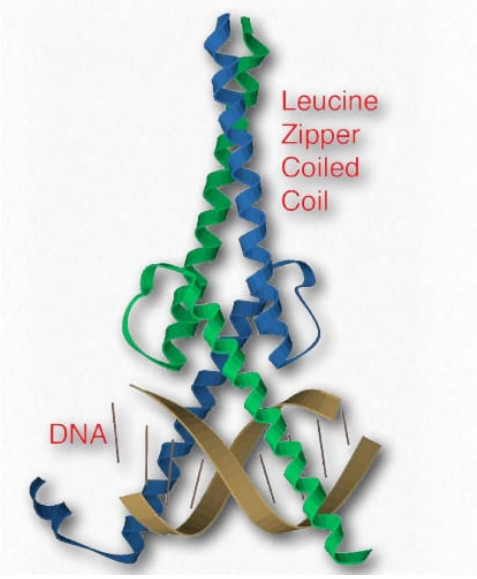


Figure 2.63 - Leucine zipper bound to DNA Wikipedia

A common feature of many eukaryotic DNA binding proteins, leucine zippers are characterized by a repeating set of leucine residues in a protein that interact like a zipper to favor dimerization. Another part of the domain has amino acids (commonly arginine and lysine) that allow it to interact with the DNA double helix (Figure 2.63). Transcription factors that contain leucine zippers include Jun-B, CREB, and AP-1 fos/ jun.

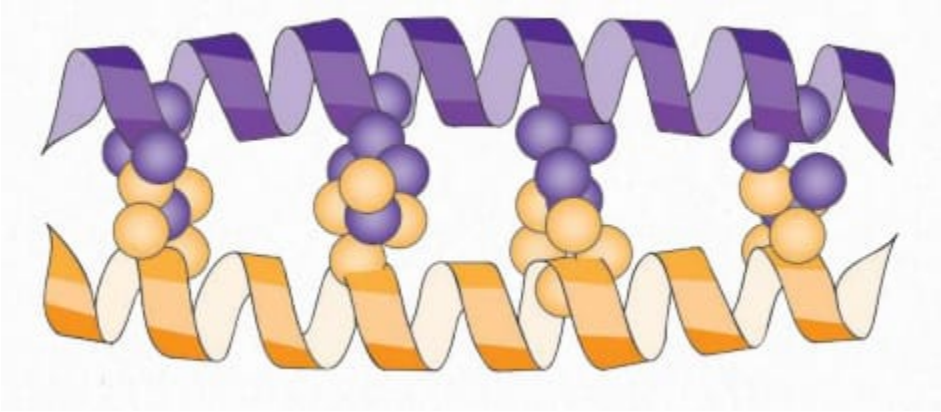


Figure 2.64 - Leucine zipper structure. Leucines are indicated by orange and purple balls. Image by Penelope Irving

Zinc fingers

The shortest structural domains are the zinc fingers, which get their name from the fact that one or more coordinated zinc ions stabilize their finger-like structure. Despite their name, some zinc fingers do not bind zinc. There are many structural domains classified as zinc fingers and these are grouped into different families. Zinc fingers were first identified as components of DNA binding transcription factors, but others are now known to bind RNA, protein, and even lipid structures. Cysteine and histidine side chains commonly play roles in coordinating the zinc.

Src SH_2 domain

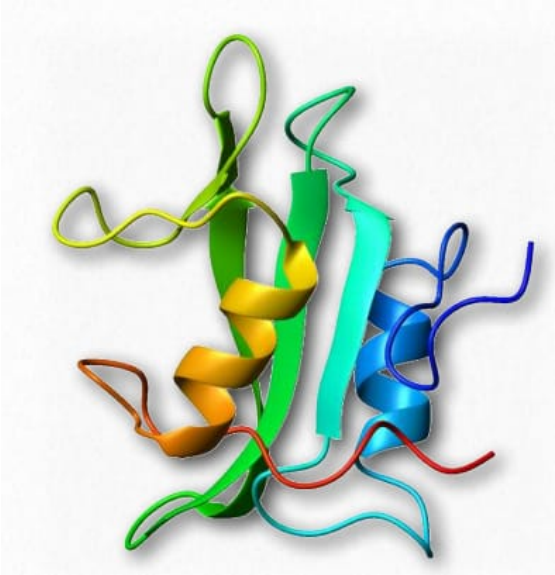


Figure 2.65 - SH2 Domain Wikipedia

The Src oncoprotein contains a conserved SH_2structural domain that recognizes and binds phosphorylated tyrosine side chains in other proteins (Figure 2.65). Phosphorylation is a fundamental activity in signaling and phosphorylation of tyrosine and interaction between proteins carrying signals is critically needed for cellular communication. The SH_2domain is found in over 100 human proteins.

Helix-turn-helix domain



Figure 2.66 - Helix-Turn-Helix Domain of a Protein Bound to DNA Wikipedia

Helix-turn-helix is a common domain found in DNA binding proteins, consisting of two α-helices separated by a small number of amino acids. As seen in Figure 2.66, the helix parts of the structural domain interact with the bases in the major groove of DNA. Individual α-helices in a protein are part of a helix-turn-helix structure, where the turn separates the individual helices.

Pleckstrin homology domain

Pleckstrin Homology (PH) domains are protein domains with important functions in the process of signaling. This arises partly from the affinity for binding phosphorylated inositides, such as PIP2 and PIP3, found in Figure 2.66 - Helix-Turn-Helix Domain of a Protein Bound to DNA Wikipedia Figure 2.65 - SH2 Domain Wikipedia biological membranes. PH domains can also bind to G-proteins and protein kinase C. The domain spans about amino acids and is found in numerous signaling proteins.

These include Akt/Rac Serine/ Threonine Protein Kinases, Btk/ltk/Tec tyrosine protein kinases, insulin receptor substrate (IRS-1), Phosphatidylinositolspecific phospholipase C, and several yeast proteins involved in cell cycle regulation.

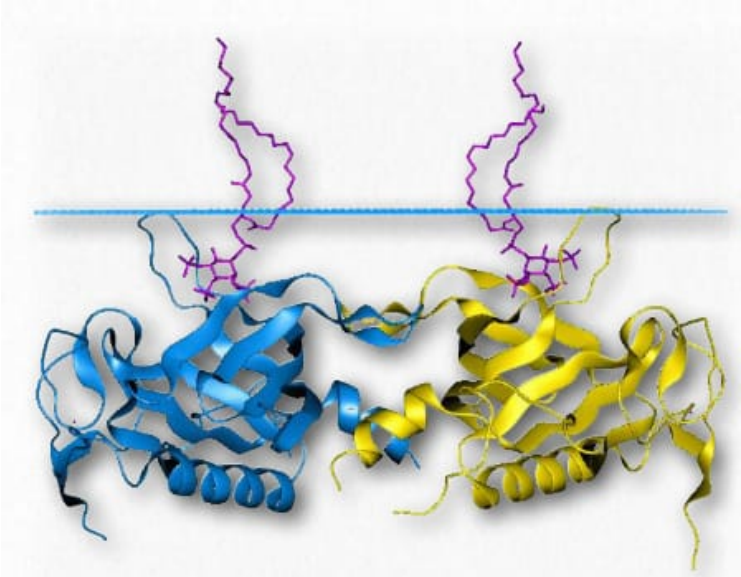


Figure 2.67 - Pleckstrin homology domain of Btk tyrosine protein kinase. The protein is embedded in a membrane (above blue line) Wikipedia

Structural globular proteins

Enzymes catalyze reactions and proteins such as hemoglobin perform important specialized functions. Evolutionary selection has reduced and eliminated waste so that we can be sure every protein in a cell has a function, even though in some cases we may not know what it is. Sometimes the structure of the proFigure 2.68 - Relationship of basement membrane to epithelium, endothelium, and connective tissue tein is its primary function because the structure provides stability, organization, connections other important properties. It is with this in mind that we present the following proteins.

Basement membrane

The basement membrane is a layered extracellular matrix of tissue comprised of protein fibers (type IV collagen) and glycosaminoglycans that separates the epithelium from other tissues (Figure 2.68). More importantly, the basement membrane acts like a glue to hold tissues together. The skin, for example, is anchored to the rest of the body by the basement membrane. Basement membranes provide an interface of interaction between cells and the environment around them, thus facilitating signaling processes. They play roles in differentiation during embryogenesis and also in maintenance of function in adult organisms.

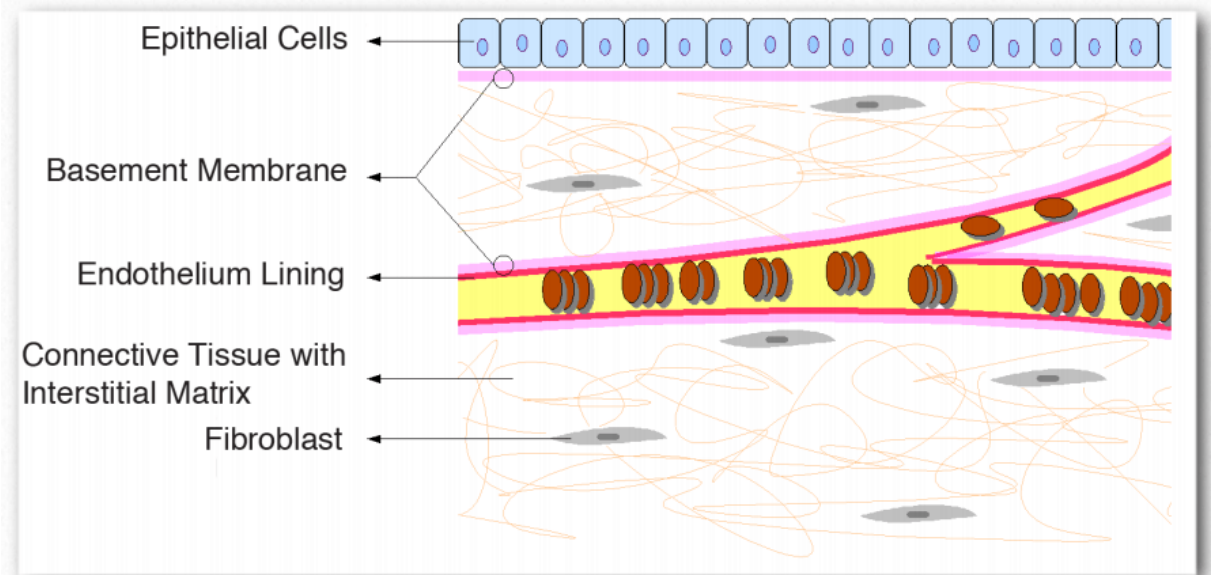


Figure 2.68 - Relationship of basement membrane to epithelium, endothelium, and connective tissue

Actin

Actin is the most abundant globular protein found in most types of eukaryotic cells, comprising as much as 20% of the weight of muscle cells. Similar proteins have been identified in bacteria (MreB) and archaeons (Ta0583). Actin is a monomeric subunit able to polymerize readily into two different types of filaments. Microfilaments are major component of the cytoskeleton and are acted on by myosin in the contraction of muscle cells (See HERE). Actin will be discussed in more detail in the next section HERE.

Intermediate Filaments

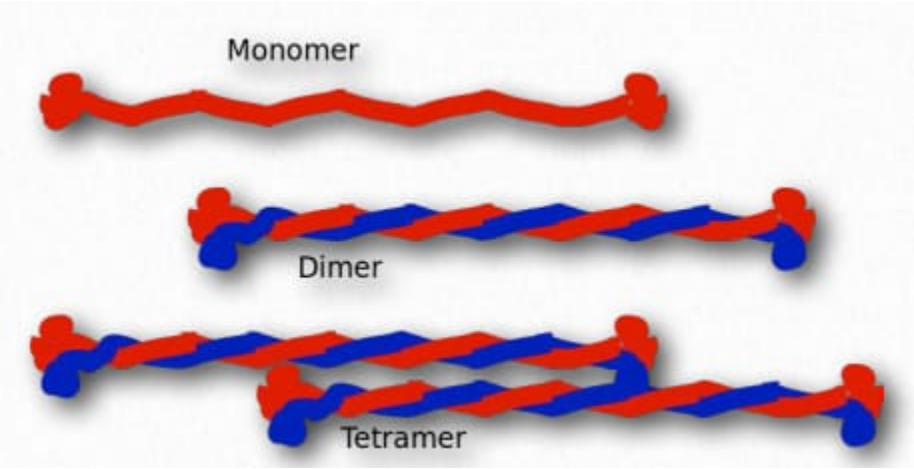


Figure 2.69 - Assembly of intermediate filaments

Intermediate filaments are a part of the cytoskeleton in many animal cells and are comprised of over 70 different proteins. They are called intermediate because their size (average diameter = 10 nm) is between that of the microfilaments (7 nm) and the microtubules (25 nm).

The intermediate filament components include fibrous proteins, such as the keratins and the lamins, which are nuclear, as well as cytoplasmic forms. Intermediate filaments give flexibility to cells because of their own physical properties. They can, for example, be stretched to several times of their original length.

Six types

There are six different types of intermediate filaments. Type I and II are acidic or basic and attract each other to make larger filaments. They include epithelial keratins and trichocytic keratins (hair components). Type III proteins include four structural

proteins - desmin, GFAP (glial fibrillary acidic protein), peripherin, and vimentin. Type IV also is a grouping of three proteins and one multiprotein structure (neurofilaments). The three proteins are α-internexin, synemin, and syncoilin. Type V intermediate filaments encompass the lamins, which give structure to the nucleus. Phosphorylation of lamins leads to their disassembly and this is important in the process of mitosis. The Type VI category includes only a single protein known as nestin.

Tubulin

A third type of filament found in cells is that of the microbutules. Comprised of a polymer of two units of a globular protein called tubulin, microtubules provide “rails” for motor proteins to move organelles and other “cargo” from one part of a cell to another. Microtubules and tubulin are discussed in more detail HERE.

Vimentin

Vimentin (Figure 2.70) is the most widely distributed protein of the intermediate filaments. It is expressed in fibroblasts, leukocytes, and blood vessel endothelial cells. The protein has a significant role maintaining the position of organelles in the cytoplasm, with attachments to the nucleus, mitochondria, and endoplasmic reticulum (Figure 2.70). Vimentin provides elasticity to cells and resilience that does not arise from the microtubules or microfilaments. Wounded mice that lack the vimentin gene survive, but take longer to heal wounds than wild type mice. Vimentin also controls the movement of cholesterol from lysosomes to the site of esterification. The result is a reduction in the amount of cholesterol stored inside of cells and has implications for adrenal cells, which must have esters of cholesterol.

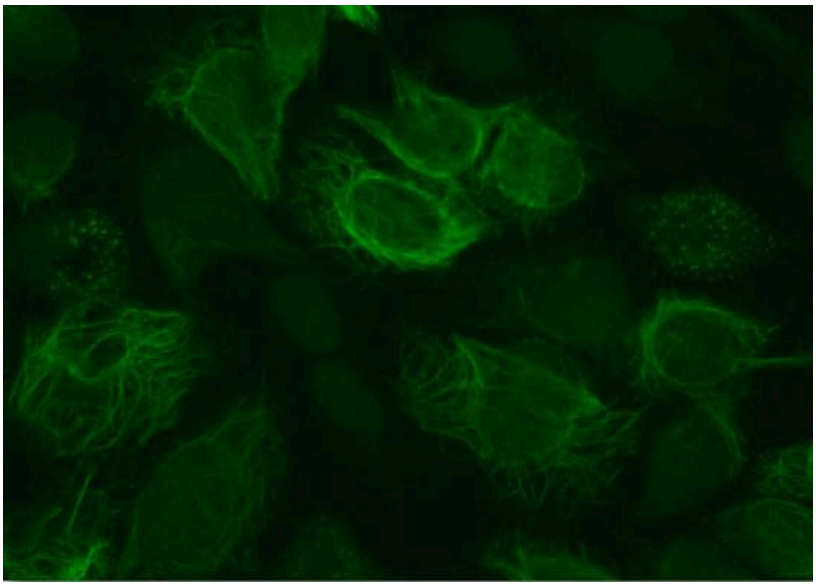


Figure 2.70 - Vimentin in cells Wikipedia

Mucin

Mucins are a group of proteins found in animal epithelial tissue that have many glycosyl residues on them and typically are of high molecular weight (1 to 10 million Da). They are gel-like in their character and are often used for lubrication. Mucus is comprised of mucins. In addition to lubrication, mucins also help to control mineralization, such as bone formation in vertebrate organisms and calcification in echinoderms. They also play roles in the immune system by helping to bind pathogens. Mucins are commonly secreted onto mucosal surfaces (nostrils, eyes, mouth, ears, stomach, genitals, anus) or into fluids, such as saliva. Because of their extensive mucosylation, mucins hold a considerable amount of water (giving them the “slimy” feel) and are resistant to proteolysis.

Vinculin

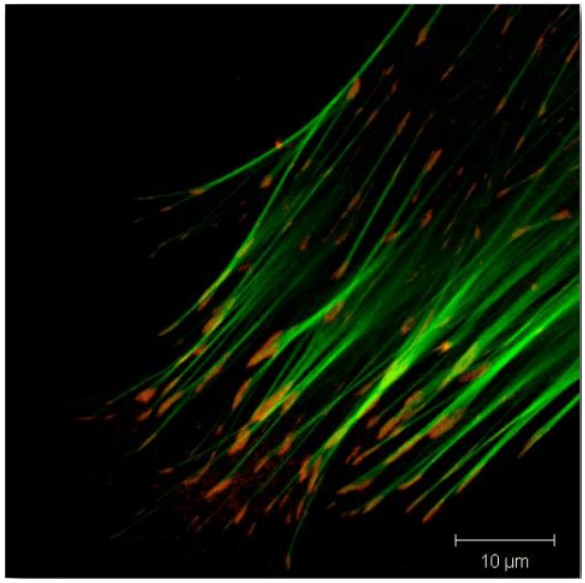


Figure 2.71 - Actin filaments (green) attached to vinculin in focal adhesion (red) Wikipedia

Vinculin (Figure 2.72) is a membrane cytoskeletal protein found in the focal adhesion structures of mammalian cells. It is found at cell-cell and cell-matrix junctions and interacts with integrins, talin, paxillins and F-actin. Vinculin is thought to assist (along with other proteins) in anchoring actin microfilaments to the membrane (Figure 2.71). Binding of vinculin to actin and to talin is regulating by polyphosphoinositides and can be inhibited by acidic phospholipids.

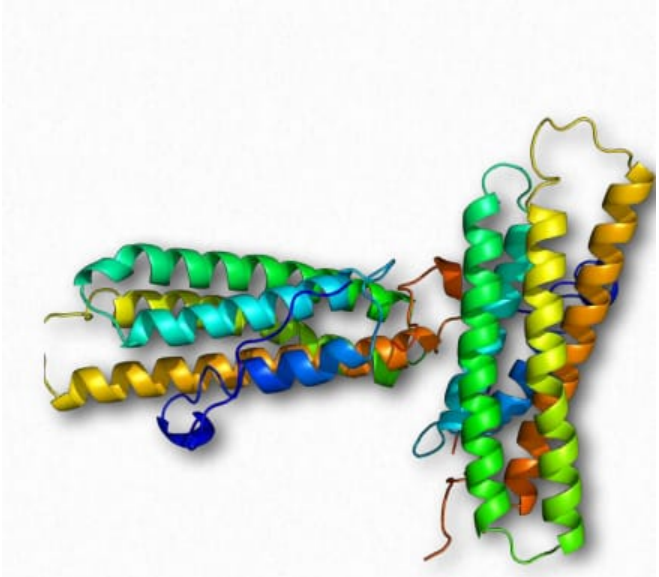


Figure 2.72 - Vinculin Wikipedia

Syndecans

Syndecans are transmembrane proteins that make a single pass with a long amino acid chain (24-25 residues) through plasma membranes and facilitate G proteincoupled receptors' interaction with Figure 2.71 - Actin filaments (green) attached to vinculin in focal adhesion (red) Wikipedia ligands, such as growth factors, fibronectin, collagens (I, III, and IV) and antithrombin-1. Syndecans typically have 3-5 heparan sulfate and chondroitin sulfate chains attached to them.

Heparan sulfate can be cleaved at the site of a wound and stimulate action of fibroblast growth factor in the healing process. The role of syndecans in cell-cell adhesion is shown in mutant cells lacking syndecan I that do not adhere well to each other. Syndecan 4 is also known to adhere to integrin. Syndecans can also inhibit the spread of tumors by the ability of the syndecan 1 ectodomain to suppress growth of tumor cells without affecting normal epithelial cells.

Defensin

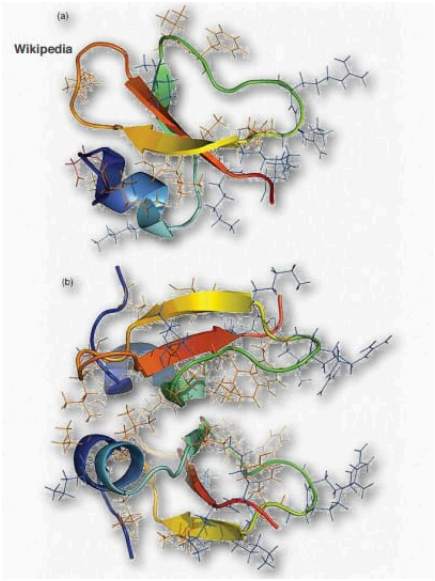


Figure 2.73 - Defensin monomer (top) and dimer (bottom) - Cationic residues in blue, hydrophobic residues in orange, and anionic residues in red Wikipedia

Defensins (Figure 2.73) are a group of small cationic proteins (rich in cysteine residues) that serve as host defense peptides in vertebrate and invertebrate organisms. They protect against infection by various bacteria, fungi, and viruses. Defensins contain between 18 and 45 amino acids with (typically) about 6- 8 cysteine residues. In the immune system, defensins help to kill bacteria engulfed by phagocytosis by epithelial cells and neutrophils. They kill 120 Figure 2.72 - Vinculin Wikipedia bacteria by acting like ionophores - binding the membrane and opening pore-like structures to release ions and nutrients from the cells.

Focal adhesions

In the cell, focal adhesions are structures containing multiple proteins that mechanically link cytoskeletal structures (actin bundles) with the extracellular matrix. They are dynamic, with proteins bringing and leaving with signals regarding the cell cycle, cell motility, and more almost constantly. Focal adhesions serve as anchors and as a signaling hub at cellular locations where integrins bind molecules and where membrane clustering events occur. Over 100 different proteins are found in focal adhesions.

Focal adhesions communicate important messages to cells, acting as sensors to update information about the status of the extracellular matrix, which, in turn, adjusts/ affects their actions. In sedentary cells, they are stabler than in cells in motion because when cells move, focal adhesion contacts are established at the “front” and removed at the rear as motion progresses. This can be very important in white blood cells’ ability to find tissue damage.

Ankyrin

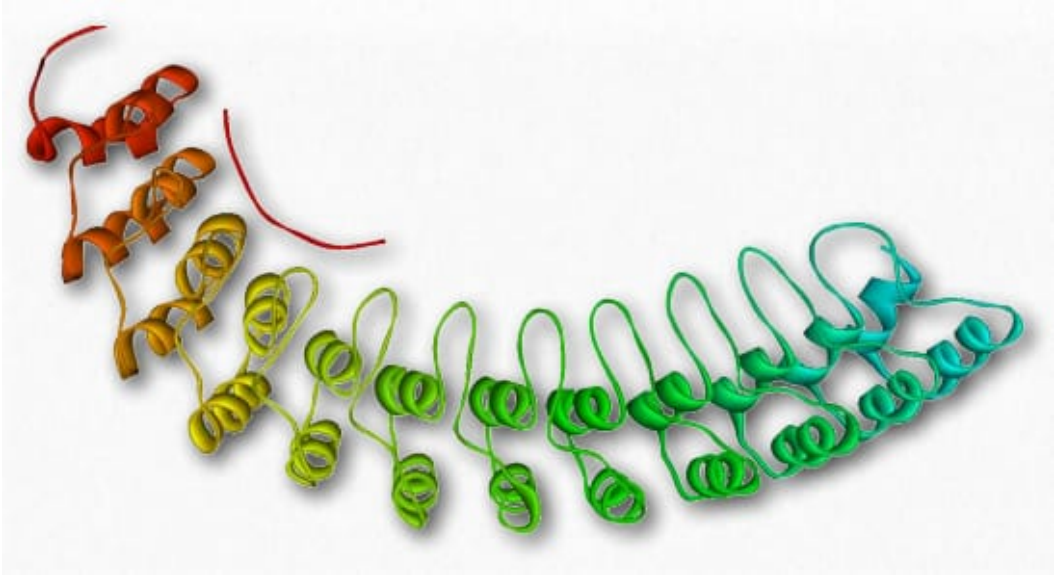


Figure 2.74 - Ankyrin's membrane-binding domain

Ankyrins (Figure 2.74) are a family of membrane adaptor proteins serving as "anchors" to interconnect integral membrane proteins to the spectrin-actin membrane cytoskeleton. Ankyrins are anchored to the plasma membrane by covalently linked palmitoyl-CoA. They bind to the β subunit of spectrin and at least a dozen groups of integral membrane proteins. The ankyrin proteins contain four functional domains: an N-terminal region with 24 tandem ankyrin repeats, a central spectrin-binding domain, a "death domain" interacting with apoptotic proteins, and a C-terminal regulatory domain that is highly varied significantly among different ankyrins.

Spectrin

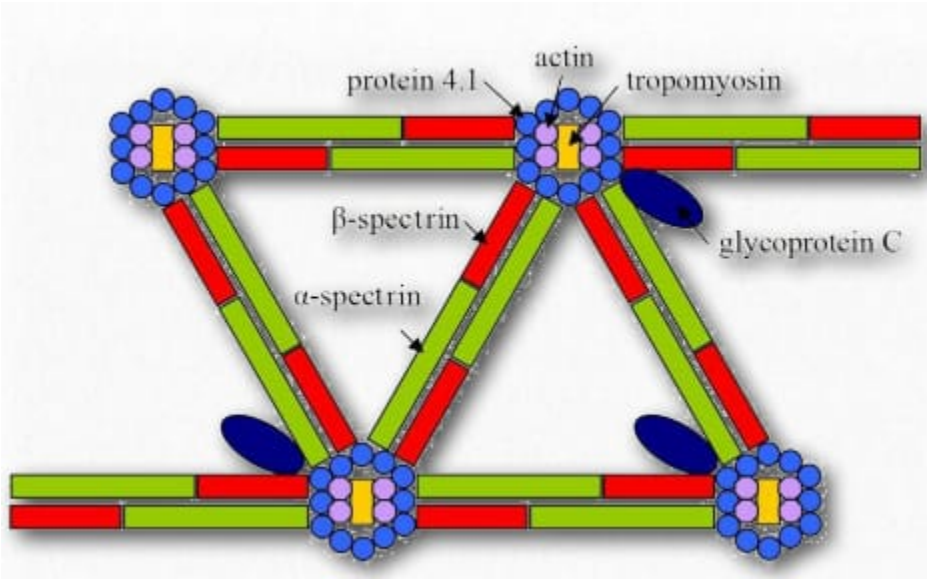


Figure 2.75 - Spectrin and other proteins in the cytoskeleton Wikipedia

Spectrin (Figures 2.75 & 2.76) is a protein of the cellular cytoskeleton that plays an important role in maintaining its structure and the integrity of the plasma membrane. In animals, spectrin gives red blood cells their shape. Spectrin is located inside the inner layer of the eukaryotic plasma membrane where it forms a network of pentagonal or hexagonal arrangements.

Spectrin fibers collect together at junctional complexes of actin and is also attached to ankyrin for stability, as well as numerous integral membrane proteins, such as glycophorin.

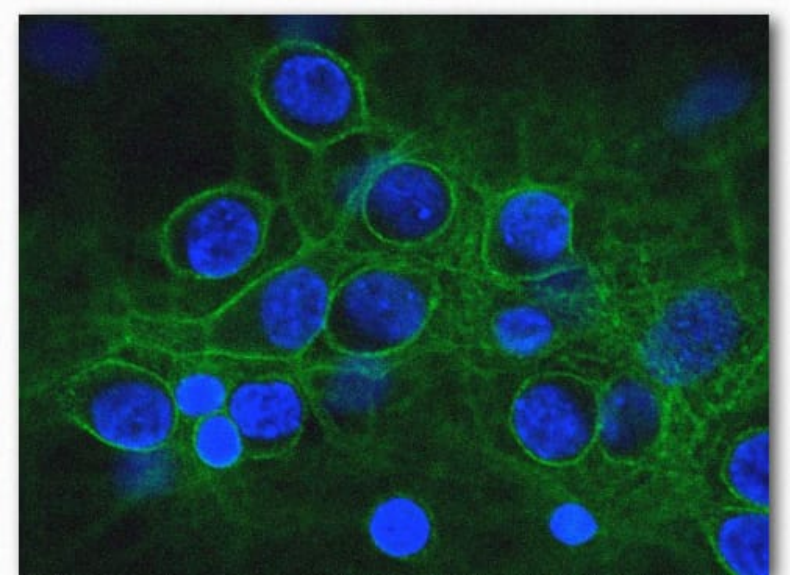


Figure 2.76 - Spectrin (green) and nuclei (Blue) Wikipedia

Integrins

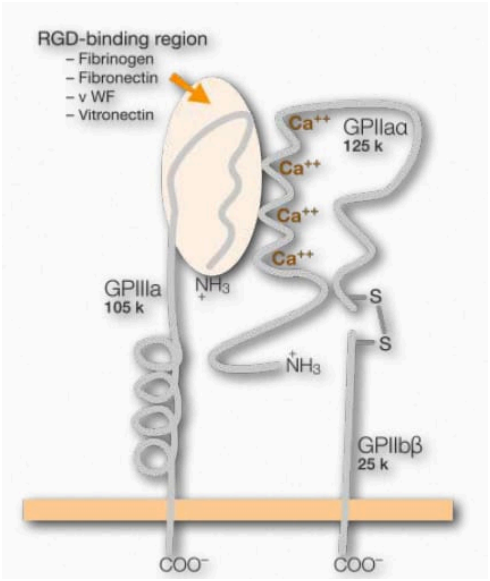


Figure 2.77 - Integrin and its binding site (on top left) Wikipedia

In multicellular organisms, cells need connections, both to each other and to the extracellular matrix. Facilitating these attachments at the cellular end are transmembrane proteins known as integrins (Figure 2.77). Integrins are found in all metazoan cells. Ligands for the integrins include collagen, fibronectin, laminin, and vitronectin. Integrins function not only in attachment, but also in communication, cell migration, virus linkages (adenovirus, for example), and blood clotting. Integrins are able to sense chemical and mechanical signals about the extracellular matrix and move that information to intracellular domains as part of the process of signal transduction. Inside the cells, responses to the signals affect cell shape, regulation of the cell cycle, movement, or changes in other cell receptors in the membrane. The process is dynamic and allows for rapid responses as may be necessary, for example in the process of blood clotting, where the integrin known as GPIbIIIa (on the surface of blood platelets) attaches to fibrin in a clot as it develops.

Integrins work along with other receptors, including immunoglobulins, other cell adhesion molecules, cadherins, selectins, and syndecans. In mammals the proteins have a large number of subunits - 18 α- and 8 β-chains. They are a bridge between its links outside the cell to the extracellular matrix (ECM) and its links inside the cell to the cytoskeleton. Integrins play central role in formation and stability of focal adhesions. These are large molecular complexes arising from clustering of integrin-ECM connections. In the process of cellular movement, integrins at the “front” of the cell (in the direction of the movement), make new

attachments to substrate and release connections to substrate in the back of the cell. These latter integrins are then endocytosed and reused.

Integrins also help to modulate signal transduction through tyrosine kinase receptors in the cell membrane by regulating movement of adapters to the plasma membrane. β1c integrin, for example, recruits the Shp2 phosphatase to the insulin growth factor receptor to cause it to become dephosphorylated, thus turning off the signal it communicates. Integrins can also help to recruit signaling molecules inside of the cell to activated tyrosine kinases to help them to communicate their signals.

Cadherins

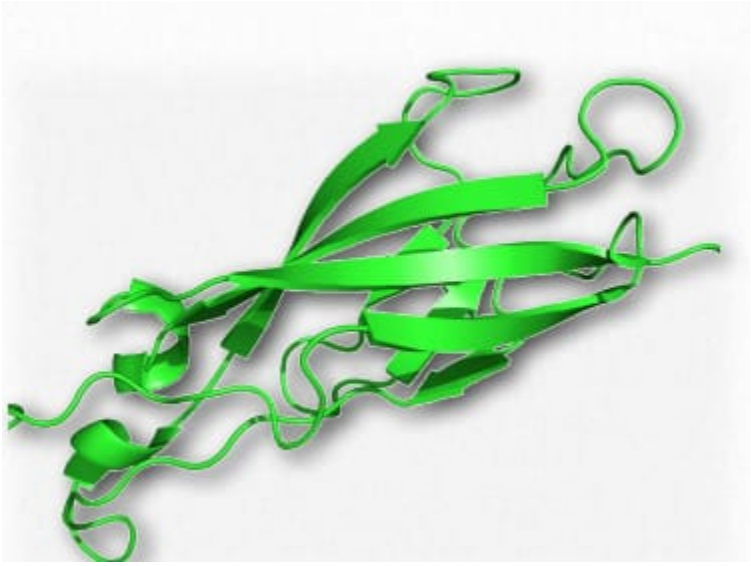


Figure 2.78 - Extracellular ectodomain of a cadherin

Cadherins (Figure 2.78) constitute a type-1 class of transmembrane proteins playing important roles in cell adhesion. They require calcium ions to function, forming adherens junctions that hold tissues together (See Figure 2.69). Cells of a specific cadherin type will preferentially cluster with each other in preference to associating with cells containing a different cadherin type. Caderins are both receptors and places for ligands to attach. They assist in the proper positioning of cells in development, separation of different tissue layers, and cell migration.

Selectins



Figure 2.79 - Selectin bound to a sugar Wikipedia

Selectins (Figure 2.79) are cell adhesion glycoproteins that bind to sugar molecules. As such, they are a type of lectin - proteins that bind sugar polymers (see HERE also). All selectins have an N-terminal calcium-dependent lectin domain, a single transmembrane domain, and an intracellular cytoplasmic tail.

There are three different types of selectins, 1) E-selectin (endothelial); 2) L (lymphocytic; and 3) P (platelets and endothelial cells. Selectins function in lymphocyte homing (adhesion of blood lymphocytes to cells in lymphoid organs), in inflammation processes, and in cancer metastasis. Near the site of inflammation, P-selectin on the surface of blood capillary cells interacts with glycoproteins on leukocyte cell surfaces. This has the effect of slowing the movement of the leukocyte. At the target site of inflammation, E- selectin on the endothelial cells of the blood vessel and L-selectin on the surface of the leukocyte bind to their respective carbohydrates, stopping the leukocyte movement. The leukocyte then crosses the wall of the capillary and begins the immune response. Selectins are involved in the inflammatory processes of asthma, psoriasis, multiple scleroris, and rheumatoid arthritis.

Laminins

Laminins are extracellular matrix glycoproteins that a major components of the basal lamina and affect cell differentiation, migration, and adhesion. They are secreted into the extracellular matrix where they are incorporated and are essential for tissue maintenance and survival. When laminins are defective, muscles may not form properly and give rise to muscular dystrophy.

Laminins are associated with fibronectin, entactin, and perlecan proteins in type IV collagen networks and bind to integrin receptors in the plasma membrane. As a consequence, laminins contribute to cellular attachment, differentiation, shape, and movement. The proteins are trimeric in structure, having one α-chain, a β-chain, and a γ-chain. Fifteen combinations of different chains are known.

Vitronectin

Vitronectin is a glycoprotein (75kDa) found in blood serum (platelets), the extracellular matrix, and in bone. It promotes the process of cell adhesion and spreading and binds to several protease inhibitors (serpins). It is secreted from cells and is believed to play roles in blood clotting and the malignancy of tumors. One domain of vitronectin binds to plasminogen activator inhibitor and acts to stabilize it. Another domain of the protein binds to cellular integrin proteins, such as the vitronectin receptor that anchors cells to the extracellular matrix.

Catenins

Catenins are a family of proteins interacting with cadherin proteins in cell adhesion (Figure 2.69). Four main types of catenins are known, α-, β-, γ-, and δ-catenin. Catenins play roles in cellular organization before development occurs and help to regulate cellular growth. α-catenin and β-catenin are found at adherens junctions with cadherin and help cells to maintain epithelial layers. Cadherins are connected to actin filaments of the cytoskeleton and catenins play the critical role. Catenins are important for the process whereby cellular division is inhibited when cells come in contact with each other (contact inhibition).

When catenin genes are mutated, cadherin cell adhesions can disappear and tumorigenesis may result. Catenins have been found to be associated with colorectal and numerous other forms of cancer.

Glycophorins

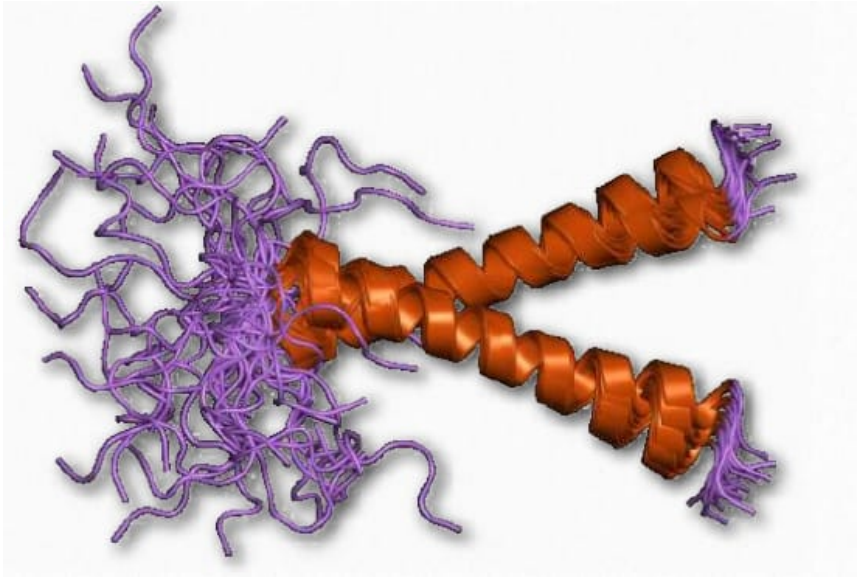


Figure 2.80 - Glycophorin a

All of the membrane proteins described so far are notable for the connections they make to other proteins and cellular structures. Some membrane proteins, though, are designed to reduce cellular connections to proteins of other cells. This is particularly important for blood cells where “stickiness” is undesirable except where clotting is concerned.

Glycophorins (Figure 2.80) are membrane-spanning sialoglycoproteins of red blood cells. They are heavily glycosylated (60%).and rich in sialic acid, giving the cells a very hydrophilic (and negatively charged) coat, which enables them to circulate in the bloodstream without adhering to other cells or the vessel walls.

Five glycophorins have been identified - four (A,B,C,and D) from isolated membranes and a fifth form (E) from coding in the human genome. The proteins are abundant, forming about 2% of the total membrane proteins in these cells. Glycophorins have important roles in regulating RBC membrane mechanical properties and shape. Because some glycophorins can be expressed in various nonerythroid tissues (particularly Glycophorin C), the importance of their interactions with the membrane skeleton may have a considerable biological significance.

Cooperativity and allosterism - quaternary structure

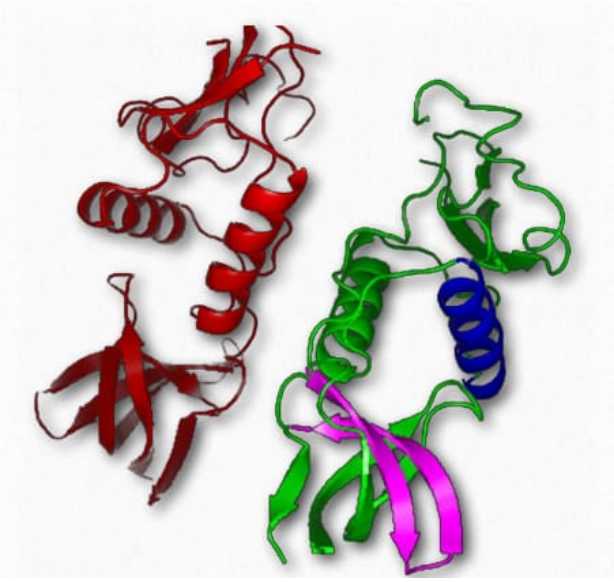


Figure 2.81 - Two polypeptide units of a protein interact in quaternary structure Wikipedia

Quaternary structure, of course describes the interactions of individual subunits of a multi-subunit protein (Figure 2.81). The result

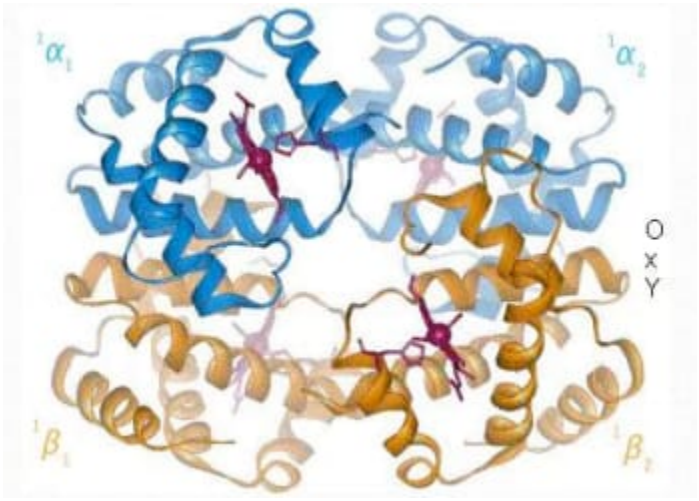
of these interactions can give rise to important biological phenomena, such as cooperative binding of substrates to a protein and allosteric effects on the action of an enzyme.

Allosteric effects can occur by a series of mechanisms, but a common feature is that binding of an effector to an enzyme subunit causes (or locks) the enzyme in either a Tstate (less activity) or an R-state (more activity). Effectors can be enzyme substrates (homotropic effectors) or non-substrates (heterotropic effectors). Allosterism will be covered in more depth in the Catalysis chapter HERE.

We begin our consideration of quaternary structure with a discussion of cooperativity, how it arises in the multi-subunit protein hemoglobin and how its properties contrast with those of the related, single subunit protein myoglobin.

Cooperativity

Cooperativity is defined as the phenomenon where binding of one ligand molecule by a protein favors the binding of additional molecules of the same type. Hemoglobin, for example, exhibits cooperativity when the binding of an oxygen molecule by the iron of the heme group in one of the four subunits causes a slight conformation change in the subunit. This happens because the heme iron is attached to a histidine side chain and binding of oxygen 'lifts' the iron along with the histidine ring (also known as the imidazole ring).



Movie 2.3 - Hemoglobin's structural changes on binding oxygen Wikipedia

Since each hemoglobin subunit interacts with and influences the other subunits, they too are induced to change shape slightly when the first subunit binds to oxygen (a transition described as going from the T-state to the R-state). These shape changes favor each of the remaining subunits binding oxygen, as well. This is very important in the lungs where oxygen is picked up by hemoglobin, because the binding of the first oxygen molecule facilitates the rapid uptake of more oxygen molecules. In the tissues, where the oxygen concentration is lower, the oxygen leaves hemoglobin and the proteins flips from the R-state back to the Tstate.

CO_2 transport

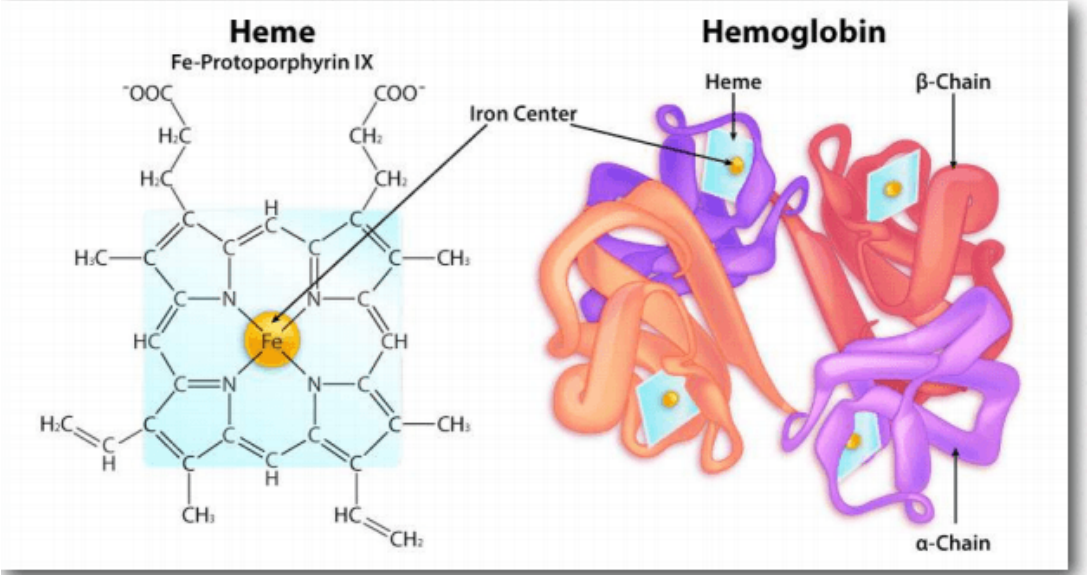


Figure 2.82 - Heme structure within hemoglobin Image by Aleia Kim

Cooperativity is only one of many fascinating structural aspects of hemoglobin that help the body to receive oxygen where it is needed and pick it up where it is abundant. Hemoglobin also assists in the transport of the product of cellular respiration (carbon dioxide) from the tissues producing it to the lungs where it is exhaled. Like the binding of oxygen to hemoglobin, binding of other molecules to hemoglobin affects its affinity for oxygen. The effect is particularly pronounced when comparing the oxygen binding characteristics of hemoglobin's four subunits with the oxygen binding of the related protein myoglobin's single subunit (Figure 2.83).

Different oxygen binding

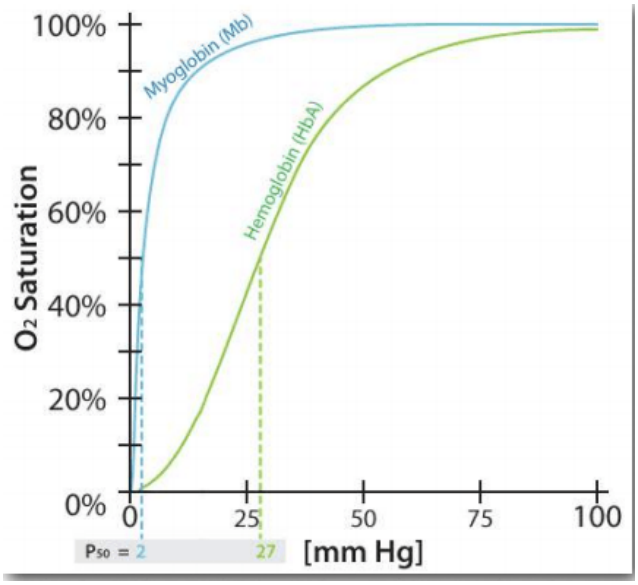


Figure by Aleia Kim

Like hemoglobin, myoglobin contains an iron in a heme group that binds to oxygen. The structure of the globin protein in myoglobin is very similar to the structure of the globins in hemoglobin and hemoglobin is thought to have evolved from myoglobin in evolutionary history. As seen in Figure 2.83, the binding curve of hemoglobin for oxygen is S-shaped (sigmoidal), whereas the binding curve for myoglobin is hyperbolic. What this tells us is that hemoglobin's affinity for oxygen is low at a low concentration oxygen, but increases as the oxygen concentration increases. Since myoglobin very quickly saturates with oxygen, even under low oxygen concentrations, it says that its affinity for oxygen is high and doesn't change.

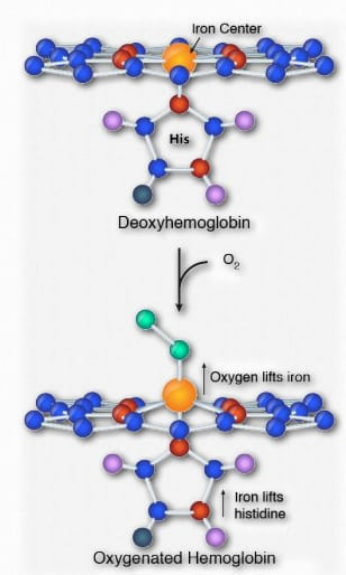
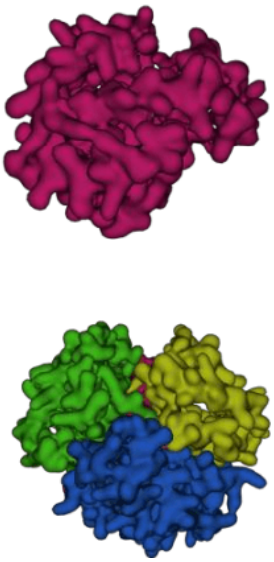


Figure 2.84 - Binding of oxygen at the heme center of hemoglobin Image by Aleia Kim

Because myoglobin has only a single subunit, binding of oxygen by that subunit can't affect any other subunits, since there are no other subunits to affect. Consequently, cooperativity requires more than one subunit. Therefore, hemoglobin can exhibit cooperativity, but myoglobin can't. It is worth noting that simply having multiple subunits does not mean cooperativity will exist. Hemoglobin is one protein that exhibits the characteristic, but many multisubunit proteins do not.



Interactive 2.2 - Hemoglobin in the presence (top) and absence (bottom) of oxygen

Storage vs. delivery

The lack of ability of myoglobin to adjust its affinity for oxygen according to the oxygen concentration (low affinity at low oxygen concentration, such as in tissues and high affinity at high oxygen concentration, such as in the lungs) means it is better suited for storing oxygen than for delivering it according to the varying oxygen needs of and animal body. As we shall see, besides cooperativity, hemoglobin has other structural features that allow it to deliver oxygen precisely where it is needed most in the body.

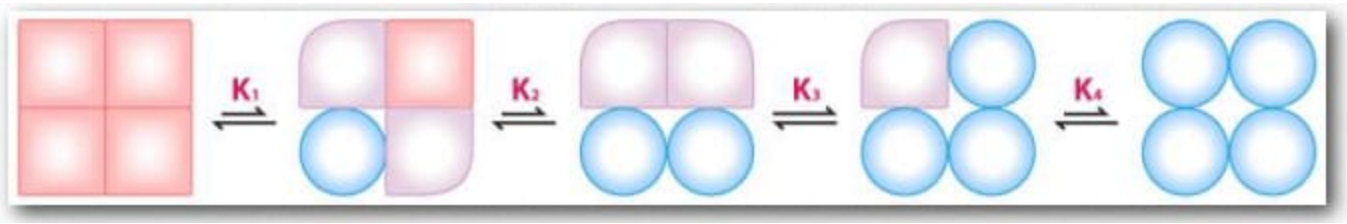


Figure 2.85 - Sequential model of binding. The sequential model is one way to explain hemoglobin's cooperativity. Squares represent no oxygen bound. Circles represent subunits bound with oxygen and rounded subunits correspond to units whose affinity for oxygen increases by interacting with a subunit that has bound oxygen. Image by Aleia Kim 131

Bohr effect

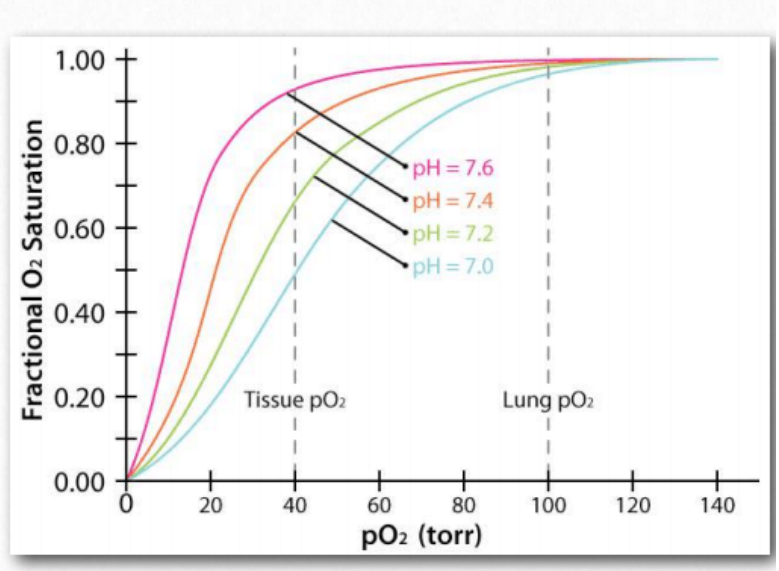


Figure 2.86 - The Bohr effect with respect to pH changes Image by Aleia Kim

The Bohr Effect was first described over 100 years ago by Christian Bohr, father of the famous physicist, Niels Bohr. Shown graphically (Figures 2.86, 2.87, and 2.88), the observed effect is that hemoglobin's affinity for oxygen decreases as the pH decreases and as the concentration of carbon dioxide increases. Binding of the protons and carbon dioxide by amino Figure 2.85 - Sequential model of binding. The sequential model is one way to explain hemoglobin's cooperativity. Squares represent no oxygen bound. Circles represent subunits bound with oxygen and rounded subunits correspond to units whose affinity for oxygen increases by interacting with a subunit that has bound oxygen. Image by Aleia Kim acid side chains in the globin proteins helps to facilitate structural changes in them. Most commonly, the amino acid affected by protons is histidine #146 of the β strands. When this happens, the ionized histidine can form an ionic bond with the side chain of aspartic acid #94, which has the effect of stabilizing the T-state (reduced oxygen binding state) and releasing oxygen. Other histidines and the amine of the amino terminal amino acids in the α-chains are also binding sites for protons.

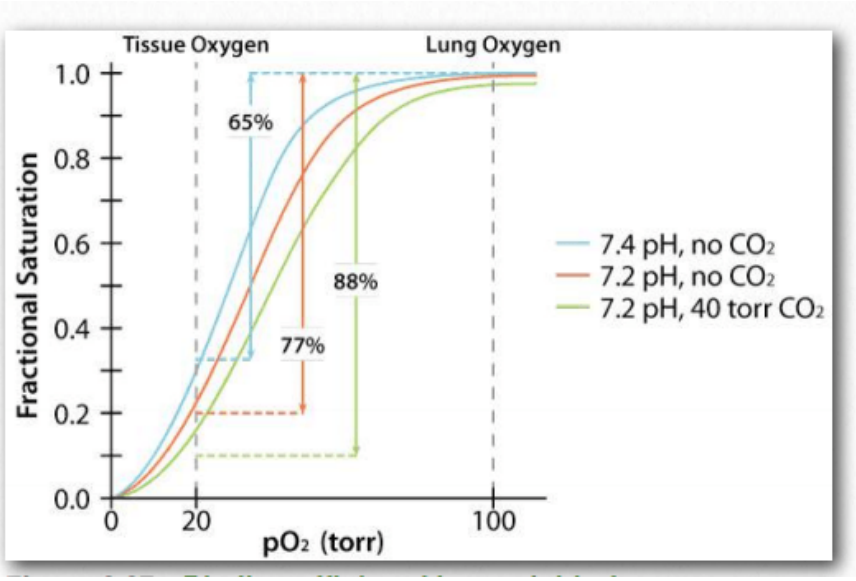


Figure 2.87 - Binding affinity of hemoglobin for oxygen under different conditions Image by Aleia Kim

2,3-BPG

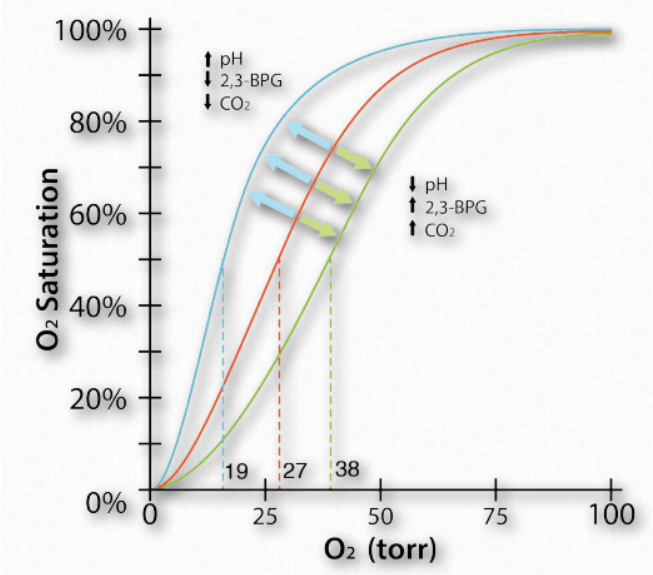


Figure 2.88 - The Bohr effect physiologically - oxygen binding curves for resting muscle (blue), active muscle (green) and reference muscle (orange) with respect to pH, 2,3-BPG, and CO2 Image by Aleia Kim

Another molecule favoring the release of oxygen by hemoglobin is 2,3- bisphosphoglycerate (also called 2,3-BPG or just BPG - Figure 2.89). Like protons and carbon dioxide, 2,3-BPG is produced by actively respiring tissues, as a byproduct of glucose metabolism. The 2,3-BPG mole cule fits into the 'hole of the donut' of adult hemoglobin (Figure 2.89). Such binding of 2,3-BPG favors the T-state (tight - low oxygen binding) of hemoglobin, which has a reduced affinity for oxygen. In the absence of 2,3-BPG, hemoglobin can more easily exist in the R-state (relaxed - higher oxygen binding), which has a high affinity for oxygen.

Smokers

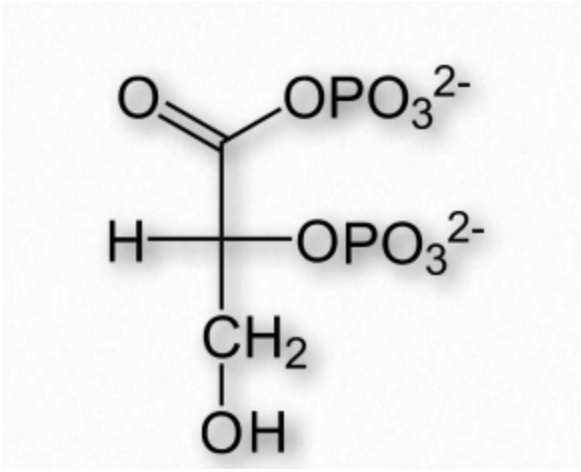


Figure 2.89 - The structure of 2,3 bisphosphoglycerate (2,3-BPG)

Notably, the blood of smokers is higher in the concentration of 2,3-BPG than non-smokers, so more of their hemoglobin remains in the T-state and thus the oxygen carrying capacity of smokers is lower than non-smokers.Another reason why smokers' oxygen carrying capacity is lower than that of non-smokers is that cigarette smoke contains carbon monoxide and this molecule, which has almost identical dimensions to molecular oxygen, effectively outcompetes with oxygen for binding to the iron atom of heme (Figure 2.90). Part of carbon monoxide's toxicity is due to its ability to bind hemoglobin and prevent oxygen from binding.

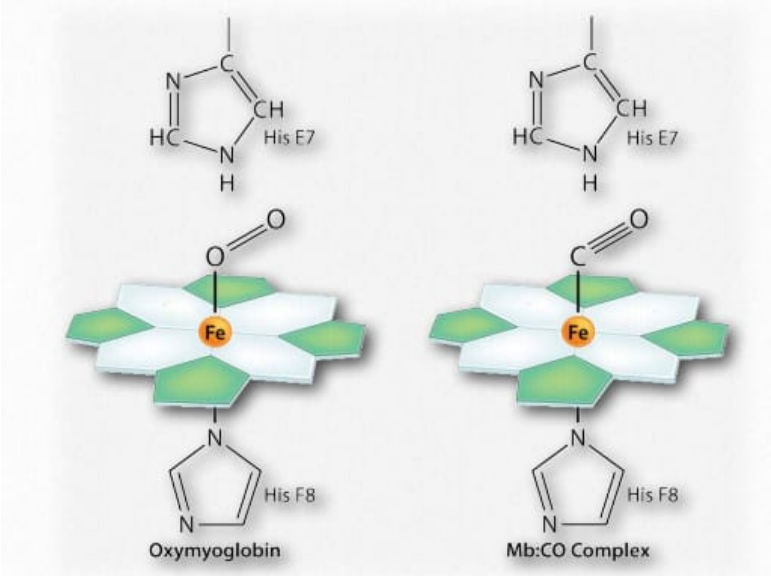


Figure 2.90 - Binding of oxygen (left) and carbon monoxide (right) by a heme group of hemoglobin Image by Aleia Kim

Carbon dioxide

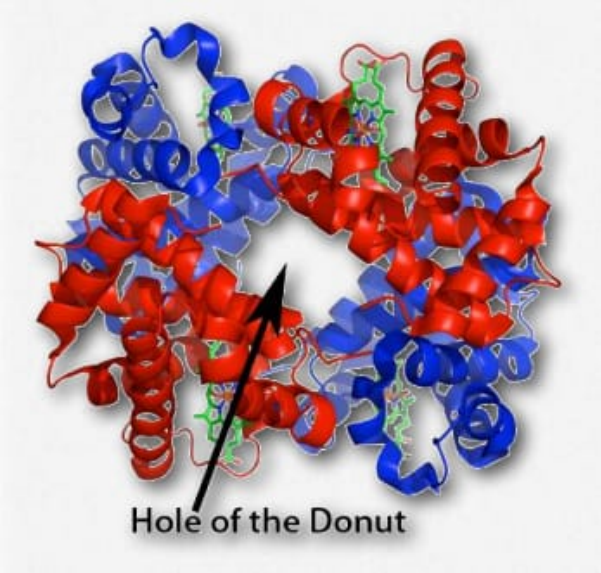


Figure 2.91 - Hemoglobin's hole of the donut for binding 2,3-BPG Wikipedia

Carbon dioxide binds to form a carbamate when binding the α-amine of each globin chain. The process of forming this structure releases a proton, which helps to further enhance the Bohr effect. Physiologically, the binding of CO2 and H+ has significance because actively respiring tissues (such as contracting muscles) require oxygen and release protons and carbon dioxide. The higher the concentration of protons and carbon dioxide, the more oxygen is released to feed the tissues that need it most.

About 40% of the released protons and about 20% of the carbon dioxide are carried back to the lungs by hemoglobin. The remainder travel as part of the bicarbonate buffering system or as dissolved CO2. In the lungs, the process reverses itself. The lungs have a higher pH than respiring tissues, so protons are released from hemoglobin and CO2 too is freed to be exhaled.

Fetal hemoglobin

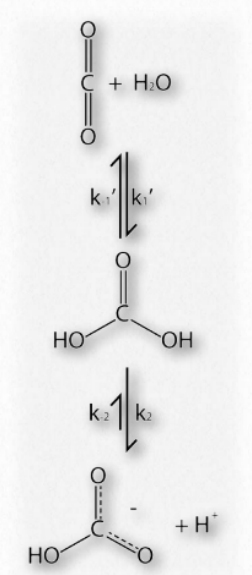


Figure 2.92 - Formation of bicarbonate from CO2 in blood

Adult hemoglobin releases oxygen when it binds 2,3- BPG. This is in contrast to fetal hemoglobin, which has a slightly different configuration (α2γ2) than adult hemoglobin (α2β2). Fetal hemoglobin has a greater affinity for oxygen than maternal hemoglobin, allowing the fetus to obtain oxygen effectively from the mother's blood. Part of the reason for fetal hemoglobin's greater affinity for oxygen is that it doesn't bind 2,3-BPG. Consequently, fetal hemoglobin remains in the R-state much more than adult hemoglobin and because of this, fetal hemoglobin has greater affinity for oxygen than adult hemoglobin and can take oxygen away from adult hemoglobin. Thus, the fetus can get oxygen from the mother.

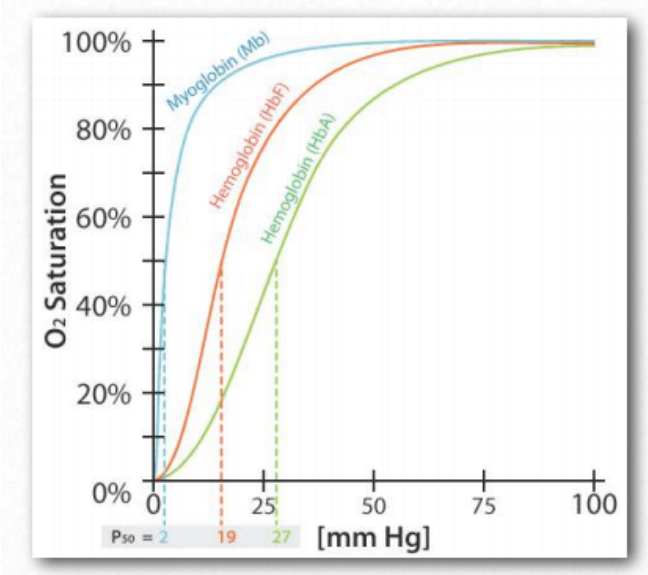


Figure 2.93 - Comparison of oxygen binding of myoglobin (blue), fetal hemoglobin (orange), and adult hemoglobin (green) Image by Aleia Kim

Sickle cell disease



Figure 2.94 - Four normal red blood cells (right) and one sickled red blood cell (left) Wikipedia

Mutations to the globin genes coding for hemoglobin can sometimes have deleterious consequences. Sickle cell disease (also called sickle cell anemia) is a genetically transmitted disease that arises from such mutations. There are different forms of the disease. It is a recessive trait, meaning that to be afflicted with it, an individual must inherit two copies of the mutated gene.

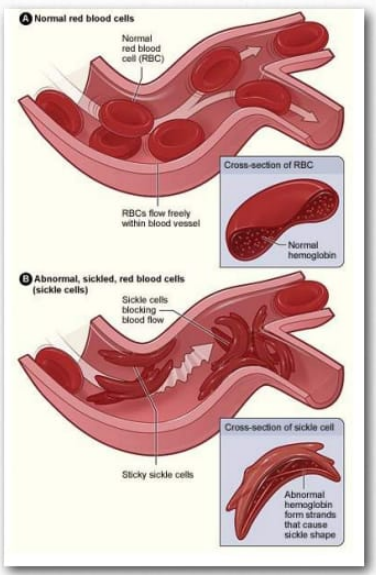


Figure 2.95 - Movement of blood in capillaries. Top - normal red blood cells. Bottom - sickled red blood cells

The predominant form of hemoglobin in adults is hemoglobin A, designated HbA (two α chains and two β chains). The mutant form is known as HbS. The most common mutation is an A to T mutation in the middle of the codon for the seventh amino acid (some counting schemes call it the sixth amino acid) of the β-chain. This results in conversion of a GAG codon to GTG and thus changes the amino acid specified at that position from a glutamic acid to a valine. This minor change places a small hydrophobic patch of amino acids on the surface of the β-globin chains.

Polymerization

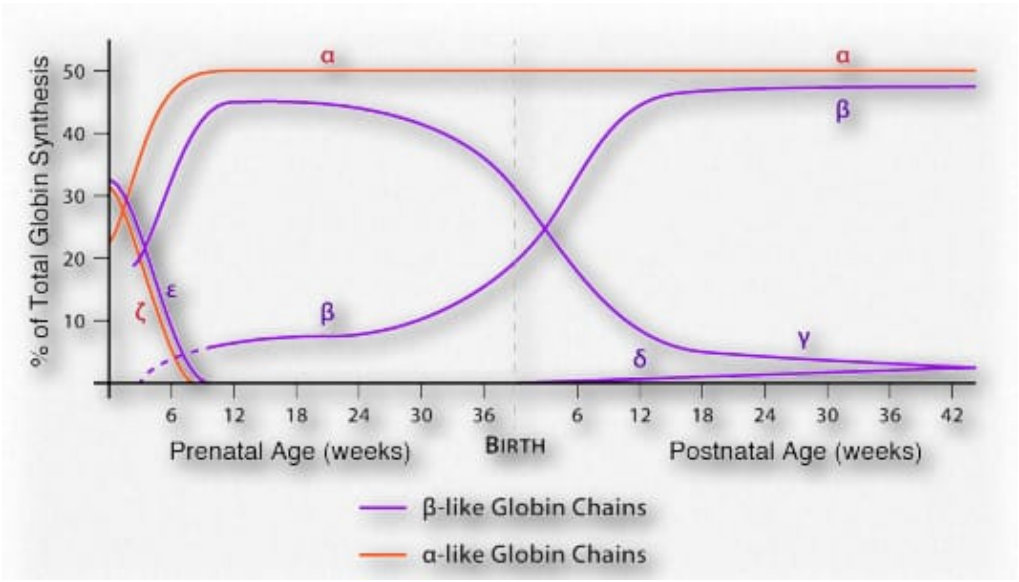


Figure 2.96 - Pattern of expression of six different globins of hemoglobin - α,β,γ,ε,δ, and ζ Image by Aleia Kim

Under conditions of low oxygen, these hydrophobic patches will associate with each other to make long polymers of hemoglobin molecules. The result is that the red blood cells containing them will change shape from being rounded to forming the shape of a sickle (Figure 2.94). Rounded red blood cells readily make it through tiny capillaries, but sickleshaped cells do not.

Worse, they block the flow of other blood cells. Tissues where these blockages occur are already low in oxygen, so stopping the flow of blood through them causes them to go quickly anaerobic, causing pain and, in some cases, death of tissue. In severe circumstances, sickled red blood cells death may result. The disease is referred to as an anemia because the sickling of the red blood cells targets them for removal by the blood monitoring system of the body, so a person with the disease has chronically reduced numbers of red blood cells.

Heterozygote advantage

Interestingly, there appears to be a selective advantage to people who are heterozygous for the disease in areas where malaria is prominent. Heterozygotes do not suffer obvious ill effects of the disease, but their red blood cells appear to be more susceptible to rupture when infected. As a consequence, the parasite gets less of a chance to reproduce and the infected person has a greater chance of survival.

The protective effect of the mutant gene, though, does not extend to people who suffer the full blown disease (homozygotes for the mutant gene). Treatments for the disease include transfusion, pain management, and avoidance of heavy exertion. The drug hydroxyurea has been linked to reduction in number and severity of attacks, as well as an increase in survival time1,2. It appears to work by reactivating expression of the fetal hemoglobin gene, which typically is not synthesized to any significant extent normally after about 6 weeks of age.

Oxygen binding

Animals have needs for oxygen that differ from all other organisms. Oxygen, of course, is the terminal electron acceptor in animals and is necessary for electron transport to work. When electron transport is functioning, ATP generation by cells is many times more efficient than when it is absent. Since abundant ATP is essential for muscular contraction and animals move around a lot - to catch prey, to exercise, to escape danger, etc., having an abundant supply of oxygen is important.

This is particularly a concern deep inside tissues where diffusion of oxygen alone (as occurs in insects) does not deliver sufficient quantities necessary for long term survival. The issue is not a problem for plants since, for the most part, their motions are largely related to growth and thus don't have rapidly changing needs/demands for oxygen that animals have. Unicellular organisms have a variety of mechanisms for obtaining oxygen and surviving without it. Two other important oxygen binding proteins besides hemoglobin are myoglobin and hemocyanin.

Myoglobin

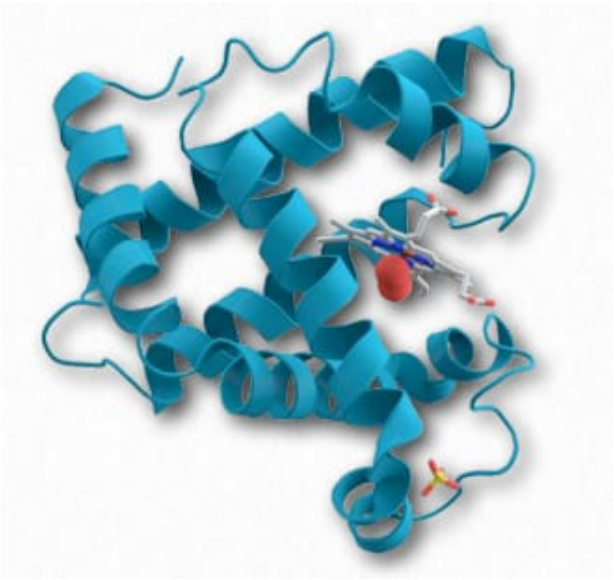


Figure 2.97 - Myoglobin bound to oxygen

Myoglobin is the primary oxygen-storage protein found in animal muscle tissues. In contrast to hemoglobin, which circulates throughout the body, myoglobin protein is only found in muscle tissue and appears in the blood only after injury. Like hemoglobin, myoglobin binds oxygen at a prosthetic heme group it contains.

The red color of meat arises from the heme of myoglobin and the browning of meat by cooking it comes from oxidation of the ferrous (Fe++) ion of myoglobin's heme to the ferric (Fe+++) ion via oxidation in the cooking process. As meat sits in our atmosphere (an oxygen-rich environment), oxidation of Fe++ to Fe+++ occurs, leaving the brown color noted above. If meat is stored in a carbon monoxide (CO) environment, CO binds to the heme group and reduces the amount of oxidation, keeping meat looking red for a longer period of time.

High affinity

Myoglobin (Figure 2.97) displays higher affinity for oxygen at low oxygen concentrations than hemoglobin and is therefore able to absorb oxygen delivered by hemoglobin under these conditions. Myoglobin's high affinity for oxygen makes it better suited for oxygen storage than delivery. The protein exists as a single subunit of globin (in contrast to hemoglobin, which contains four subunits) and is related to the subunits found in hemoglobin. Mammals that dive deeply in the ocean, such as whales and seals, have muscles with particularly high abundance of myoglobin. When oxygen concentration in muscles falls to low levels, myoglobin releases its oxygen, thus functioning as an oxygen "battery" that delivers oxygen fuel when needed and holding onto it under all other conditions. Myoglobin holds the distinction of being the first protein for which the 3D structure was determined by X-ray crystallography by John Kendrew in 1958, an achievement for which he later won the Nobel Prize.

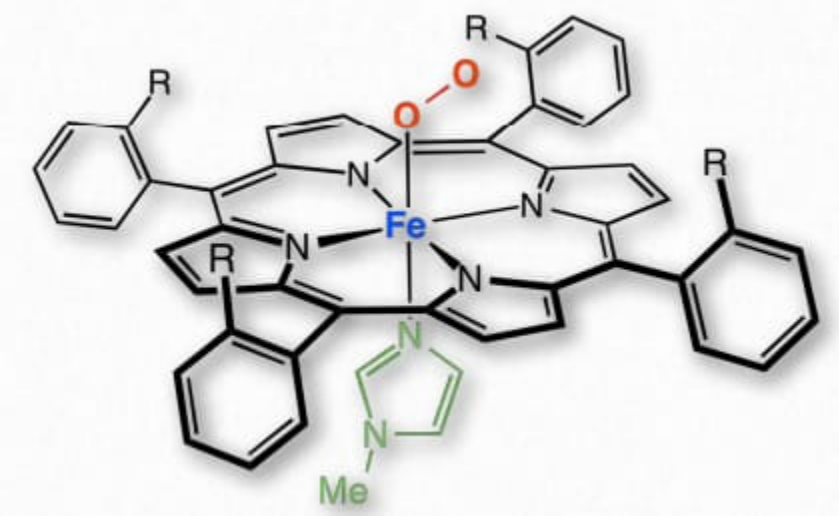


Figure 2.98 - Oxygen bound at heme of myoglobin Wikipedia

Hemocyanin

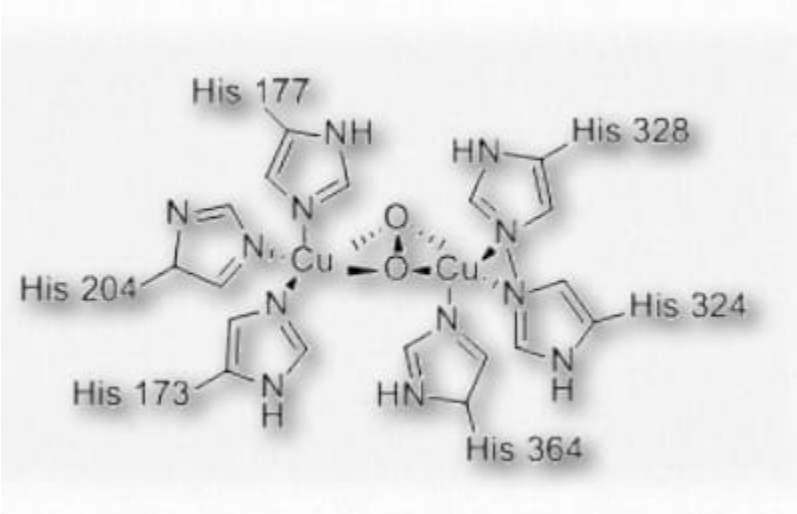


Figure 2.99 - Oxygen binding in hemocyanin Wikipedia

Hemocyanin is the protein transporting oxygen in the bodies of molluscs and arthropods. It is a coppercontaining protein found not within blood cells of these organisms, but rather is suspended in the circulating hemolymph they possess. The oxygen binding site of hemocyanin contains a pair of copper(I) cations directly coordinated to the protein by the imidazole rings of six histidine side chains.



Figure 2.100 - Hemocyanin (purple) in a red rock crab Wikipedia

Most, but not all hemocyanins bind oxygen non-cooperatively and are less efficient than hemoglobin at transporting oxygen. Notably, the hemocyanins of horseshoe crabs and some other arthropods do, in fact, bind oxygen cooperatively. Hemocyanin contains many subunit proteins, each with two copper atoms that can bind one oxygen molecule (O_2). Subunit proteins have atomic masses of about 75 kilodaltons (kDa). These may be arranged in dimers or hexamers depending on species. Superstructures comprised of dimer or hexamer complexes are arranged in chains or clusters and have molecular weights of over 1500 kDa.

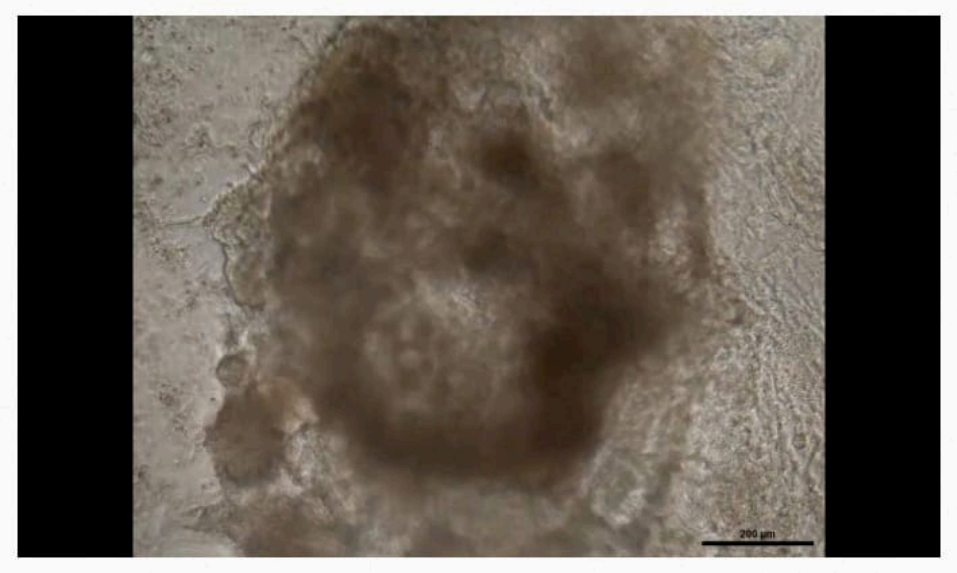
This page titled 2.4: Structure and Function- Proteins II is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.5: Structure and Function- Protein Function II

Source: BiochemFFA_2_4.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

To this point, the proteins we have discussed have not been catalysts (enzymes). The majority of proteins in cells, however, catalyze reactions. In this section we begin our discussion of a subclass of proteins that catalyze reactions releasing energy and convert it into mechanical force. These operate at the cellular and organismal level and are known as motor proteins. Motor proteins rely on globular structural proteins, so it is important that we describe how these cellular "railways" are assembled before discussing the motor proteins themselves. There are two relevant fibrous structures serving as rails for motor proteins. They are:

1. **microfilaments** (composed of an actin polymer) and
2. **microtubules** (composed of a polymer of tubulin.



Actin

The monomeric unit of actin is called G-actin (globular actin) and the polymer is known as F-actin (filamentous actin). Filaments of Factin comprise the smallest filaments of cells known as microfilaments (Figure 2.10̸1). Actin is essential for muscular contraction and also has diverse roles in cellular signaling and maintenance of cell junctions. In conjunction with other proteins, actin has numerous interactions with the cell membrane. The β- and γ-forms of actin are components of the cytoskeleton and facilitate motility inside of cells. α-actin is important in muscle tissues, where it is used by myosin in the mechanical process of contraction (See HERE).

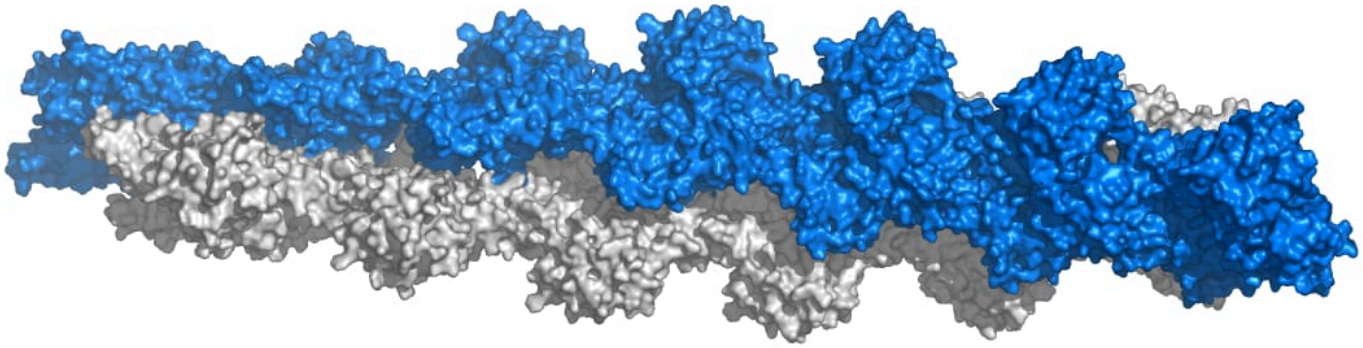


Figure 2.101 - Model of actin filaments. (CC BY-SA 3.0; Thomas Splettstoesser).

Monomeric and polymeric forms of actin play roles in cellular activities relating to motion. Two parallel F-actin strands can pair with each other and create a double helical structure with 2.17 subunits per turn of the helix. Helical F-actin in muscles contains tropomyosin, which covers the actin binding sites for myosin in resting muscles to prevent contraction. Other proteins bound to actin muscle filaments include the troponins (I, T, and C).

Actin Cellular Action

Examples of actin action at the cellular level include cell motility, cytokinesis, intracellular transport of vesicles and organelles, and cell shape. Each actin monomer is bound to a molecule of ATP or ADP and the presence of one of these is essential for proper G-actin functioning.

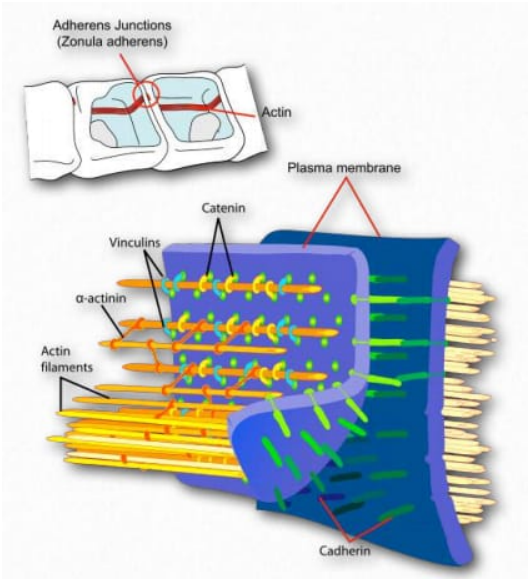


Figure 2.102 - Attachment of actin at the cell membrane complex known as the adherens junction Wikipedia

The role of ATP

In the monomer, actin is more commonly bound to ATP, whereas in the filaments, it is typically bound to ADP. Actin is an inefficient ATPase, breaking the molecule down slowly, but the catalysis speeds up as much as 40,000 fold when the monomer begins to polymerize. Actin also has a binding site for divalent cations - either calcium or magnesium. F-

Actin binds to structural proteins at the adherens junction (Figure 2.102). These include α-actinin, vinculin (provides a membrane connection and connections to the catenins and cadherin).

Polymerization

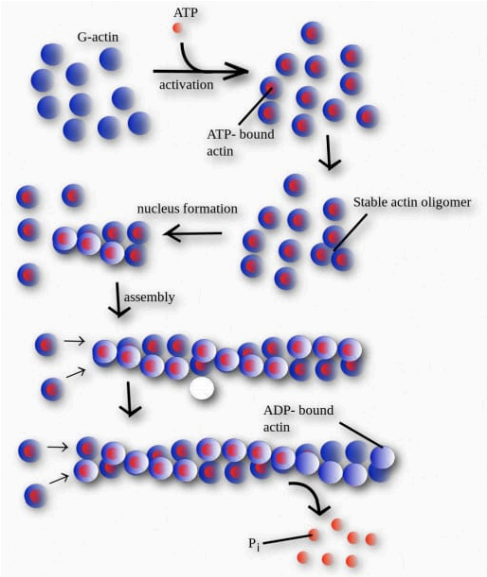


Figure 2.103 - Polymerization of G-actin monomers into Factin polymers

Polymerization of actin begins with a nucleating event (Figure 2.103). One factor known to affect the process is known as the Arp 2/3 complex. It does this by mimicking an actin dimer, starting an autocatalytic process of actin assembly. The Arp 2/3 complex plays roles both in the initiation of polymerization of new actin filaments as well as the formation of branches in the filaments.

Two proteins play roles in modulating polymer growth. Thymosin functions on the end of actin filaments to control growth. Profilin works on G-actin monomers exchanging ADP for ATP, promoting addition of monomers to a growing chain.

F-actin filaments are held together by relatively weak bonds compared to the covalent bonds of the monomers of nucleic acids, thus allowing for easier disassembly when desired. Actin's amino acid sequence is optimized, having diverged only a relatively small amount (20%) between algae and humans. Mutations in the actin gene result in muscular diseases and/or deafness.

Tubulin

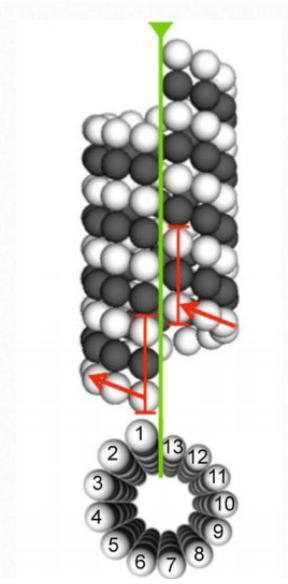


Figure 2.104 Microtubule structure Wikipedia

Tubulin proteins are the monomeric building blocks of eukaryotic microtubules (Figure 2.104 & 2.105). Bacterial (TubZ) and archaeon (FtsZ) equivalents are known. The α-tubulin and β-tubulin proteins polymerize to make microtubule structures in the cytoplasm of cells. Microtubules are major components of the cytoskeleton of eukaryotic cells, providing structural support, transport within the cell, and functions necessary for segregation of DNAs during cell division.

Dimerization of the α-tubulin and β-tubulin proteins is necessary for polymerization and requires that the subunits bind to GTP. Microtubules only grow in one direction. β- tubulin is found on the plus end of the tubule (growth end = plus end) and α-tubulin is exposed on the other end (non-growth end = minus end). Dimers of α-tubulin/β-tubulin are incorporated into growing microtubules in this orientation. If a dimer is bound to GDP instead of GTP, it tends to be unstable and fall apart, whereas those bound to GTP stably assemble into microtubules.

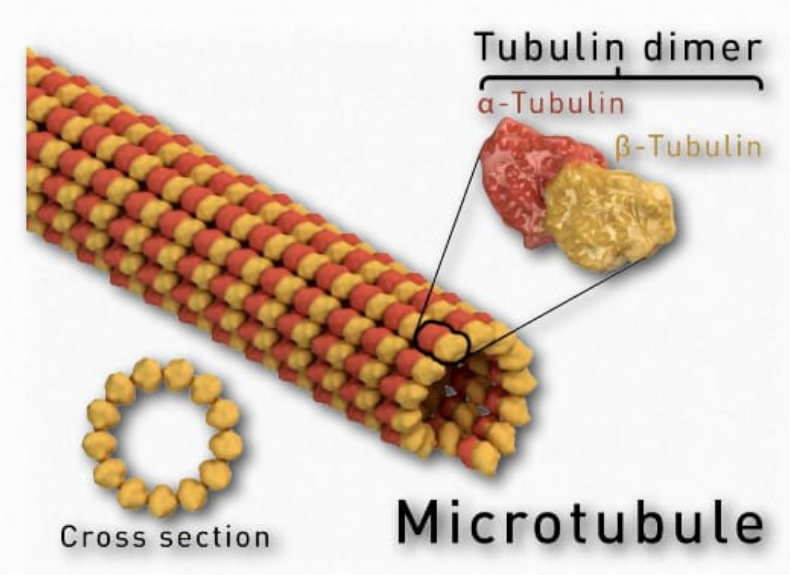


Figure 2.105 - Microtubule anatomy Wikipedia

Microtubules

Microtubules, along with microfilaments and intermediate filaments (see HERE) constitute the cytoskeleton of cells. Found in the cytoplasm, they are found in eukaryotic cells, as well as some bacteria. Microtubules help to give cells structure. They comprise the inner structure of flagella and cilia and provide rail-like surfaces for the transport of materials within cells.

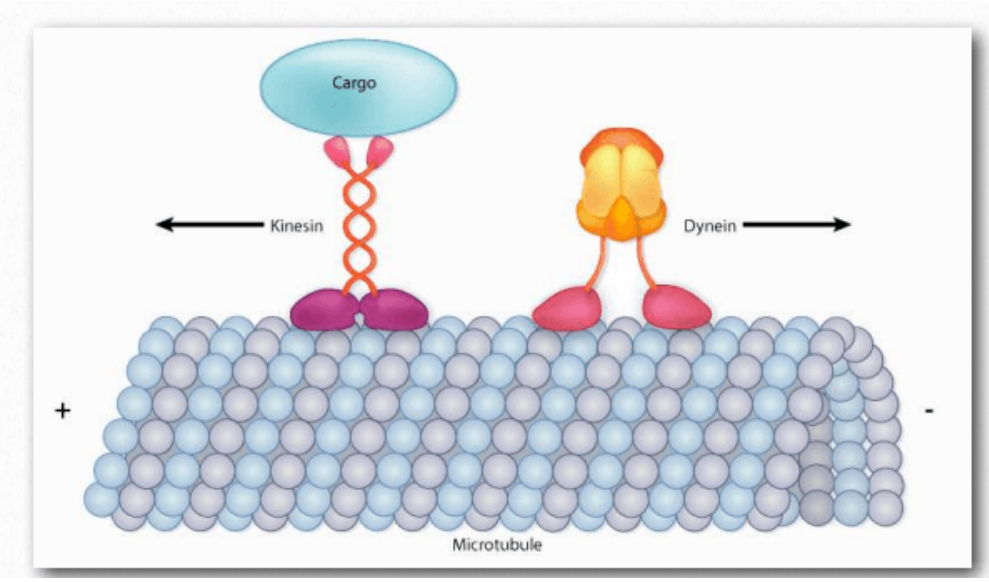


Figure 2.106 - Kinesin and dynein “walk” along microtubules, but move in opposite directions Image by Aleia Kim

Polymerization of α- tubulin and β-tubulin to form microtubules occurs after a nucleating event. Individual units get arranged in microtubule organizing centers (MTOCs), an example of which is the centrosome. Centrosomes are focal points of connection of microtubules. Basal bodies of cilia and flagella also help to organize microtubules.

Motor proteins

From the transport of materials within a cell to the process of cytokinesis where one cell splits into two in mitosis, a cell has needs for motion at the molecular level. Secretory vesicles and organelles must be transported. Chromosomes must be separated in mitosis and meiosis.

The proteins dynein and kinesin (Figure 2.106) are necessary for intracellular movement. These motor proteins facilitate the movement of materials inside of cells along microtubule “rails”. These motor proteins are able to move along a portion of the cytoskeleton by converting chemical energy into motion with the hydrolysis of ATP. An exception is flagellar rotation, which uses energy provided from a gradient created by a proton pump.

Kinesins and dyneins

As noted, kinesins and dyneins navigate in cells on microtubule tracks (Figure 2.108 & Movie 2.4). Most kinesins move in the direction of the synthesis of the microtubule (+ end movement), which is generally away from the cell center and the opposite direction of movement of dyneins, which are said to do retrograde transport toward the cell center. Both proteins provide movement functions necessary for the processes of mitosis and meiosis. These include spindle formation, chromosome separation, and shuttling of organelles, such as the mitochondria, Golgi apparatuses, and vesicles.

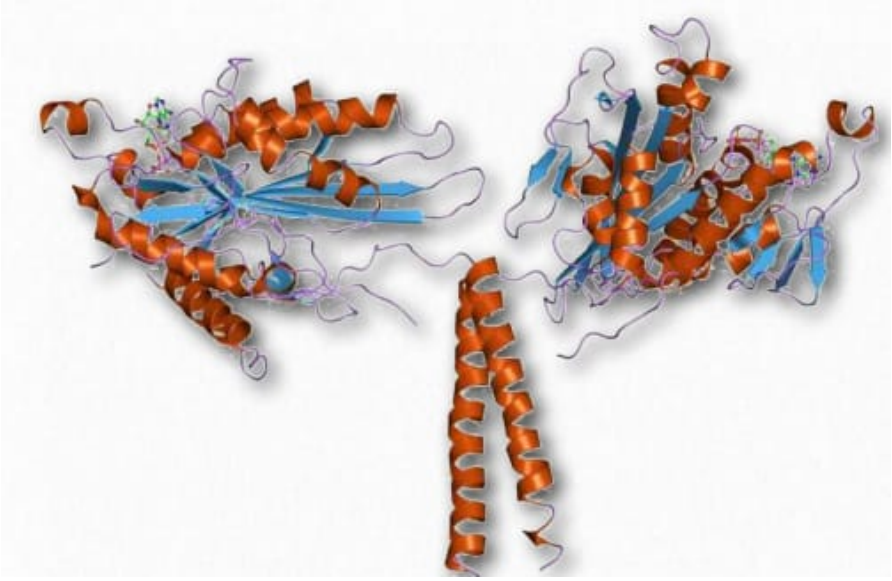


Figure 2.107 - Kinesin. “Feet” are at the top.

Kinesins are comprised of two heavy chains and two light chains. The head motor domains of heavy chains (in the feet) use energy of ATP hydrolysis to do mechanical work for the movement along the microtubules. There are at least fourteen distinct kinesin families and probably many related ones in addition.

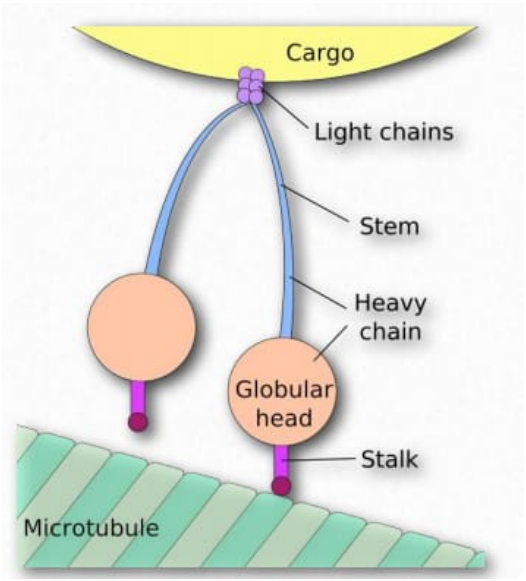
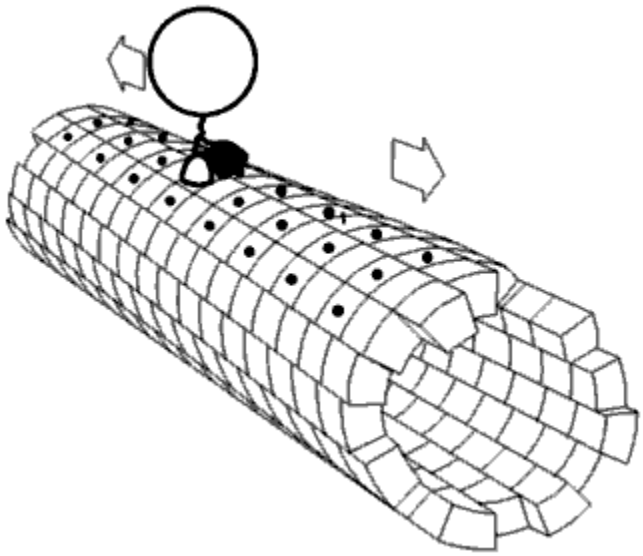


Figure 2.105 are the stalk and globular head of the structure here. Wikipedia

Dyneins are placed into two groups - cytoplasmic and axonemal (also called ciliary or flagellar dyneins - Figure 2.109). Dyneins are more complex in structure than kinesins with many small polypeptide units. Notably, plants do not have dynein motor proteins, but do contain kinesins.



Movie 2.4 The motor protein kinesin walking down a microtubule. (Public Domain; zp706).

Myosin

An important group of motor proteins in the cell is the myosins. Like kinesins and dyneins, myosins use energy from hydrolysis of ATP for movement. In this case, the movement is mostly not along microtubules, but rather along microfilaments comprised of a polymer of actin (F-actin). Movement of myosin on actin is best known as the driving force for muscular contraction. Myosins are a huge family of proteins, all of which bind to actin and all of which involve motion. Eighteen different classes of myosin proteins are known.

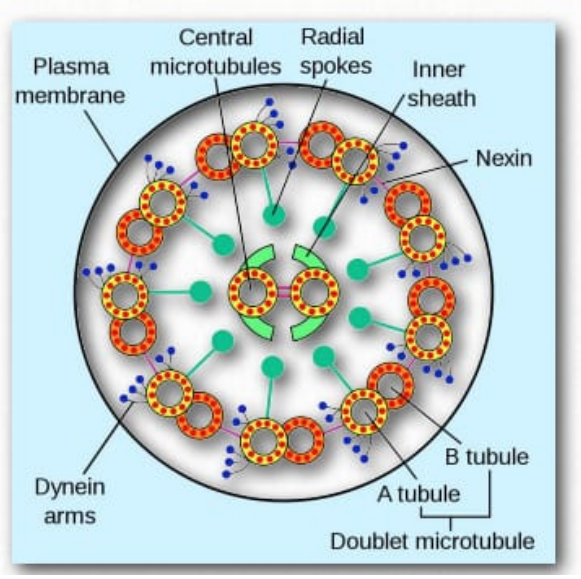


Figure 2.109 - Dynein in an axoneme Wikipedia

Myosin II is the form responsible for generating muscle contraction. It is an elongated protein formed from two heavy chains with motor heads and two light chains. Each myosin motor head binds actin and has an ATP binding site. The myosin heads bind and hydrolyze ATP. This hydrolysis produces the energy necessary for myosin to walk toward the plus end of an actin filament.

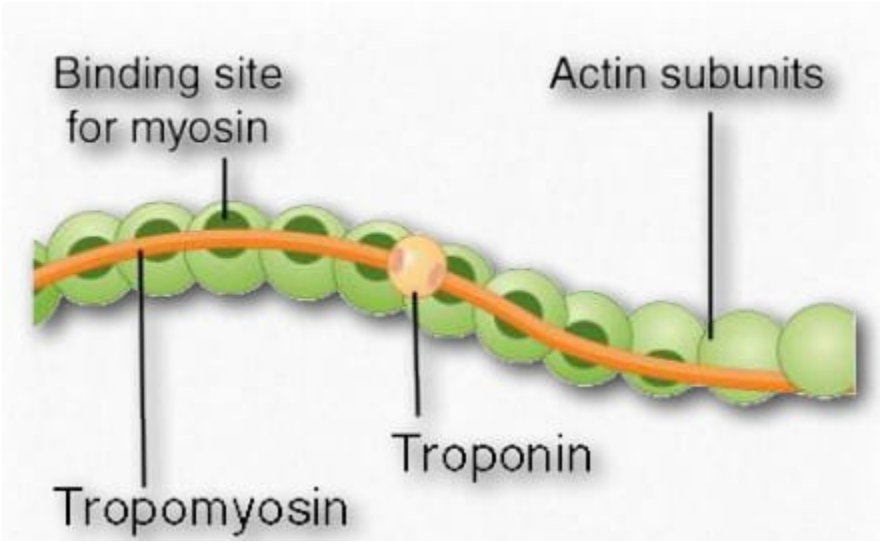


Figure 2.110 - Actin filament anatomy Wikipedia

Non-muscle myosin IIs provide contraction needed to power the action of cytokinesis. Other myosin proteins are involved in movement of non-muscle cells. Myosin I is involved in intracellular organization. Myosin V performs vesicle and organelle transport. Myosin XI provides movement along cellular microfilament networks to facilitate organelle and cytoplasmic streaming in a particular direction.

Structure

Myosins have six subunits, two heavy chains and four light chains. Myosin proteins have domains frequently described as a head and a tail (Figure 2.111). Some also describe an intermediate hinge region as a neck. The head portion of myosin is the part that binds to actin. It uses energy from ATP hydrolysis to move along the actin filaments. In muscles, myosin proteins form aggregated structures referred to as thick filaments. Movements are directional.

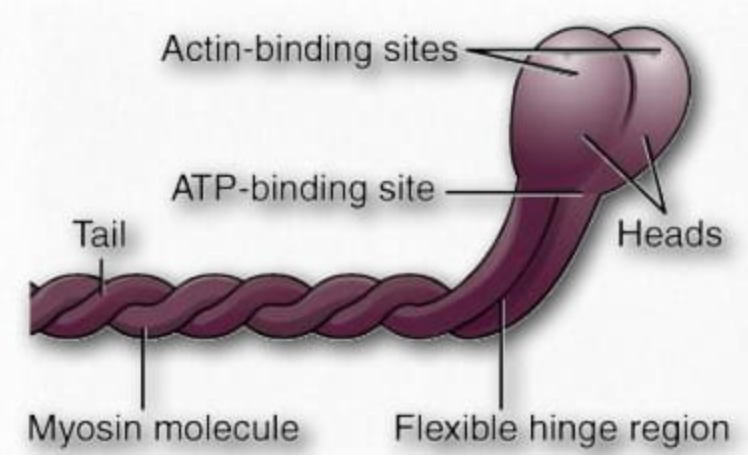


Figure 2.111 - Myosin protein anatomy Wikipedia

Structural considerations of muscular contraction

Before we discuss the steps in the process of muscular contraction, it is important to describe anatomical aspects of muscles and nomenclature.

There are three types of muscle tissue - skeletal (striated), smooth, and (in vertebrates) cardiac. We shall concern ourselves mostly here with skeletal muscle tissue. Muscles may be activated by the central nervous system or, in the case of smooth and cardiac muscles, may contract involuntarily. Skeletal muscles may be slow twitch or fast twitch.

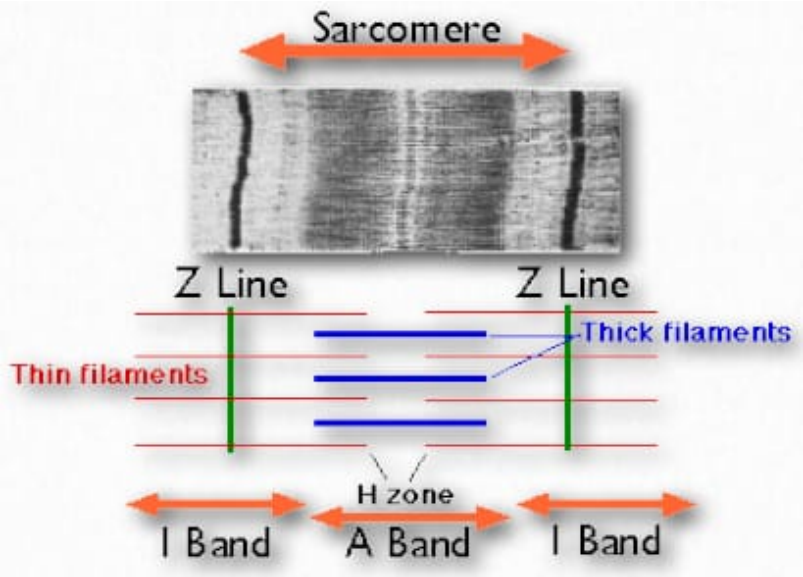


Figure 2.112 - Structural components of muscle. Wikipedia

Sarcomeres

Sarcomeres are described as the basic units comprising striated muscles and are comprised of thick (myosin) and thin (actin) filaments and a protein called titin. The filaments slide past each other in muscular contraction and then backwards in muscular relaxation. They are not found in smooth muscles.

Under the microscope, a sarcomere is the region between two Z-lines of striated muscle tissue (Figure 2.112). The Z-line is the distinct, narrow, dark region in the middle of an I-band. Within the sarcomere is an entire Aband with its central H-zone. Within the Hzone are located tails of myosin fibers, with the head pointed outwards from there projecting all the way to the I-band. The outside of the Aband is the darkest and it gets lighter moving towards the center.

Within the Iband are located thin filaments that are not occupied with thick myosin filaments. The Aband contains intact thick filaments overlaying thin filaments except in the central H zone, which contains only thick filaments. In the center of the H-zone is a line, known as the M-line. It contains connecting elements of the cellular cytoskeleton. In muscular contraction, myosin heads walk along pulling their tails over the actin thin filaments, using energy from hydrolysis of ATP and pulling them towards the center of the sarcomere.



Figure 2.113 - Skeletal muscle longitudinal section. Wikipedia

Sarcolemma

The sarcolemma (also known as the myolemma) is to muscle cells what the plasma membrane is to other eukaryotic cells - a barrier between inside and outside. It contains a lipid bilayer and a glycocalyx on the outside of it. The glyocalyx contains polysaccharides and connects with the basement membrane. The basement membrane serves a s scaffolding to connect muscle fibers to. This connection is made by transmembrane proteins bridging the actin cytoskeleton on the inside of the cell with the basement membrane on the outside. On the ends of the muscle fibers, each sarcolemma fuses with a tendon fiber and these, in turn, adhere to bones.

Sarcoplasmic reticulum

The sarcoplasmic reticulum (Figure 2.114) is a name for the structure found within muscle cells that is similar to the smooth endoplasmic reticulum found in other cells. It contains a specialized set of proteins to meet needs unique to muscle cells. The organelle largely serves as a calcium “battery,” releasing stored calcium to initiate muscular contraction when stimulated and taking up calcium when signaled at the end of the contraction cycle. It accomplishes these tasks using calcium ion channels for release of the ion and specific calcium ion pumps to take it up.

Movement direction

All myosins but myosin VI move towards the + end (the growing end) of the microfilament. The neck portion serves to link the head and the tail. It also a binding site for myosin light chain proteins that form part of a macromolecular complex with regulatory functions. The tail is the point of attachment of molecules or other “cargo” being moved. It can also connect with other myosin subunits and may have a role to play in controlling movement.

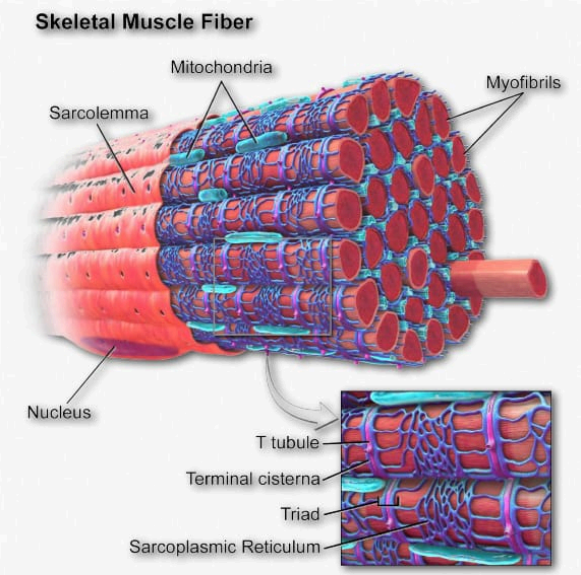


Figure 2.114 - Anatomy of a muscle fiber Wikipedia

Muscular contraction

The sliding filament model has been proposed to describe the process of muscular tension/contraction. In this process a repeating set of actions slide a thin actin filament over a thick myosin filament as a means of creating tension/ shortening of the muscle fiber.

Steps in the process occur as follows:

A. A signal from the central nervous system (action potential) arrives at a motor neuron, which it transmits towards the neuromuscular junction (see more on the neurotransmission part of the process HERE)

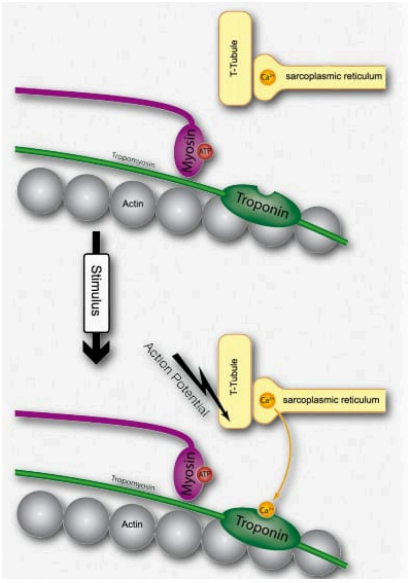


Figure 2.115 - 1. Activation of a muscle cell by release of calcium (step H) Wikipedia

B. At the end of the axon, the nerve signal stimulates the opening of calcium channels at the axon terminus causing calcium to flow into the terminal.

C. Movement of calcium into the axon of the nerve causes acetylcholine (a neurotransmitter) in synaptic vesicles to fuse with the plasma membrane. This causes the acetylcholine to be expelled into the synaptic cleft between the axon and the adjacent skeletal muscle fiber.

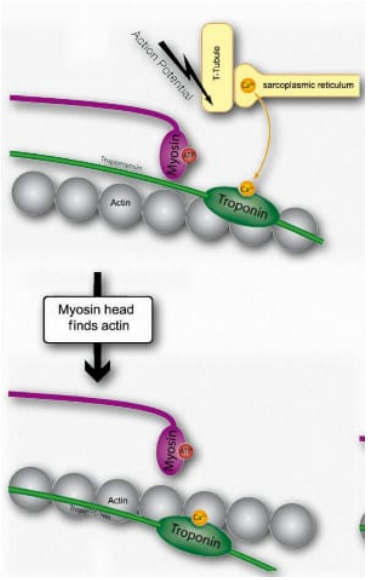


Figure 2.116 - 2. Calcium binding by troponin allows myosin to access actin sites (I). Wikipedia

D. Acetylcholine diffuses across the synapse and then binds to nicotinic acetylcholine receptors on the neuromuscular junction, activating them.

E. Activation of the receptor stimulates opening gates of sodium and potassium channels, allowing sodium to move into the cell and potassium to exit. The polarity of the membrane of the muscle cell (called a sarcolemma - Figure 2.111) changes rapidly (called the end plate potential).

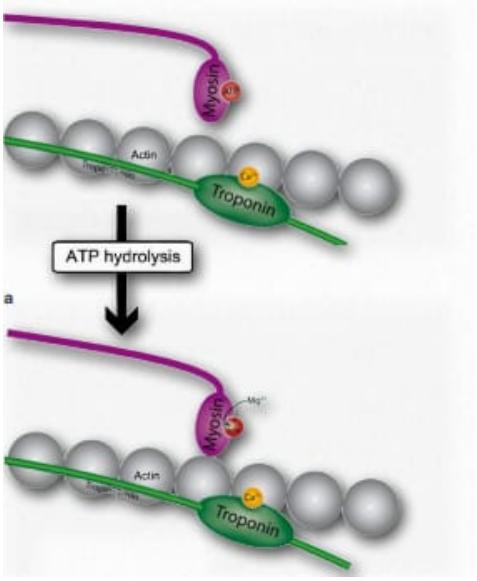


Figure 2.117 - 3. ATP cleavage by myosin allows actin attachment (J) Wikipedia

F. Change in the end plate potential results in opening of voltage sensitive ion channels specific for sodium or potassium only to Figure 2.117 - 3. ATP cleavage by myosin allows actin attachment (J) Wikipediaopen, creating an action potential (voltage change) that spreads throughout the cell in all directions.

G. The spreading action potential depolarizes the inner muscle fiber and opens calcium channels on the sarcoplasmic reticulum (Figure 2.115).

H. Calcium released from the sarcoplasmic reticulum binds to troponin on the actin filaments (Figure 2.115).

I. Troponin alters the structure of the tropomyosin to which is it bound. This causes tropomyosin to move slightly, allowing access

to myosin binding sites on the microfilament (also called thin filament) that it was covering (Figure 2.116).

J. Myosin (bound to ATP) cleaves the ATP to ADP and Pi, which it holds onto in its head region and then attaches itself to the exposed binding sites on the thin filaments causing inorganic phosphate to be released from the myosin followed by ADP (Figure 2.117).

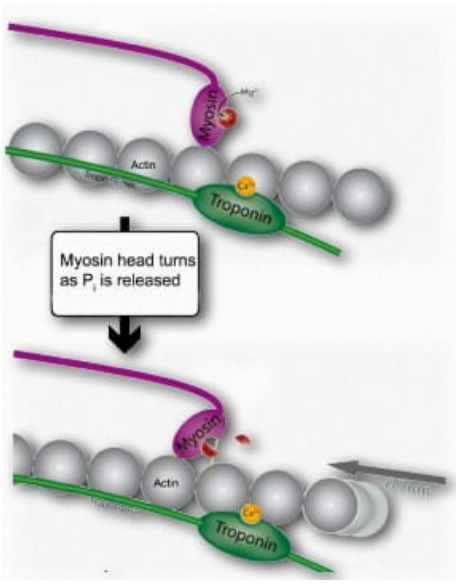


Figure 2.118 - 4. Release of Pi causes myosin hinge to bend. Thin filament pulled left (K). Wikipedia

K. Release of ADP and Pi is tightly coupled to a bending of the myosin hinge, resulting in what is called the power stroke. This causes the thin filament to move relative to the thick fibers of myosin (Figures 2.118 & 2.119).

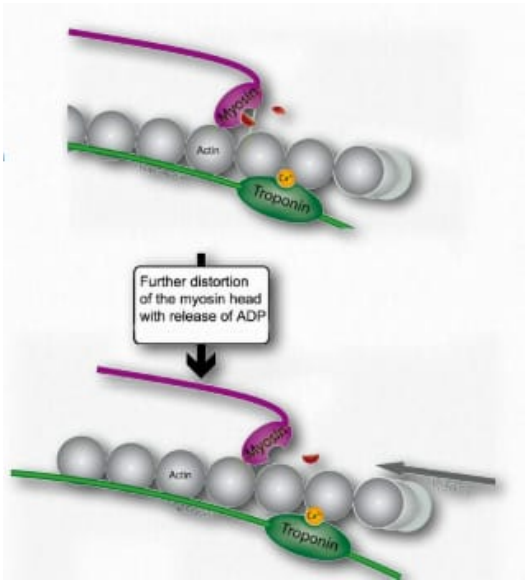


Figure 2.119 - 5. Release of ADP favors further bending of hinge and movement of thin filament leftward (K). Wikipedia

L. Such movement of the thin filaments causes the Z lines to be pulled closer to each other. This results in shortening of the sarcomere as a whole (Figure 2.122) and narrowing of the I band and the H zones (Figure 2.123). M. If ATP is available, it binds to myosin, allowing it to let go of the actin (Figures 2.120 & 2.121). If ATP is not available, the muscle will remain locked in this state. This is the cause of rigor mortis in death - contraction without release of muscles

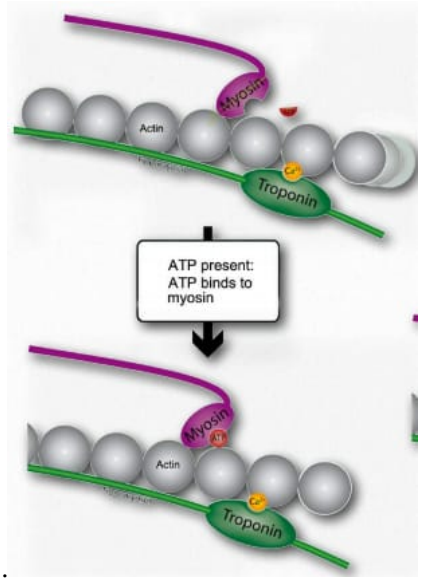


Figure 2.120 - When ATP is present, it binds to myosin (M). Wikipedia

N. After myosin has bound the ATP, it hydrolyzes it, producing ADP and Pi, which are held by the head. Hydrolysis of ATP resets the hinge region to its original state, unbending it. This unbent state is also referred to as the cocked position.

O.If tropomyosin is still permitting access to binding sites on actin, the process repeats so long as ATP is available and calcium remains at a high enough concentration to permit it to bond to troponin.

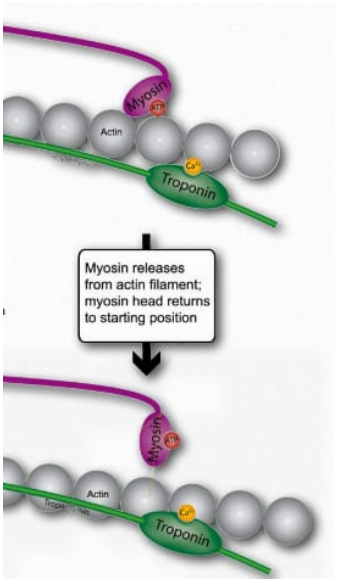


Figure 2.121 - Binding of ATP favors release of myosin from actin site (N) Wikipedia

Relaxation of the muscle tension occurs as the action potential in the muscle cell dissipates. This happens because all of the following things happen 1) the nerve signal stops; 2) the neurotransmitter is degraded by the enzyme acetylcholinesterase; and 3) the calcium concentration declines because it is taken up by the sarcoplasmic reticulum.

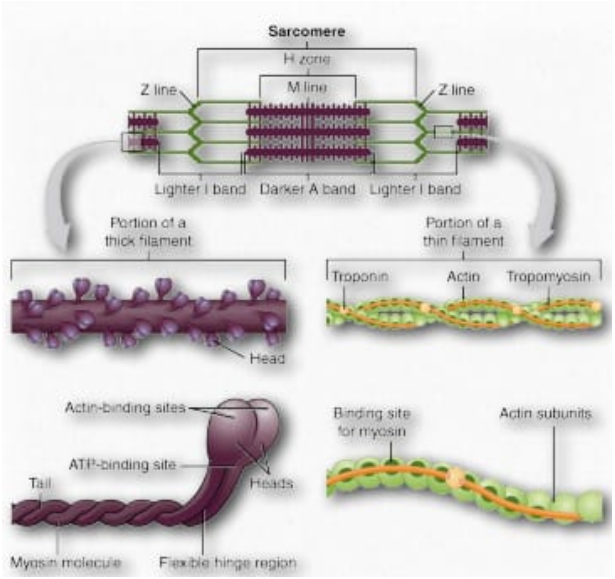


Figure 2.122 - Sarcomere Anatomy Wikipedia

It should be noted that the sarcoplasmic reticulum is always taking up calcium. Only when its calcium gates are opened by the action potential is it unable to reduce cellular calcium concentration. As the action potential decreases, then the calcium gates close and the sarcoplasmic reticulum “catches up” and cellular calcium concentrations fall. At that point troponin releases calcium, tropomyosin goes back to covering myosin binding sites on actin, myosin loses its attachment to actin and the thin filaments slide back to their original positions relative to the myosin thick filaments.

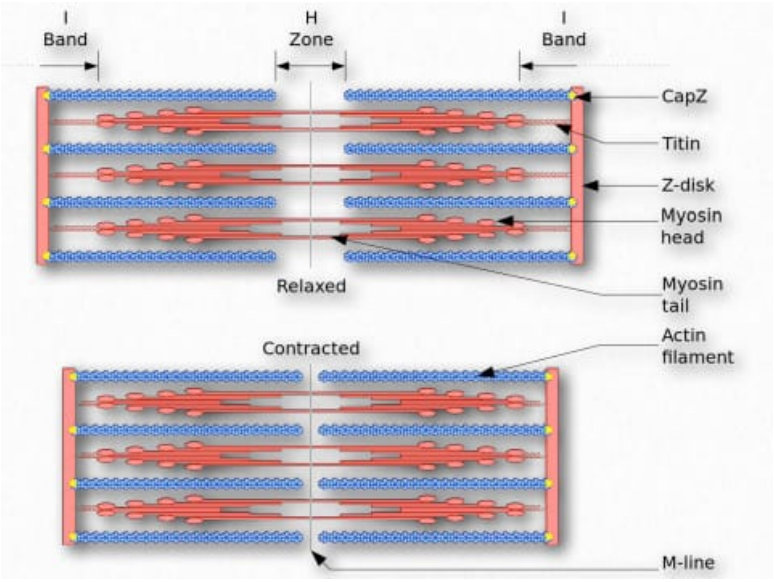


Figure 2.123 - The Sliding Filament Model of Muscular Contraction Wikipedia

Tropomyosin

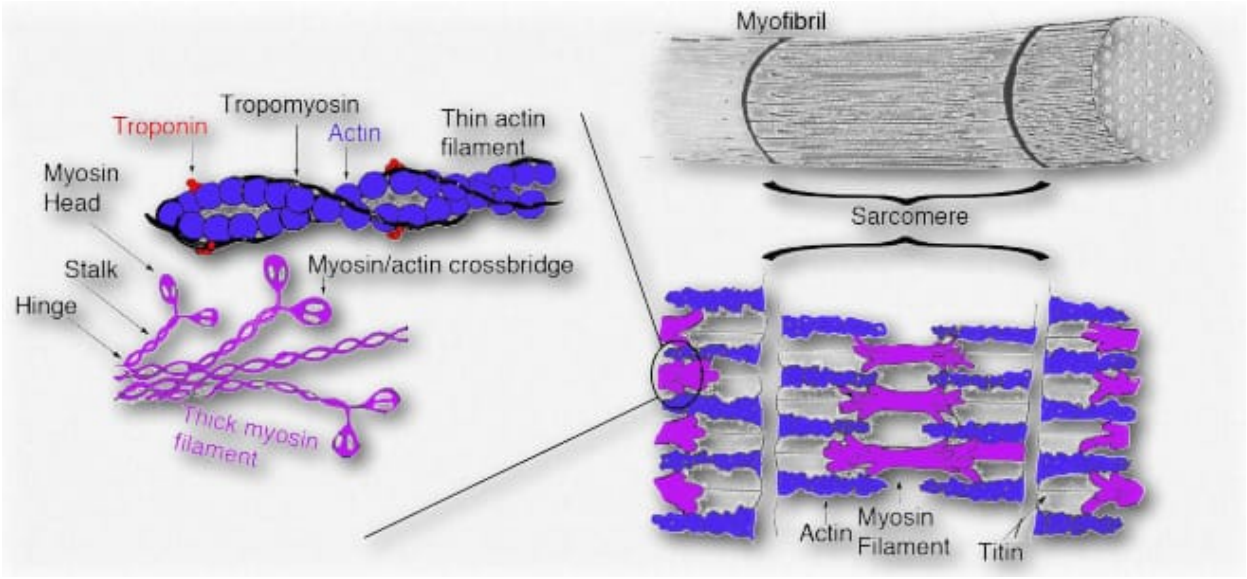


Figure 2.124 - Tropomyosin and troponin in muscle anatomy Wikipedia

Tropomyosins are proteins that interact with actin thin filaments to help regulate their roles in movement, both in muscle cells and non-muscle cells (Figure 2.124). Tropomyosins interact to form head-to-toe dimers and perch along the α-helical groove of an actin filament. The isoforms of tropomyosin that are in muscle cells control interactions between myosin and the actin filament within the sarcomere and help to regulate contraction of the muscle. In other cells, nonmuscle tropomyosins help to regulate the cytoskeleton's functions.

The interactions of tropomyosin with the cytoskeleton are considerably more complicated than what occurs in muscle cells. Muscle cells have five tropomyosin isoforms, but in the cytoskeleton of non-muscle cells, there are over 40 tropomyosins.

Troponin

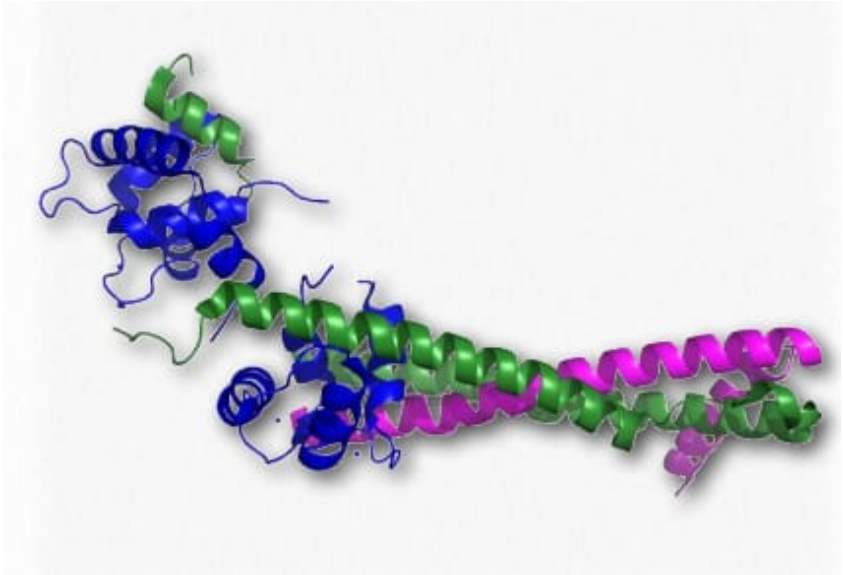


Figure 2.125 - Troponin complex of muscle. Blue = troponin C, magenta = troponin T , green = troponin I

The troponins involved in muscular contraction are actually a complex of three proteins known as troponin I, troponin C, and troponin T (Figure 2.125). They associate with each other and with tropomyosin on actin filaments to help regulate the process of muscular contraction. Troponin I prevents binding of myosin's head to actin and thus prevents the most important step in contraction.

Troponin C is a unit that binds to calcium ions. Troponin T is responsible for binding all three proteins to tropomyosin. Troponins

in the bloodstream are indicative of heart disorders. Elevation of troponins in the blood occurs after a myocardial infarction and can remain high for up to two weeks.

Actinin

Actinin is a skeletal muscle protein that attaches filaments of actin to Z-lines of skeletal muscle cells. In smooth muscle cells, it also connects actin to dense bodies.

Titin

Titin (also known as connectin) is the molecular equivalent of a spring that provides striated muscle cells with elasticity. It is the third most abundant protein in muscle cells. The protein is enormous, with 244 folded individual protein domains spread across 363 exons (largest known number), with the largest known exon (17,106 base pairs long), and it is the largest protein known (27,000 to 33,000 amino acids, depending on splicing).

Unstructured sequences

The folded protein domains are linked together by unstructured sequences. The unstructured regions of the protein allow for unfolding when stretching occurs and refolding upon relaxation. Titin connects the M and Z lines in the sarcomere (Figure 2.123). Tension created in titin serves to limit the range of motion of the sarcomere, giving rise to what is called passive stiffness.

Skeletal and cardiac muscles have slight amino acid sequence variations in their ti tin proteins and these appear to relate to differences in the mechanical characteristics of each muscle.

Energy backup for muscle energy

Myoglobin was described as a molecular batter for oxygen. Muscle cells have a better of their own for ATP. The is important for animals, but not for plants because a plant's need for energy is different than an animal's. Plants do not need to access energy sources as rapidly as animals do, nor do they have to maintain a constant internal temperature. Plants can neither flee predators, nor chase prey. These needs of animals are much more immediate and require that energy stores be accessible on demand. Muscles, of course, enable the motion of animals and the energy required for muscle contraction is ATP. To have stores of energy readily available, muscles have, in addition to ATP, creatine phosphate for energy and glycogen for quick release of glucose to make more energy. The synthesis of creatine phosphate is a prime example of the effects of concentration on the synthesis of high energy molecules. For example, creatine phosphate has an energy of hydrolysis of -43.1 kJ/mol whereas ATP has an energy of hydrolysis of -30.5 kJ/mol. Creatine phosphate, however, is made from creatine and ATP in the reaction shown in Figure 2.126. How is this possible?

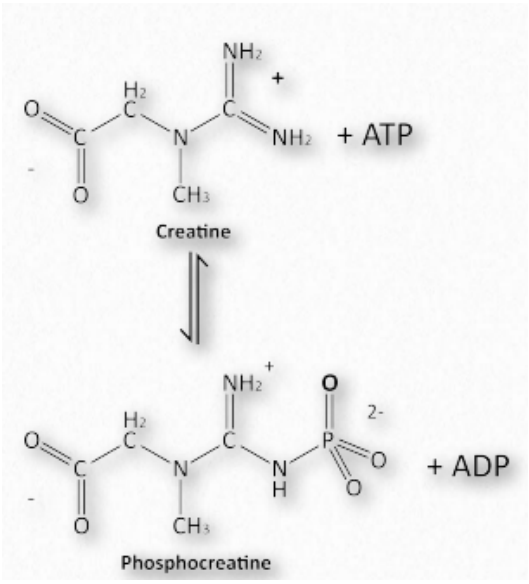


Figure 2.126 - Phosphorylation of creatine (phosphocreatine) - making of a creatine phosphate battery Image by Aleia Kim

The ΔG°' of this reaction is +12.6 kJ/mol, reflecting the energies noted above. In a resting muscle cell, ATP is abundant and ADP

is low, driving the reaction downward, creating creatine phosphate. When muscular contraction commences, ATP levels fall and ADP levels climb. The above reaction then reverses and proceeds to synthesize ATP immediately. Thus, creatine phosphate acts like a battery, storing energy when ATP levels are high and releasing it almost instantaneously to create ATP when its levels fall.

This page titled 2.5: Structure and Function- Protein Function II is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.6: Structure and Function - Nucleic Acids

Source: BiochemFFA_2_5.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

The nucleic acids, DNA and RNA, may be thought of as the information molecules of the cell. In this section, we will examine the structures of DNA and RNA, and how these structures are related to the functions these molecules perform.



We will begin with DNA, which is the hereditary information in every cell, that is copied and passed on from generation to generation. The race to elucidate the structure of DNA was one of the greatest stories of 20th century science. Discovered in 1869 by Friedrich Miescher, DNA was identified as the genetic material in experiments in the 1940s led by Oswald Avery, Colin MacLeod, and Maclyn McCarty. X-ray diffraction work of Rosalind Franklin and the observations of Erwin Chargaff were combined by James Watson and Francis Crick to form a model of DNA that we are familiar with today. Their famous paper, in the April 25, 1953 issue of Nature, opened the modern era of molecular biology. Arguably, that one-page paper has had more scientific impact per word than any other research article ever published. Today, every high school biology student is familiar with the double helical structure of DNA and knows that G pairs with C and A with T.

The double helix, made up of a pair of DNA strands, has at its core, bases joined by hydrogen bonds to form base pairs - adenine always paired with thymine, and guanine invariably paired with cytosine. Two hydrogen bonds are formed between adenine and thymine, but three hydrogen bonds hold together guanine and cytosine (Figure 2.127).

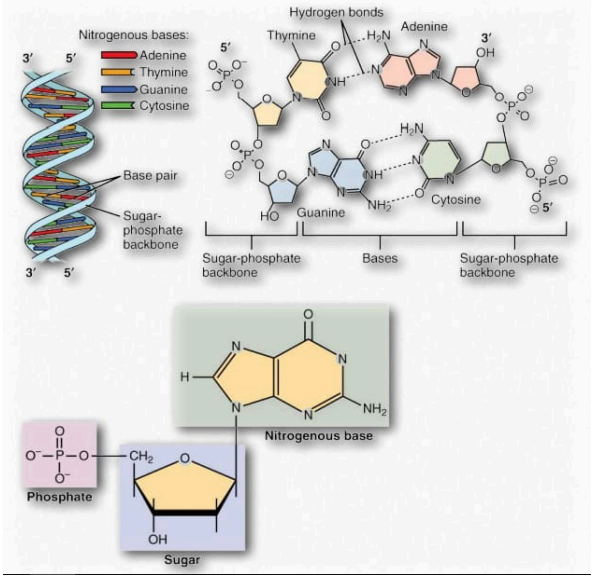


Figure 2.127 - A DNA duplex with base pairs, a closeup of base pairing, and a closeup of a nucleotide Wikipedia

The complementary structure immediately suggested to Watson and Crick how DNA might be (and in fact, is) replicated and it further explains how information is DNA is transmitted to RNA for the synthesis of proteins. In addition to the hydrogen bonds between bases of each strand, the double helix is held together by hydrophobic interactions of the stacked, non-polar bases. Crucially, the sequence of the bases in DNA carry the information for making proteins. Read in groups of three, the sequence of the bases directly specifies the sequence of the amino acids in the encoded protein.

Structure

A DNA strand is a polymer of nucleoside monophosphates held together by phosphodiester bonds. Two such paired strands make up the DNA molecule, which is then twisted into a helix. In the most common Bform, the DNA helix has a repeat of 10.5 base pairs per turn, with sugars and phosphate forming the covalent phosphodiester “backbone” of the molecule and the adenine, guanine, cytosine, and thymine bases oriented in the middle where they form the now familiar base pairs that look like the rungs of a ladder.

Building blocks

The term nucleotide refers to the building blocks of both DNA (deoxyribonucleoside triphosphates, dNTPs) and RNA (ribonucleoside triphosphates, NTPs). In order to discuss this important group of molecules, it is necessary to define some terms.

Nucleotides contain three primary structural components. These are a nitrogenous base, a pentose sugar, and at least one phosphate. Molecules that contain only a sugar and a nitrogenous base (no phosphate) are called nucleosides. The nitrogenous bases found in nucleic acids include adenine and guanine (called purines) and cytosine, uracil, or thymine (called pyrimidines). There are two sugars found in nucleotides - deoxyribose and ribose (Figure 2.128). By convention, the carbons on these sugars are labeled 1’ to 5’. (This is to distinguish the carbons on the sugars from those on the bases, which have their carbons simply labeled as 1, 2, 3, etc.) Deoxyribose differs from ribose at the 2’ position, with ribose having an OH group, where deoxyribose has H.

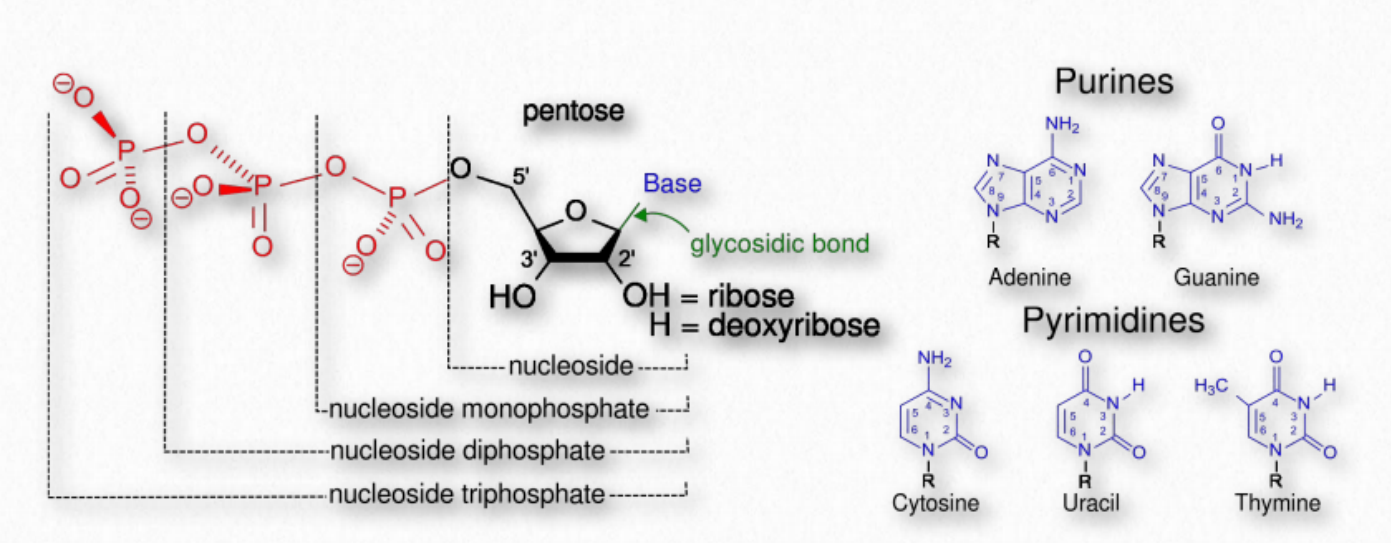


Figure 2.128 - Nucleotides, nucleosides, and bases

Nucleotides containing deoxyribose are called deoxyribonucleotides and are the forms found in DNA. Nucleotides containing ribose are called ribonucleotides and are found in RNA. Both DNA and RNA contain nucleotides with adenine, guanine, and cytosine, but with very minor exceptions, RNA contains uracil nucleotides, whereas DNA contains thymine nucleotides. When a base is attached to a sugar, the product, a nucleoside, gains a new name.

- **uracil-containing** = uridine (attached to ribose) / deoxyuridine (attached to deoxyribose)
- **thymine-containing** = ribothymidine (attached to ribose) / thymidine (attached to deoxyribose)
- **cytosine-containing** = cytidine (attached to ribose - Figure 2.129) / deoxycytidine (attached to deoxyribose)
- **guanine-containing** = guanosine (attached to ribose) / deoxyguanosine (attached to deoxyribose)
- **adenine-containing** = adenosine (attached to ribose) / deoxyadenosine (attached to deoxyribose)

Of these, deoxyuridine and ribothymidine are the least common. The addition of one or more phosphates to a nucleoside makes it a nucleotide. Nucleotides are often referred to as nucleoside phosphates, for this reason. The number of phosphates in the nucleotide is indicated by the appropriate prefixes (mono, di or tri).

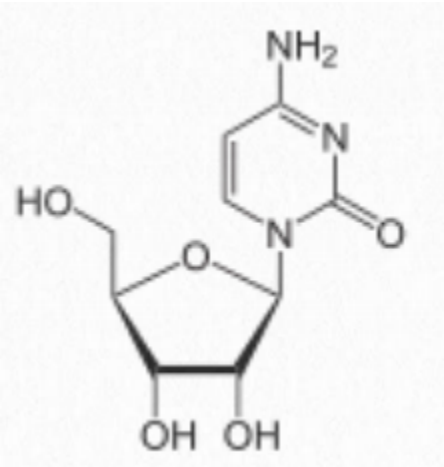


Figure 2.129 Cytidine

Thus, cytidine, for example, refers to a nucleoside (no phosphate), but cytidine monophosphate refers to a nucleotide (with one phosphate). Addition of second and third phosphates to a nucleoside monophosphate requires energy, due to the repulsion of negatively charged phosphates and this chemical energy is the basis of the high energy triphosphate nucleotides (such as ATP) that fuel cells.

Note: Ribonucleotides as Energy Sources

Though ATP is the most common and best known cellular energy source, each of the four ribonucleotides plays important roles in providing energy. GTP, for example, is the energy source for protein synthesis (translation) as well as for a handful of metabolic reactions. A bond between UDP and glucose makes UDP-glucose, the building block for making glycogen. CDP is similarly linked to several different molecular building blocks important for glycerophospholipid synthesis (such as CDP-diacylglycerol).

The bulk of ATP made in cells is not from directly coupled biochemical metabolism, but rather by the combined processes of electron transport and oxidative phosphorylation in mitochondria and/or photophosphorylation that occurs in the chloroplasts of photosynthetic organisms. Triphosphate energy in ATP is transferred to the other nucleosides/nucleotides by action of enzymes called kinases. For example, nucleoside diphosphokinase (NDPK) catalyzes the following reaction

$$\mathrm{ATP} + \mathrm{NDP} \longleftrightarrow \mathrm{ADP} + \mathrm{NTP} \tag{2.6.1}$$

where 'N' of "NDP" and "NTP corresponds to any base. Other kinases can put single phosphates onto nucleosides or onto nucleoside monophosphates using energy from ATP.

Deoxyribonucleotides

Individual deoxyribonucleotides are derived from corresponding ribonucleoside diphosphates via catalysis by the enzyme known as ribonucleotide reductase (RNR). The deoxyribonucleoside diphosphates are then converted to the corresponding triphosphates (dNTPs) by the addition of a phosphate group. Synthesis of nucleotides containing thymine is distinct from synthesis of all of the other nucleotides and will be discussed later.

Building DNA strands

Each DNA strand is built from dNTPs by the formation of a phosphodiester bond, catalyzed by DNA polymerase, between the 3'OH of one nucleotide and the 5' phosphate of the next. The result of this directional growth of the strand is that the one end of the strand has a free 5' phosphate and the other a free 3' hydroxyl group (Figure 2.130). These are designated as the 5' and 3' ends of the strand.

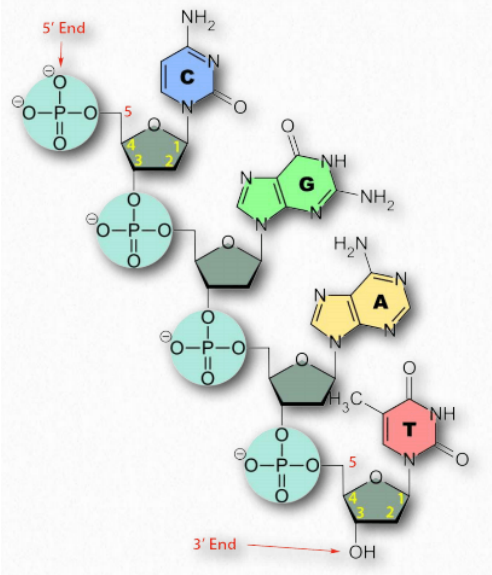


Figure 2.130 - 5'-3' Polarity of a DNA strand

Figure 2.131 shows two strands of DNA (left and right). The strand on the left, from 5' to 3' reads T-C-G-A, whereas the strand on the right, reading from 5' to 3' is T-C-G-A. The strands in a double-stranded DNA are arranged in an anti-parallel fashion with the 5' end of one strand across from the 3' end of the other.

Hydrogen bonds

Hydrogen bonds between the base pairs hold a nucleic acid duplex together, with two hydrogen bonds per A-T pair (or per A-U pair in RNA) and three hydrogen bonds per G-C pair. The B-form of DNA has a prominent major groove and a minor groove tracing the path of the helix (Figure 2.132). Proteins, such as transcription factors bind in these grooves and access the hydrogen bonds of the base pairs to “read” the sequence therein.

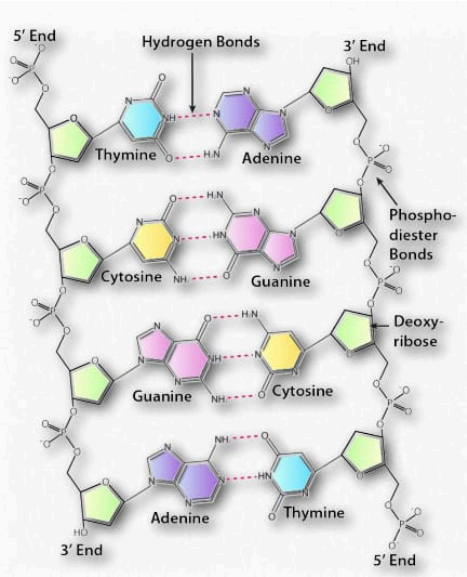


Figure 2.131 - Anti-parallel orientation of a DNA duplex, phosphodiester backbone, and base pairing Image by Aleia Kim

Other forms of DNA besides the B-form (Movie 2.5) are known (Figure 2.133). One of these, the A-form, was identified by Rosalind Franklin in the same issue of Nature as Watson and Crick’s paper. Though the A-form structure is a relatively minor form of DNA and resembles the B-form, it turns out to be important in the duplex form of RNA and in RNA-DNA hybrids. Both the A form and the B-form of DNA have the helix oriented in what is termed the right-handed form.

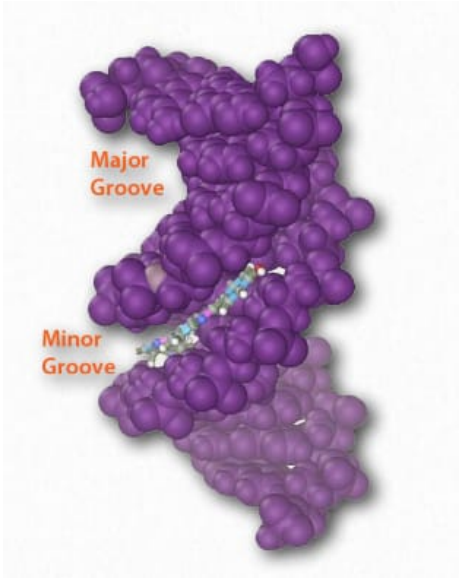
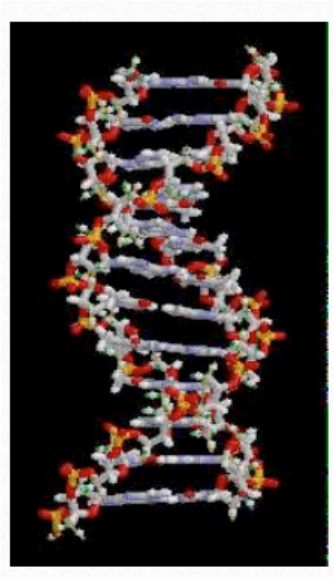


Figure 2.132 - Major and minor grooves of DNA. The minor groove has been bound by a dye Wikipedia



Movie 2.5 - B-form DNA duplex rotating in space Wikipedia

Z-DNA

The A-form and the B-form stand in contrast to another form of DNA, known as the Z-form. ZDNA, as it is known, has the same base-pairing rules as the B and A forms, but instead has the helices twisted in the opposite direction, making a left-handed helix (Figure 2.133). The Z-form has a sort of zig-zag shape, giving rise to the name Z-DNA.

In addition, the helix is rather stretched out compared to the A- and B-forms. Why are there different topological forms of DNA? The answer relates to both superhelical tension and sequence bias. Sequence bias means that certain sequences tend to favor the “flipping” of Bform DNA into other forms. ZDNA forms are favored by long stretches of alternating Gs and Cs. Superhelical tension is discussed below.

Superhelicity



Figure 2.133 - From left to right, the A-, B-, and Zforms of DNA

Short stretches of linear DNA duplexes exist in the B-form and have 10.5 base pairs per turn. Double helices of DNA in the cell can vary in the number of base pairs per turn they contain. There are several reasons for this. For example, during DNA replication,

strands of DNA at the site of replication get unwound at the rate of 6000 rpm by an enzyme called helicase. The effect of such local unwinding at one place in a DNA has the effect increasing the winding ahead of it. Unrelieved, such 'tension' in a DNA duplex can result in structural obstacles to replication.

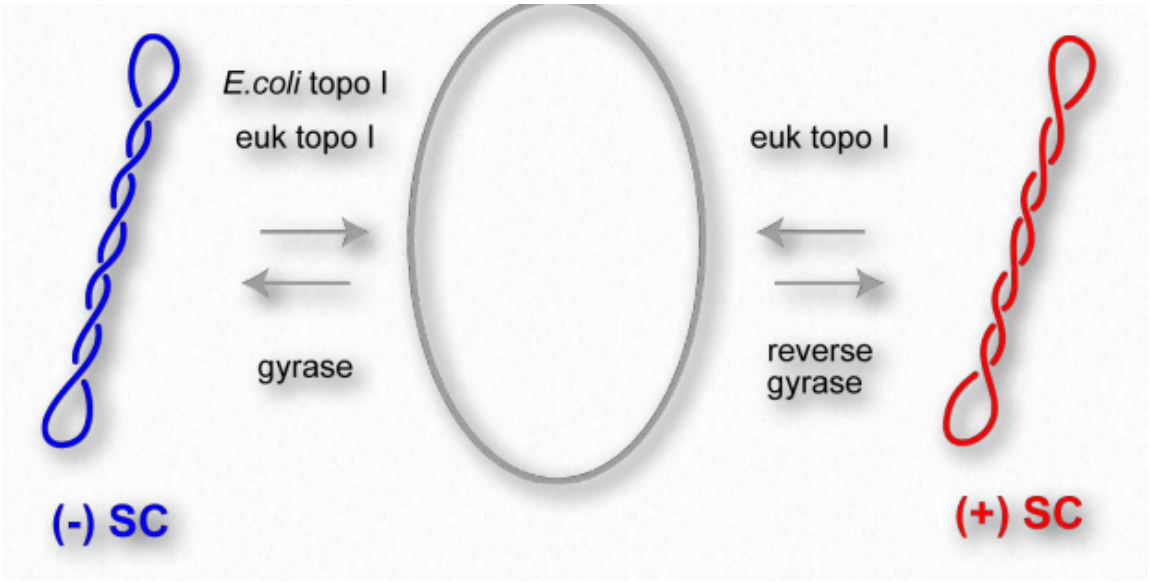


Figure 2.134 - Negatively supercoiled, relaxed circular, and positively supercoiled DNA interconverted by topoisomerase enzymes Wikipedia

Such adjustments can occur in three ways. First, tension can provide the energy for 'flipping' DNA structure. Z-DNA can arise as a means of relieving the tension. Second, DNA can 'supercoil' to relieve the tension (Figures 2.134 & 2.135). In this method, the duplex crosses over itself repeatedly, much like a rubber band will coil up if one holds one section in place and twists another part of it. Third, enzymes called topoisomerases can act to relieve or, in some cases, increase the tension by adding or removing twists in the DNA.

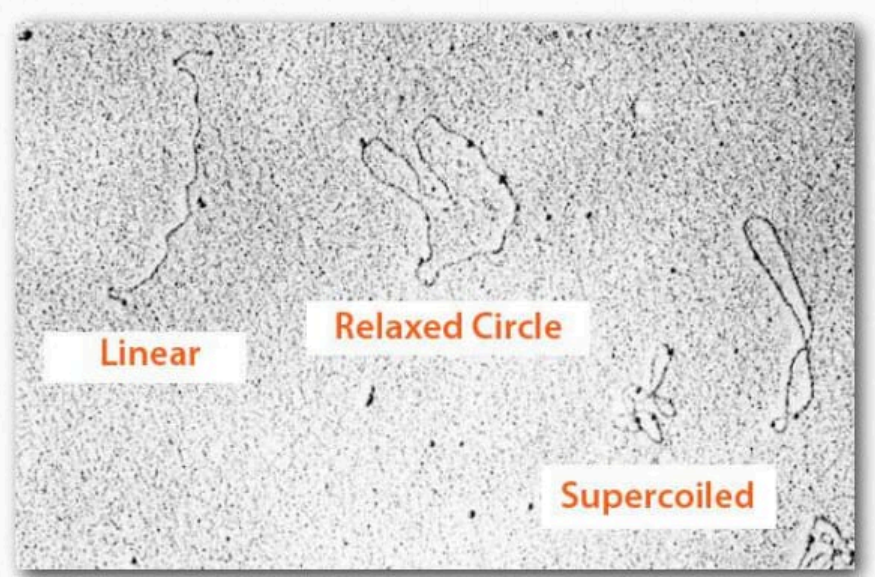


Figure 2.135 - Three topological forms of DNA Wikipedia

Topological isomers

As noted, so-called "relaxed" DNA has 10.5 base pairs per turn. Each turn corresponds to one twist of the DNA. Using enzymes, it is possible to change the number of base pairs per turn. In either the case of increasing or decreasing the twists per turn, tension is introduced into the DNA structure. If the tension cannot be relieved, the DNA duplex will act to relieve the strain, as noted. This is most easily visualized for circular DNA, though long linear DNA (such as found in eukaryotic chromosomes) or DNAs

constrained in other ways will exhibit the same behavior.

Parameters

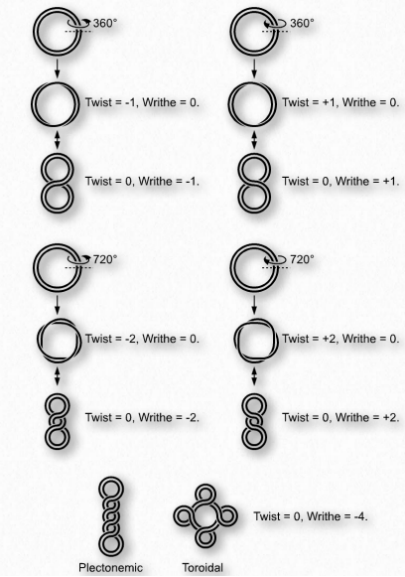


Figure 2.136 - Twists and writhes - DNA duplex coiling omitted for simplicity Wikipedia

To understand topologies, we introduce the concepts of 'writhe' and 'linking number'. First, imagine either opening a closed circle of DNA and either removing one twist or adding one twist and then re-forming the circle. Since the strands have no free ends, they cannot relieve the induced tension by re-adding or removing the twists at their ends, respectively. Instead, the tension is relieved by "superhelices" that form with crossing of the double strands over each other (figure 8 structures in Figure 2.136). Though it is not apparent to visualize, each crossing of the double strands in this way allows twists to be increased or decreased correspondingly. Thus, superhelicity allows the double helix to reassume 10.5 base pairs per turn by adding or subtracting twists as necessary and replacing them with writhes.

We write the equation L= T + W where T is the number of twists in a DNA, W is the number of writhes, and L is the linking number. The linking number is therefore the sum of the twists and writhes. Interestingly, inside of cells, DNAs typically are in a supercoiled form. Supercoiling affects the size of the DNA (compacts it) and also the expression of genes within the DNA, some having enhanced expression and some having reduced expression when supercoiling is present. Enzymes called topoisomerases alter the superhelical density of DNAs and play roles in DNA replication, transcription, and control of gene expression. They work by making cuts in one strand (Type I topoisomerases) or both strands (Type II topoisomerases) and then add or subtract twists as appropriate to the target DNA. After that process is complete, the topoisomerase re-ligates the nick/cut it had made in the DNA in the first step.

Topoisomerases may be the targets of antibiotics. The family of antibiotics known as fluoroquinolones work by interfering with the action of bacterial type II topoisomerases. Ciprofloxacin also preferentially targets bacterial type II topoisomerases. Other topoisomerase inhibitors target eukaryotic topoisomerases and are used in anti-cancer treatments.

RNA

The structure of RNA (Figure 2.137) is very similar to that of a single strand of DNA. Built of ribonucleotides, joined together by the same sort of phosphodiester bonds as in DNA, RNA uses uracil in place of thymine. In cells, RNA is assembled by RNA polymerases, which copy a DNA template in the much same way that DNA polymerases replicate a parental strand. During the synthesis of RNA, uracil is used across from an adenine in the DNA template. The building of messenger RNAs by copying a DNA template is a crucial step in the transfer of the information in DNA to a form that directs the synthesis of protein. Additionally, ribosomal and transfer RNAs serve important roles in "reading" the information in the mRNA codons and in polypeptide synthesis. RNAs are also known to play important roles in the regulation of gene expression.

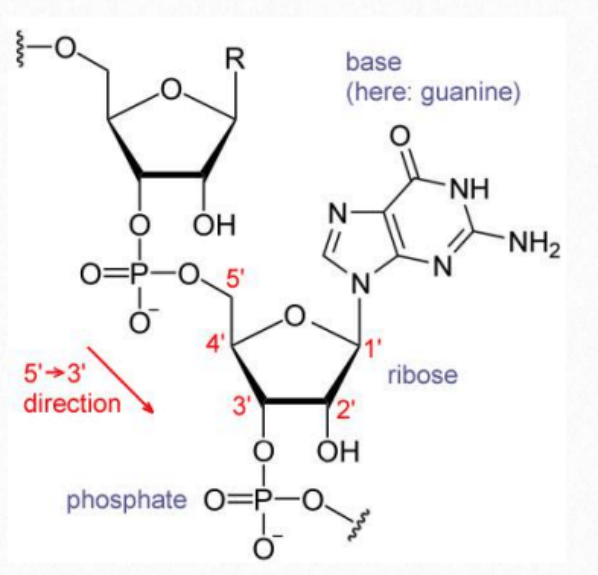


Figure 2.137 A section of an RNA molecule Wikipedia

RNA world

The discovery, in 1990, that RNAs could play a role in catalysis, a function once thought to be solely the domain of proteins, was followed by the discovery of many more so-called ribozymes- RNAs that functioned as enzymes. This suggested the answer to a long-standing chicken or egg puzzle - if DNA encodes proteins, but the replication of DNA requires proteins, how did a replicating system come into being? This problem could be solved if the first replicator was RNA, a molecule that can both encode information and carry out catalysis. This idea, called the “RNA world” hypothesis, suggests that DNA as genetic material and proteins as catalysts arose later, and eventually prevailed because of the advantages they offer. The lack of a 2’OH on deoxyribose makes DNA more stable than RNA. The double-stranded structure of DNA also provides an elegant way to easily replicate it. RNA catalysts, however, remain, as remnants of that early world. In fact, the formation of peptide bonds, essential for the synthesis of proteins, is catalyzed by RNA.

Secondary structure



Figure 2.138 - tRNA Images - 3D projection (left) and 2D projection (inset) Wikipedia

With respect to structure, RNAs are more varied than their DNA cousin. Created by copying regions of DNA, cellular RNAs are synthesized as single strands, but they often have self-complementary regions leading to “foldbacks” containing duplex regions. These are most easily visualized in the ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs) (Figure 2.138), though other

RNAs, including messenger RNAs (mRNAs), small nuclear RNAs (snRNAs), microRNAs (Figure 2.139), and small interfering RNAs (siRNAs) may each have double helical regions as well.



Figure 2.139 - MicroRNA stem loops. Wikipedia

Base pairing

Base pairing in RNA is slightly different than DNA. This is due to the presence of the base uracil in RNA in place of thymine in DNA. Like thymine, uracil forms base pairs with adenine, but unlike thymine, uracil can, to a limited extent, also base pair with guanine, giving rise to many more possibilities for pairing within a single strand of RNA.

These additional base pairing possibilities mean that RNA has many ways it can fold upon itself that single-stranded DNA cannot. Folding, of course, is critical for protein function, and we now know that, like proteins, some RNAs in their folded form can catalyze reactions just like enzymes. Such RNAs are referred to as ribozymes. It is for this reason scientists think that RNA was the first genetic material, because it could not only carry information, but also catalyze reactions. Such a scheme might allow certain RNAs to make copies of themselves, which would, in turn, make more copies of themselves, providing a positive selection.

Stability

RNA is less chemically stable than DNA. The presence of the 2' hydroxyl on ribose makes RNA much more prone to hydrolysis than DNA, which has a hydrogen instead of a hydroxyl. Further, RNA has uracil instead of thymine. It turns out that cytosine is the least chemically stable base in nucleic acids. It can spontaneously deaminate and in turn is converted to a uracil. This reaction occurs in both DNA and RNA, but since DNA normally has thymine instead of uracil, the presence of uracil in DNA indicates that deamination of cytosine has occurred and that the uracil needs to be replaced with a cytosine. Such an event occurring in RNA would be essentially undetectable, since uracil is a normal component of RNA. Mutations in RNA have much fewer consequences than mutations in DNA because they are not passed between cells in division.

Catalysis

RNA structure, like protein structure, has importance, in some cases, for catalytic function. Like random coils in proteins that give rise to tertiary structure, single-stranded regions of RNA that link duplex regions give these molecules a tertiary structure, as well. Catalytic RNAs, called ribozymes, catalyze important cellular reactions, including the formation of peptide bonds in ribosomes (Figure 2.114). DNA, which is usually present in cells in strictly duplex forms (no tertiary structure, per se), is not known to be involved in catalysis.

RNA structures are important for reasons other than catalysis. The 3D arrangement of tRNAs is necessary for enzymes that attach amino acids to them to do so properly. Further, small RNAs called siRNAs found in the nucleus of cells appear to play roles in both gene regulation and in cellular defenses against viruses. The key to the mechanisms of these actions is the formation of short foldback RNA structures that are recognized by cellular proteins and then chopped into smaller units. One strand is copied and used to base pair with specific mRNAs to prevent the synthesis of proteins from them.

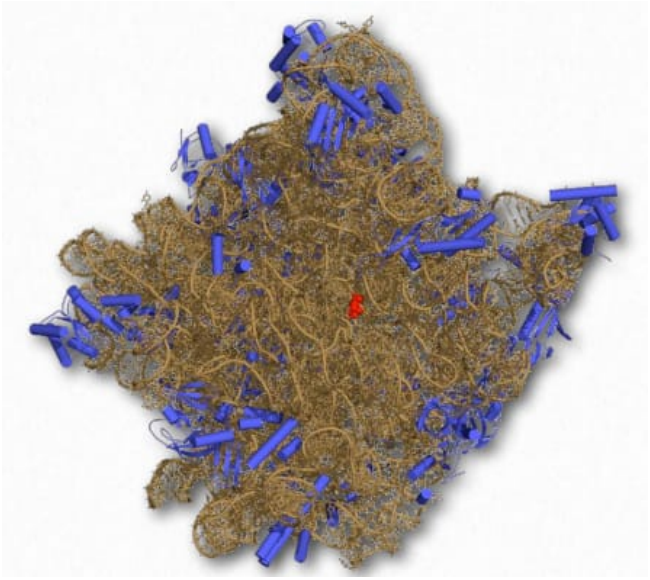


Figure 2.140 - Structure of the 50S ribosomal subunit. rRNA shown in brown. Active site in red Wikipedia

Denaturing nucleic acids

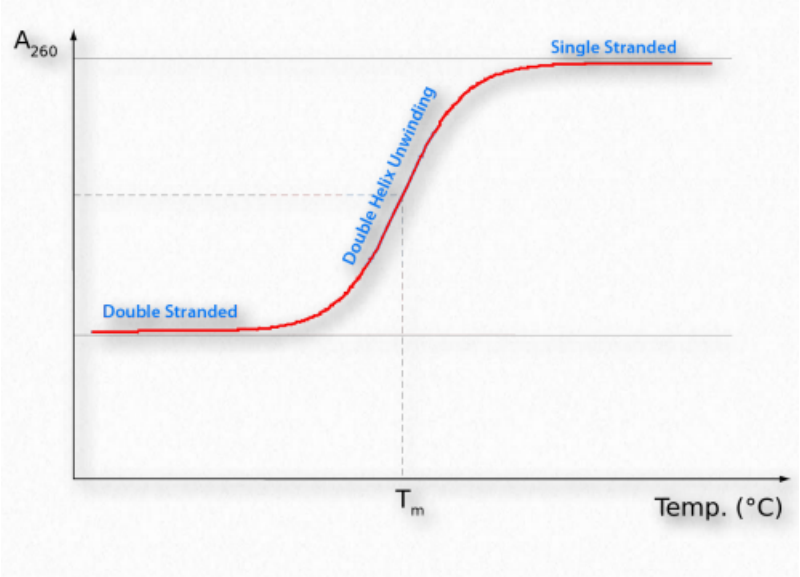


Figure 2.141 - The hyperchromic effect Wikipedia

Like proteins, nucleic acids can be denatured. Forces holding duplexes together include hydrogen bonds between the bases of each strand that, like the hydrogen bonds in proteins, can be broken with heat or urea. (Another important stabilizing force for DNA arises from the stacking interactions between the bases in a strand.) Single strands absorb light at 260 nm more strongly than double strands. This is known as the hyperchromic effect (Figure 2.141)and is a consequence of the disruption of interactions among the stacked bases. The changes in absorbance allow one to easily follow the course of DNA denaturation. Denatured duplexes can readily renature when the temperature is lowered below the “melting temperature” or Tm, the temperature at which half of the DNA strands are in duplex form. Under such conditions, the two strands can re-form hydrogen bonds between the complementary sequences, returning the duplex to its original state. For DNA, strand separation and rehybridization are important for the technique known as the polymerase chain reaction (PCR). Strand separation of DNA duplexes is accomplished in the method by heating them to boiling. Hybridization is an important aspect of the method that requires single stranded primers to “find” matching sequences on the template DNA and form a duplex. Considerations for efficient hybridization (also called annealing) include temperature, salt concentration, strand concentration, and magnesium ion levels (for more on PCR, see HERE).

DNA packaging

DNA is easily the largest macromolecule in a cell. The single chromosome in small bacterial cells, for example, can have a molecular weight of over 1 billion Daltons. If one were to take all of the DNA of human chromosomes from a single cell and lay them end to end, they would be over 7 feet long. Such an enormous molecule demands careful packaging to fit within the confines of a nucleus (eukaryotes) or a tiny cell (bacteria). The chromatin system of eukaryotes is the best known, but bacteria, too, have a system for compacting DNA.

DNA in Bacteria

In bacteria, there is no nucleus for the DNA. Instead, DNA is contained in a structure called a nucleoid (Figure 2.142). It contains about 60% DNA with much of the remainder comprised of RNAs and transcription factors. Bacteria do not have histone proteins that DNA wrap around, but they do have proteins that help organize the DNA in the cell - mostly by making looping structures.

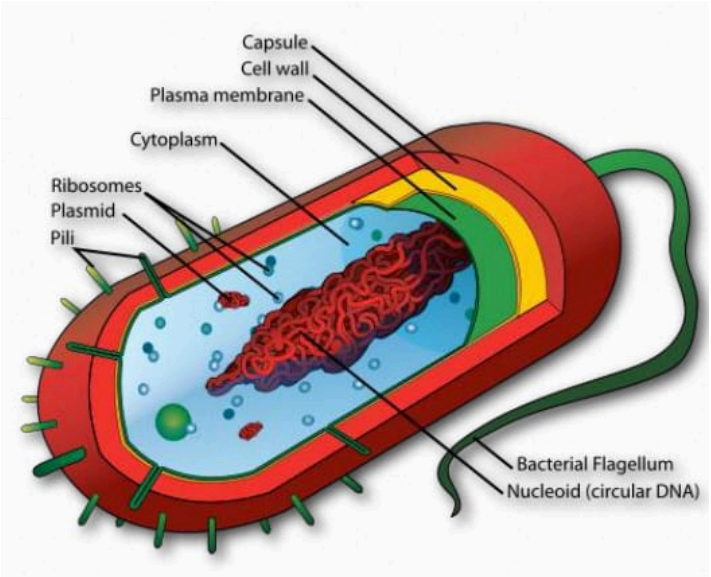


Figure 2.142 - Anatomy of a bacterial cell.

These proteins are known as Nucleoid Associated Proteins and include ones named HU, H-NS, Fis, CbpA, and Dps. Of these, HU most resembles eukaryotic histone H2B and binds to DNA non-specifically. The proteins associate with the DNA and can also cluster, which may be the origin of the loops. It is likely these proteins play a role in helping to regulate transcription and respond to DNA damage. They may also be involved in recombination.

Eukaryotes

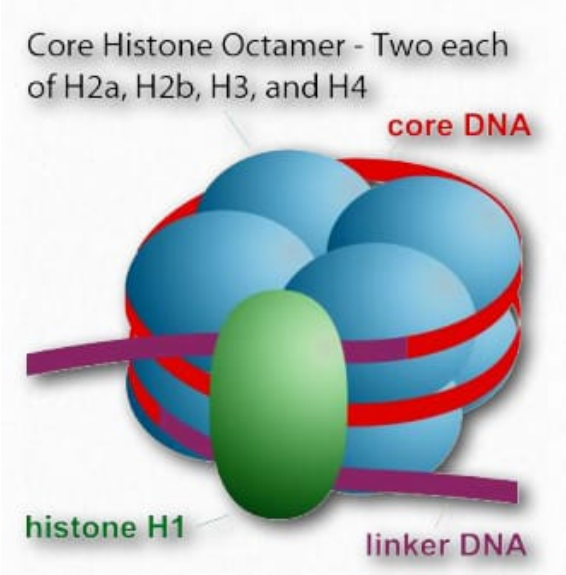


Figure 2.143 - Structure of a nucleosome Wikipedia

The method eukaryotes use for compacting DNA in the nucleus is considerably different, and with good reason - eukaryotic DNAs are typically much larger than prokaryotic DNAs, but must fit into a nucleus that is not much bigger than a prokaryotic cell. Human DNA, for example, is about 1000 times longer than c DNA. The strategy employed in eukaryotic cells is that of spooling - DNA is coiled around positively charged proteins called histones. These proteins, whose sequence is very similar in cells as diverse as yeast and humans, come in four types, dubbed H1, H2a, H2b, H3, and H4. A sixth type, referred to as H5 is actually an isoform of H1 and is rare. Two each of H2a, H2b, H3, and H4 are found in the core structure of what is called the fundamental unit of chromatin - the nucleosome (Figure 2.143).

Octamer

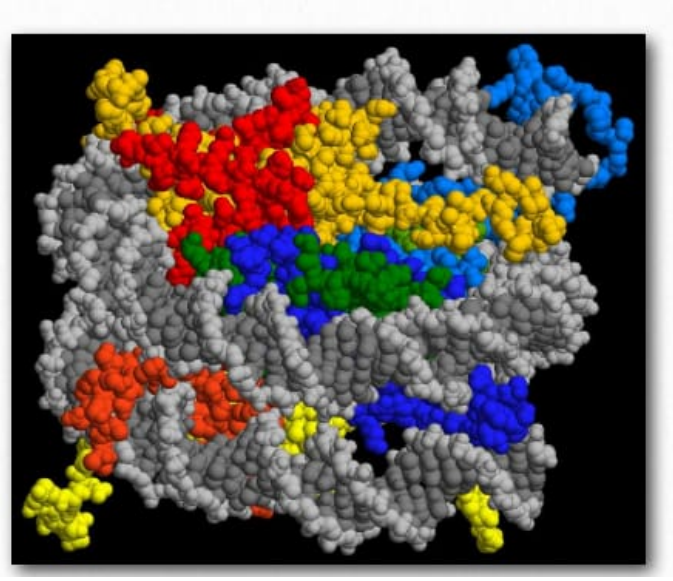


Figure 2.144 - Detailed core particle structure - DNA in gray, histones in red, yellow, green, blue Wikipedia

The core of 8 proteins is called an octamer. The stretch of DNA wrapped around the octamer totals about 147 base pairs and makes 1 2/3 turns around it. This complex is referred to as a core particle (Figure 2.144). A linker region of about 50-80 base pairs separate core particles. The term nucleosome then refers to a a core particle plus a linker region (Figure 2.143). Histone H1 sits near the junction of the incoming DNA and the histone core. It is often referred to as the linker histone. In the absence of H1, non-condensed nucleosomes resemble “beads on a string” when viewed in an electron microscope.

Histones



Figure 2.145 - Amino acid sequence (1- letter code) of histone H3 of S. cerevisiae. Arginines (R) and lysines (K) shown in red.

Histone proteins are similar in structure and are rich in basic amino acids, such as lysine and arginine (Figure 2.145). These amino acids are positively charged at physiological pH, with enables them to form tight ionic bonds with the negatively charged phosphate backbone of DNA.

For DNA, compression comes at different levels (Figure 2.146). The first level is at the nucleosomal level. Nucleosomes are stacked and coiled into higher order structures. 10 nm fibers are the simplest higher order structure (beads on a string) and they grow in complexity. 30 nm fibers consist of stacked nucleosomes and they are packed tightly. Higher level packing produces the metaphase chromosome found in meiosis and mitosis.

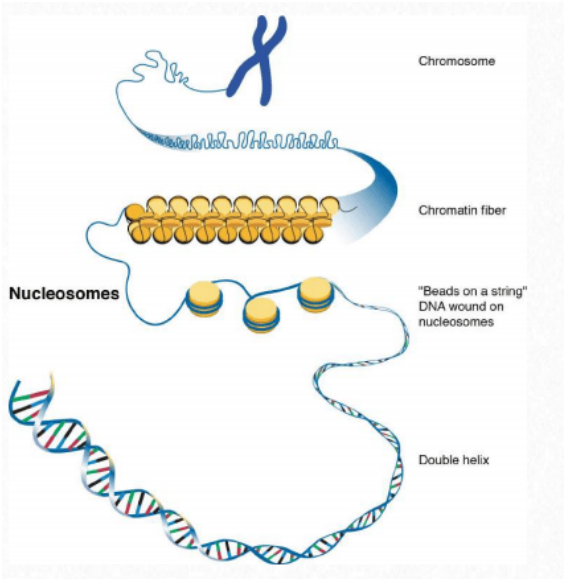


Figure 2.146 Packaging of DNA into a chromosome

The chromatin complex is a logistical concern for the processes of DNA replication and (particularly) gene expression where specific regions of DNA must be transcribed. Altering chromatin structure is therefore an essential function for transcriptional activation in eukaryotes. One strategy involves adding acetyl groups to the positively charged lysine side chains to “loosen their grip” on the negatively charged DNA, thus allowing greater access of proteins involved in activating transcription to gain access to the DNA. The mechanisms involved in eukaryotic gene expression are

Ames test

The Ames test (Figure 2.147) is an analytical method that allows one to determine whether a compound causes mutations in DNA (is mutagenic) or not. The test is named for Dr. Bruce Ames, a UC Berkeley emeritus professor who was instrumental in creating it. In the procedure, a single base pair of a selectable marker of an organism is mutated in a plasmid to render it nonfunctional. In the example, a strain of Salmonella is created that lacks the ability to grow in the absence of histidine. Without histidine, the organism will not grow, but if that one base in the plasmid’s histidine gene gets changed back to its original base, a functional gene will be made and the organism will be able to grow without histidine.

A culture of the bacterium lacking the functional gene is grown with the supply of histidine it requires. It is split into two vials. To one of the vials, a compound that one wants to test the mutagenicity of is added. To the other vial, nothing is added. The bacteria in each vial are spread onto plates lacking histidine. In the absence of mutation, no bacteria will grow. The more colonies of

bacteria that grow, the more mutation happened. Note that even the vial without the possible mutagenic compound will have a few colonies grow, as a result of mutations unlinked to the potential mutagen.

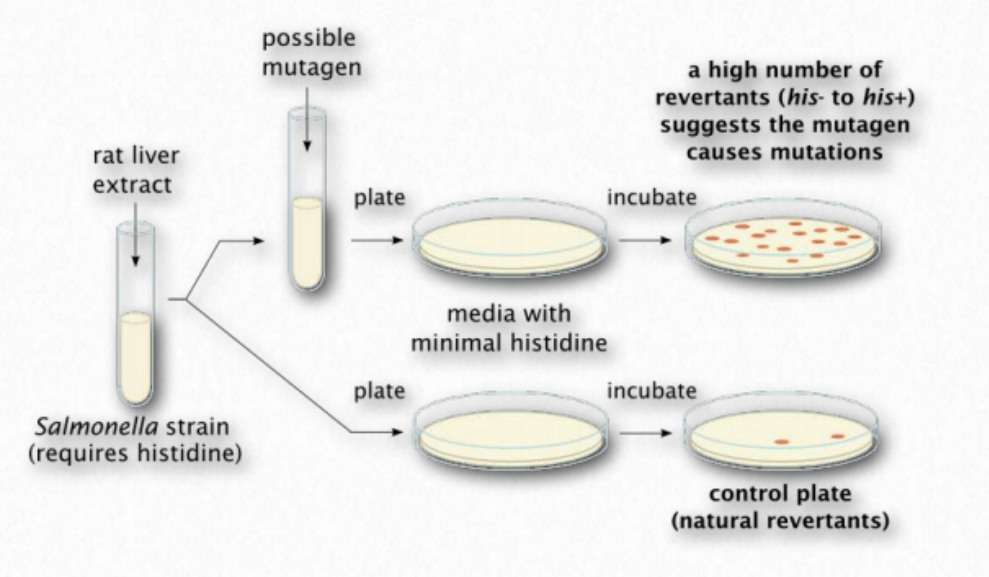


Figure 2.147 - The Ames test Wikipedia

Mutation happens in all cells at a low level. If the plate with the cells from the vial with the compound has more colonies than the cells from the control vial (no compound), then that would be evidence that the compound causes more mutations than would normally occur and it is therefore a mutagen. On the other hand, if there was no significant difference in the number of colonies on each plate, then that would suggest it is not mutagenic. The test is not perfect - it identifies about 90% of known mutagens - but its simplicity and inexpensive design make it an excellent choice for an initial screen of a compound.

This page titled 2.6: Structure and Function - Nucleic Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.7: Structure and Function- Carbohydrates



Figure 2.171 - Repeating unit of amylose

Endogenous glycation, on the other hand, arises with a frequency that is proportional to the concentration of free sugar in the body. These occur most frequently with fructose, galactose, and glucose in that decreasing order and are detected in the bloodstream. Both proteins and lipids can be glycated and the accumulation of endogenous advanced glycation endproducts (AGEs) is associated with Type 2 diabetes, as well as in increases in cardiovascular disease (damage to endothelium, cartilage, and fibrinogen), peripheral neuropathy (attack of myelin sheath), and deafness (loss of myelin sheath).

The formation of AGEs increases oxidative stress, but is also thought to be exacerbated by it. Increased oxidative stress, in turn causes additional harm. Damage to collagen in blood cells causes them to stiffen and weaken and is a factor in hardening of the arteries and formation of aneurysms, respectively. One indicator of diabetes is increased glycation of hemoglobin in red blood cells, since circulating sugar concentration are high in the blood of diabetics. Hemoglobin glycation is measured in testing for blood glucose control in diabetic patients.

Homopolymer	Monomeric Unit
Glycogen	Glucose
Cellulose	Glucose
Amylose	Glucose
Callose	Glucose
Chitin	N-acetylglucosamine
Xylan	Xylose
Mannan	Mannose
Chrysolaminarin	Glucose

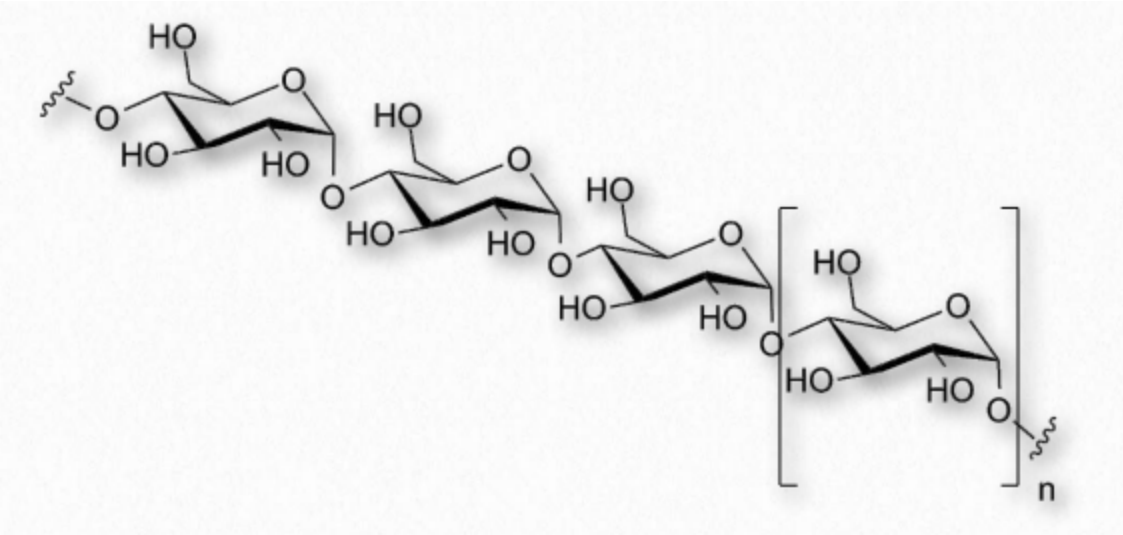


Figure 2.189). Along with the proteoglycan called lubricin, hyaluronic acid turns water into lubricating material. Hyaluronic acid is present as a coat around each cell of articular cartilage and forms complexes with proteoglycans that absorb water, giving resilience (resistance to compression) to cartilage. Aging causes a decrease in size of hyaluronans, but an increase in concentration.

Function in skin

Hyaluronic acid is a major component of skin and has functions in tissue repair. With exposure to excess UVB radiation, cells in the dermis produce less hyaluronan and increase its degradation.

For some cancers the plasma level of hyaluronic acid correlates with malignancy. Hyaluronic acid levels have been used as a marker for prostate and breast cancer and to follow disease progression. The compound can to used to induce healing after cataract surgery. Hyaluronic acid is also abundant in the granulation tissue matrix that replaces a fibrin clot during the healing of wounds. In wound healing, it is thought that large polymers of hyaluronic acid appear early and they physically make room for white blood cells to mediate an immune response.

Breakdown

Breakdown of hyaluronic acid is catalyzed by enzymes known as hyaluronidases. Humans have seven types of such enzymes, some of which act as tumor suppressors. Smaller hyaluronan fragments can induce inflammatory response in macrophages and dendritic cells after tissue damage. They can also perform proangiogenic functions.

Proteoglycans

Glycosaminoglycans are commonly found attached to proteins and these are referred to as proteoglycans. Linkage between the protein and the glycosaminoglycan is made through a serine side-chain. Proteoglycans are made by glycosylation of target proteins in the Golgi apparatus.

This page titled 2.7: Structure and Function- Carbohydrates is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.8: Structure and Function - Lipids and Membranes

Source: BiochemFFA_2_7.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy



Lipids are a diverse group of molecules that all share the characteristic that at least a portion of them is hydrophobic. Lipids play many roles in cells, including serving as energy storage (fats/oils), constituents of membranes (glycerophospholipids, sphingolipids, cholesterol), hormones (steroids), vitamins (fat soluble), oxygen/ electron carriers (heme), among others. For lipids that are very hydrophobic, such as fats/ oils, movement and storage in the aqueous environment of the body requires special structures. Other, amphipathic lipids, such as glycerophospholipids and sphingolipids spontaneously organize themselves into lipid bilayers when placed in water. Interestingly, major parts of many lipids can be derived from acetyl-CoA.

Fatty acids

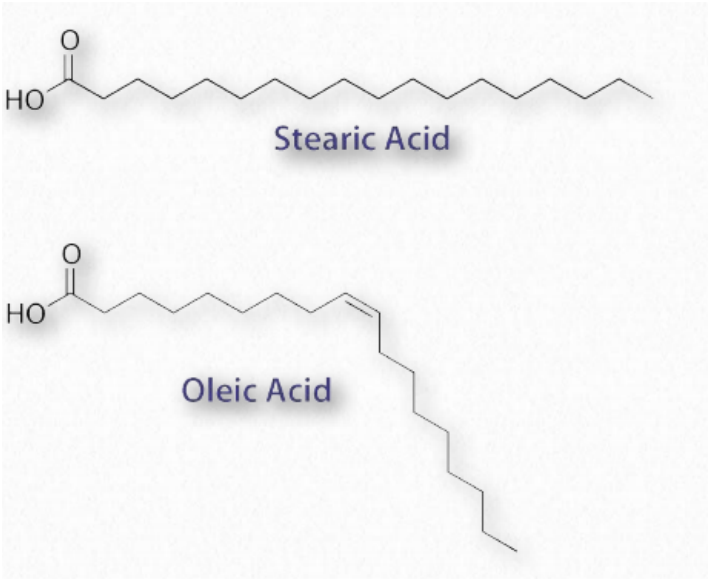


Figure 2.190 - Saturated fatty acid (stearic acid) and unsaturated fatty acid (oleic acid)

The most ubiquitous lipids in cells are the fatty acids. Found in fats, glycerophospholipids, sphingolipids and serving as as

membrane anchors for proteins and other biomolecules, fatty acids are important for energy storage, membrane structure, and as precursors of most classes of lipids. Fatty acids, as can be seen from Figure 2.190 are characterized by a polar head group and a long hydrocarbon tail. Fatty acids with hydrocarbon tails that lack any double bonds are described as saturated, while those with one or more double bonds in their tails are known as unsaturated fatty acids. The effect of double bonds on the fatty acid tail is to introduce a kink, or bend, in the tail, as shown for oleic acid.

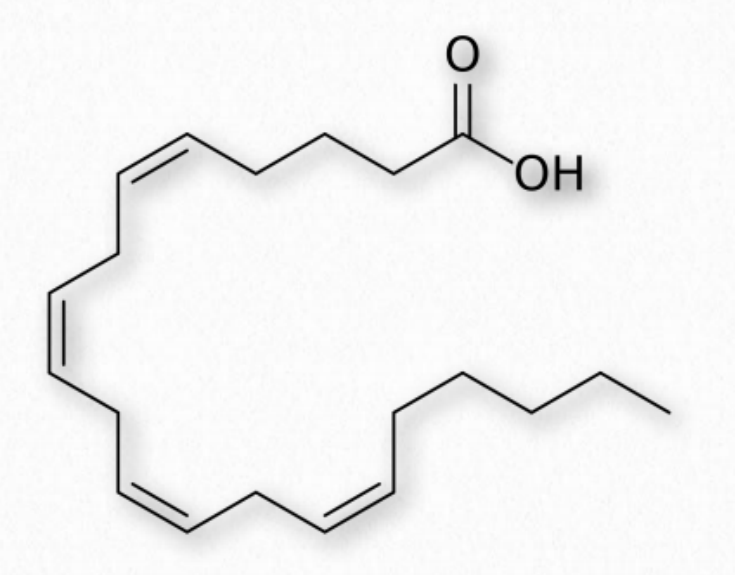


Figure 2.191 - Arachidonic acid - A polyunsaturated fatty acid Wikipedia

Stearic acid, a saturated fatty acid, by contrast has a straight hydrocarbon tail. Figures 2.190-2.194 show the most common saturated and unsaturated fatty acids. Fatty acids with unsaturated tails have a lower melting temperature than those with saturated tails of the same length. Shorter tails also decrease melting temperature. These properties carry over to the fats/oils containing them.

Common name	Chemical structure	
Caprylic acid	$CH_3(CH_2)_6COOH$	8
Capric acid	$CH_3(CH_2)_8COOH$	10
Lauric acid	$CH_3(CH_2)_{10}COOH$	12
Myristic acid	$CH_3(CH_2)_{12}COOH$	14
Palmitic acid	$CH_3(CH_2)_{14}COOH$	16
Stearic acid	$CH_3(CH_2)_{16}COOH$	18
Arachidic acid	$CH_3(CH_2)_{18}COOH$	20
Behenic acid	$CH_3(CH_2)_{20}COOH$	22
Lignoceric acid	$CH_3(CH_2)_{22}COOH$	24
Cerotic acid	$CH_3(CH_2)_{24}COOH$	26

Figure 2.192 - Saturated fatty acids. Number of carbons in right column Wikipedia

Fatty acids with more than one double bond are called polyunsaturated. Plants are excellent sources of unsaturated and polyunsaturated fatty acids. The position of the double bond(s) in fatty acids has important considerations both for their synthesis and for their actions in the body. Biochemically, the double bonds found in fatty acids are predominantly in the cis configuration. So-called trans fats arise as a chemical by-product of partial hydrogenation of vegetable oil.

Name	Double Bond Info	
Myristoleic acid	*cis*-Δ^9	14:1
Palmitoleic acid	*cis*-Δ^9	16:1
Sapienic acid	*cis*-Δ^6	16:1
Oleic acid	*cis*-Δ^9	18:1
Elaidic acid	*trans*-Δ^9	18:1
Vaccenic acid	*trans*-Δ^{11}	18:1
Linoleic acid	*cis,cis*-Δ^9,Δ^{12}	18:2
Linoelaidic acid	*trans,trans*-Δ^9,Δ^{12}	18:2
α-Linolenic acid	*cis,cis,cis*-$\Delta^9,\Delta^{12},\Delta^{15}$	18:3
Arachidonic acid	*cis,cis,cis,cis*-$\Delta^5\Delta^8,\Delta^{11},\Delta^{14}$	20:4
Eicosapentaenoic acid	*cis,cis,cis,cis,cis*-$\Delta^5,\Delta^8,\Delta^{11},\Delta^{14},\Delta^{17}$	20:5
Erucic acid	*cis*-Δ^{13}	22:1
Docosahexaenoic acid	*cis,cis,cis,cis,cis,cis*-$\Delta^4,\Delta^7,\Delta^{10},\Delta^{13},\Delta^{16},\Delta^{19}$	22:6

Figure 2.193 - Unsaturated fatty acids. Right column Indicates number of carbons and double bonds Wikipedia

In humans, consumption of trans fat raises low density lipoprotein (LDL) levels and lowers high density lipoprotein (HDL) levels. Each is thought to contribute to the risk of developing coronary artery disease. The most Figure 2.194 - Fatty acid models. Carboxyl end labeled in red Wikipedia common fatty acids in our body include palmitate, stearate, oleate, linolenate, linoleate, and arachidonate. Two notable shorter fatty acids are nonanoic (9 carbons) and decanoic acid (10 carbons), both of which appear to have anti-seizure effects. Decanoic acid directly inhibits excitatory neurotransmission in the brain and may contribute to the anticonvulsant effect of the ketogenic diet.

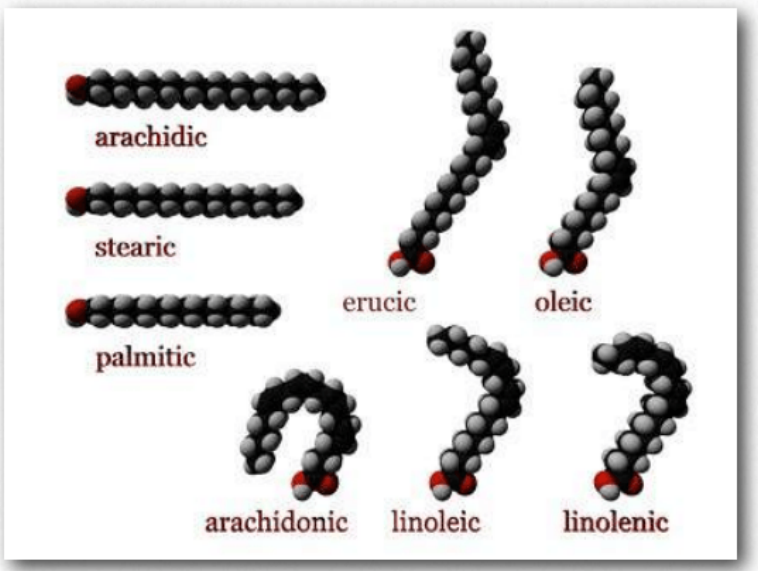


Figure 2.194 - Fatty acid models. Carboxyl end labeled in red Wikipedia

Numbering

Figure 2.195 shows two different systems for locating double bonds in a fatty acid. The ω system counts carbons starting with the methyl end (shown in red) while the Δ system counts from the carboxyl end (shown in blue). For example, an ω-3 (omega 3) fatty acid would have a double bond at the third carbon from the methyl end. In the Δ system, a fatty acid that has a cis double bond at carbon 6, counting from the carboxyl end, would be written as cis-Δ6.

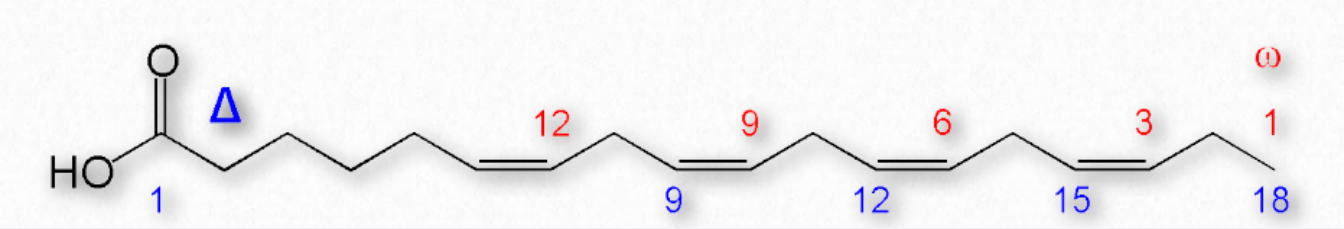


Figure 2.195 - Δ and ω numbering systems for fatty acids Image by Pehr Jacobson

Fatty acids are described as essential fatty acids if they must be in the diet (can't be synthesized by the organism) and nonessential fatty acids if the organism can synthesize them. Humans and other animals lack the desaturase enzymes necessary to make double bonds at positions greater than Δ-9, so fatty acids with double bonds beyond this position must be obtained in the diet. Linoleic acid and linolenic acid, both fall in this category. Related unsaturated fatty acids can be made from these fatty acids, so the presence of linoleic and linolenic acids in the diet eliminates the need to have all unsaturated fatty acids in the diet. Both linoleic and linolenic acid contain 18 carbons, but linoleic acid is an ω-6 fatty acid, whereas linolenic acid is an ω-3 fatty acid. Notably, ω-6 fatty acids tend to be proinflammatory, whereas ω-3 fatty acids are lesser so.

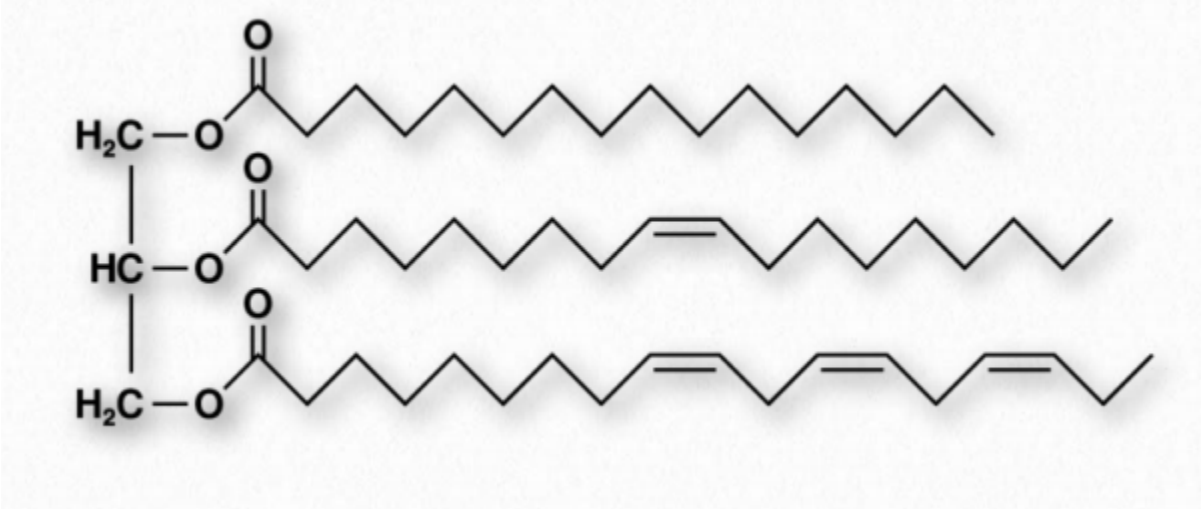


Figure 2.196 - Structure of a fat/oil

Fats/oils

Fats and oils are the primary energy storage forms of animals and are also known as triacylglycerols and triglycerides, since they consist of a glycerol molecule linked via ester bonds to three fatty acids (Figure 2.196). Fats and oils have the same basic structure. We give the name fat to those compounds that are solid at room temperature and the name oil to those that are liquid at room temperature. Note that biological oils are not the same as petroleum oils.

Increasing the number of unsaturated fatty acids (and the amount of unsaturation in a given fatty acid) in a fat decreases the melting temperature of it. Organisms like fish, which live in cool environments, have fats with more unsaturation and this is why fish oil contains polyunsaturated fatty acids.

Adipocytes

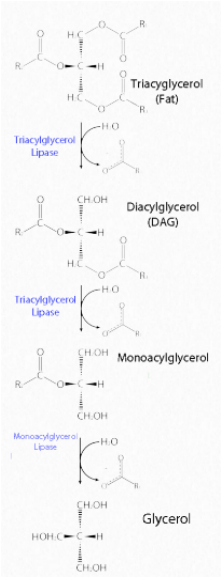


Figure 2.197 - Lipase action on a fat Image by Aleia Kim

Fats are stored in the body in specialized cells known as adipocytes. Enzymes known as lipases release fatty acids from fats by hydrolysis reactions (Figure 2.197). Triacylglycercol lipase (pancreatic - Figure 2.198) is able to cleave the first two fatty acids from the fat. A second enzyme, monoacylglycerol lipase, cleaves the last fatty acid. Fats can be synthesized by replacing the phosphate on phosphatidic acid with a fatty acid.



Figure 2.198 - Pancreatic lipase

Glycerophospholipids

Glycerophospholipids (phosphoglycerides) are important components of the lipid bilayer of cellular membranes. Phosphoglycerides are structurally related to fats, as both are derived from phosphatidic acid (Figure 2.199). Phosphatidic acid is a simple glycerophospholipid that is usually converted into phosphatidyl compounds. These are made by esterifying various groups, such as ethanolamine, serine, choline, inositol, and others (Figure 2.200) to the phosphate of phosphatidic acid. All of these compounds form lipid bilayers in aqueous solution , due to their amphiphilic nature.

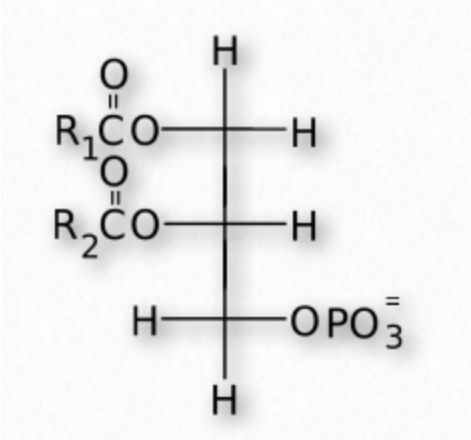


Figure 2.199 - Structure of phosphatidic acid. R1 and R2 are alkyl groups of fatty acids.

Phosphatidylethanolamines

Since all glycerolipids can have a variety of fatty acids at positions 1 and 2 on the glycerol, they all are families of compounds. The phosphatidylethanolamines are found in all living cells and are one of the most common phosphatides, making up about 25% of them. They are common constituents of brain tissue and in the spinal cord, making up as much as 45% of the total phospholipids. Phosphatidylethanolamines are asymmetrically distributed across membranes, being preferentially located on the inner leaflet (closest to the cytoplasm) of the plasma membrane. Metabolically, phosphatidylethanloamines are precursors of phosphatidylcholines. Phosphatidylserines Phosphatidylserines are another group of phosphatidyl compounds that are preferentially distributed across the lipid bilayer of the plasma membrane. Like the phosphatidylethanolamines, phosphatidylserines are preferentially located on the inner leaflet of the plasma membrane. When apoptosis (cell suicide) occurs, the preferential distribution is lost and the phosphatidylserines appear on the outer leaflet where they serve as a signal to macrophages to bind and destroy the cell.

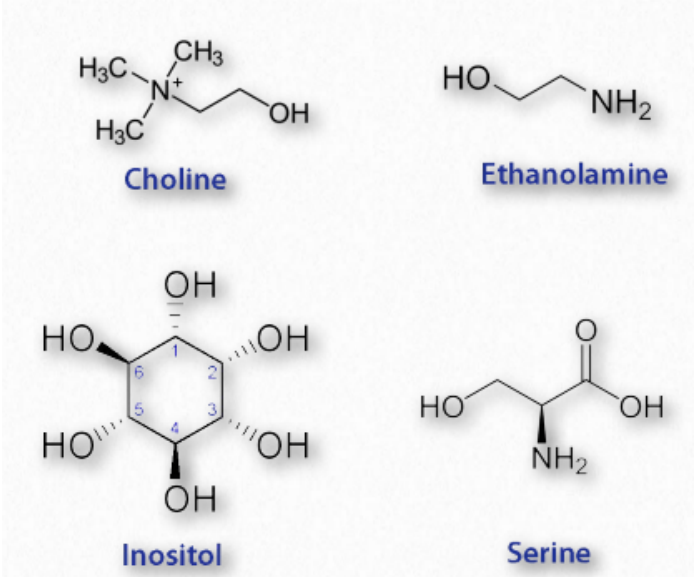


Figure 2.200 - Four common components of phosphatides Wikipedia

Phosphatidylcholines

Phosphatidylcholines (Figure 2.201) are another group of important membrane components. They tend to be found more commonly on the outer leaflet of the plasma membrane. Nutritionally, the compounds are readily obtained from eggs and soybeans. Phosphatidylcholines are moved across membranes by Phosphatidylcholine transfer protein (PCTP). This protein, which is sensitive to the levels of phosphatidylcholines, acts to stimulate the activity of a thioesterase (breaks thioester bonds, such as acyl-CoAs) and activates PAX3 transcription factors.

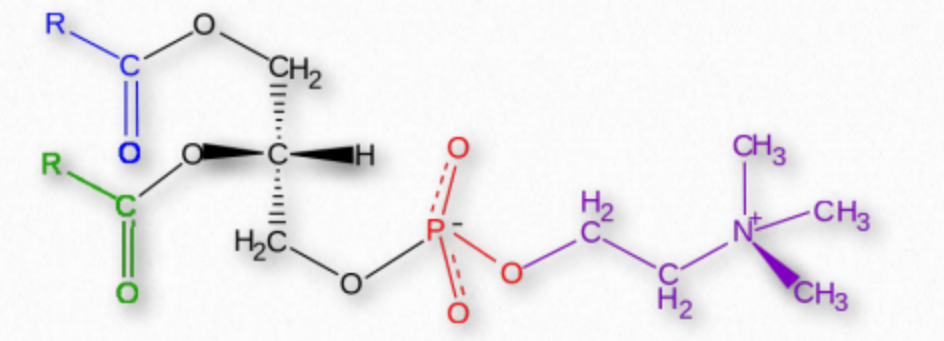


Figure 2.201 - Phosphatidylcholine

Cardiolipins

Cardiolipins are an unusual set of glycerophospholipids in containing two diacylglycerol backbones joined in the middle by a diphosphoglycerol (Figure 2.202). It is an important membrane lipid, constituting about 20% of the inner mitochondrial membrane and is found in organisms from bacteria to humans. In both plants and animals, it is found almost totally in the inner mitochondrial membrane.

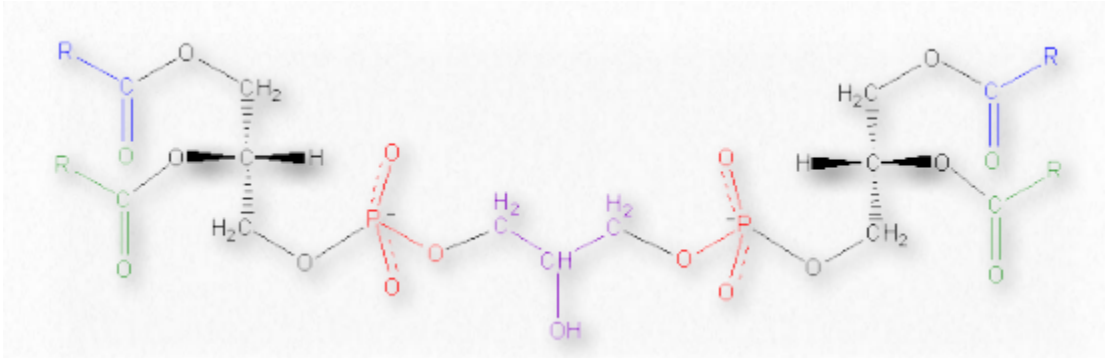


Figure 2.202 - Cardiolipin

The molecules appear to be required for both Complex IV and Complex III of the electron transport chain to maintain its structure. The ATP synthase enzyme (Complex V) of the oxidative phosphorylation system also binds four molecules of cardiolipin. It has been proposed that cardiolipin functions as a proton trap in the process of proton pumping by Complex IV.

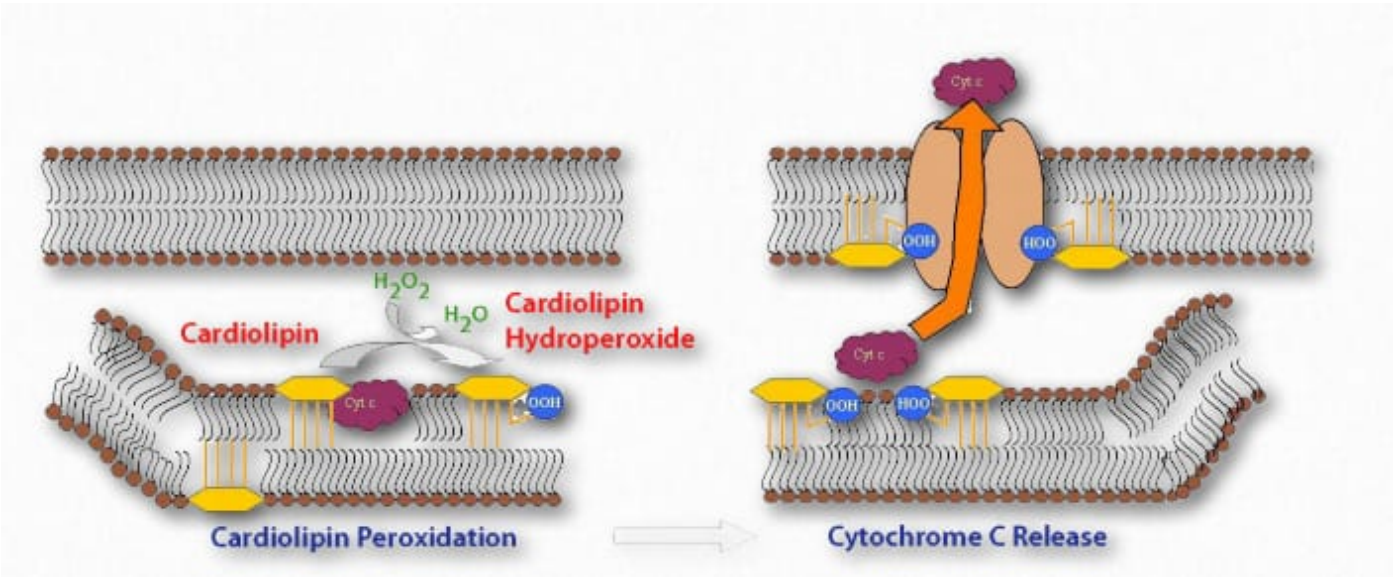


Figure 2.203 - Cardiolipin oxidation and the release of cytochrome C in apoptosis

Cardiolipin also plays a role in apoptosis. As shown in Figure 2.203, oxidation of cardiolipin by a cardiolipin-specific oxygenase causes cardiolipin to move from the inner mitochondrial membrane to the outer one, helping to form a permeable pore and facilitating the transport of cytochrome c out of the intermembrane space and into the cytoplasm - a step in the process of apoptosis.

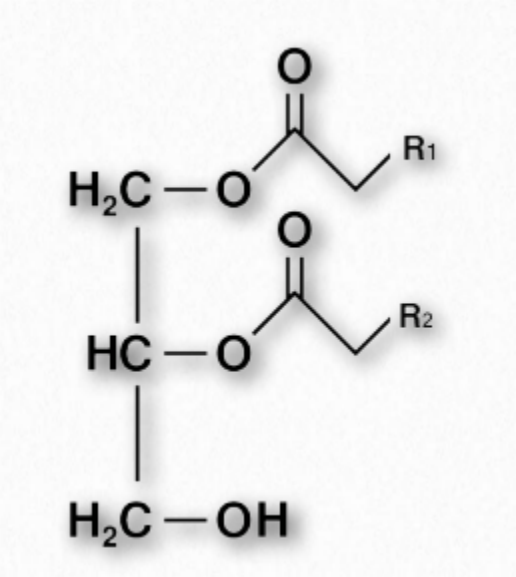


Figure 2.204 - Diacylglycerol

Diacylglycerol

Diacylglycerol (also called diglyceride and DAG - Figure 2.204) is an important intermediate in metabolic pathways. It is produced, for example, in the first step of the hydrolysis of fat and is also produced when membrane lipids, such as PIP2 (phosphatidylinositol-4,5-bisphosphate) are hydrolyzed by phospholipase C in a signaling cascade.

DAG is itself a signaling compound, binding to protein kinase C to activate it to phosphorylate substrates. Synthesis of DAG begins with glycerol-3-phosphate, which gains two fatty acids from two acyl-CoAs to form phosphatidic acid. Dephosphorylation of phosphatidic acid produces DAG. DAG can also be rephosphorylated by DAG kinase to re-make phosphatidic acid or another fatty acid can be added to make fat.

Inositol

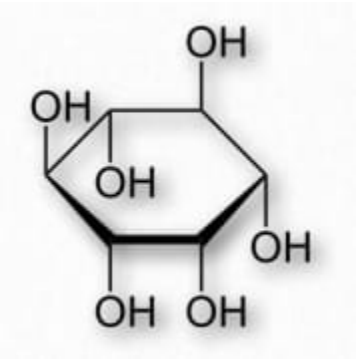


Figure 2.205 - Inositol

Though technically not a lipid itself, inositol is found in many lipids. Inositol is a derivative of cyclohexane containing six hydroxyl groups - one on each carbon (Figure 2.205. It has nine different stereoisomers of which one, cis-1,2,3,5-trans-4,6-cyclohexanehexol (called myo-inositol) is the most common. It has a sweet taste (half that of sucrose).

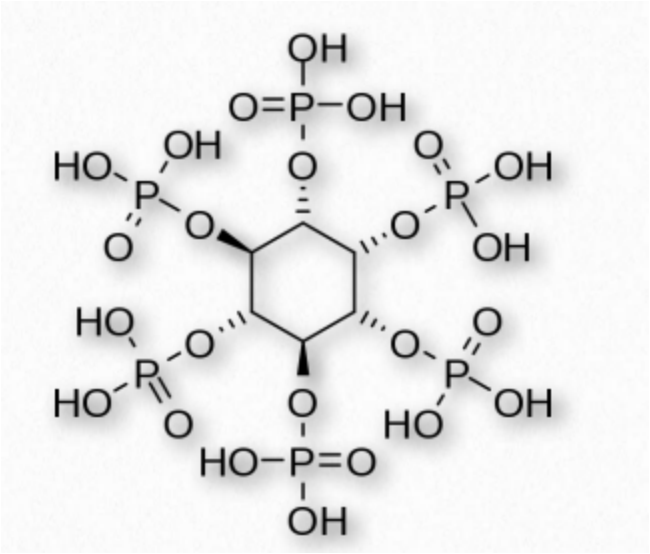


Figure 2.206 - Phytic acid

Numerous phosphorylated forms of the compound exist, from a single phosphate to six (one on each carbon). Phytic acid, for example, in plants, has six phosphates (Figure 2.206) that it uses to store phosphate. Inositol is produced from glucose and was once considered vitamin B8, but is made by the body in adequate amounts, so it is not now considered a vitamin. Phosphorylated forms of inositol are found in phosphoinositides, such as PIP2 and PIP3, both of which are important in signaling processes. Some of these include insulin signaling, fat catabolism, calcium regulation, and assembly of the cytoskeleton.

Phosphoinositides

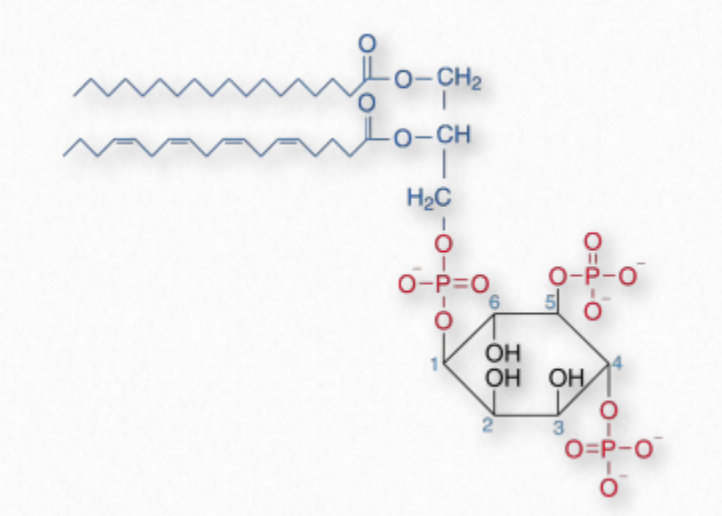


Figure 2.207 - Structure of PIP_2

Compounds based on phosphatidylinositol (PI) are often called phosphoinositides. These compounds have important roles in signaling and membrane trafficking. Hydroxyls on carbons 3,4, and 5 of the inositol ring are targets for phosphorylation by a variety of kinases. Seven different combinations are used. Steric hindrance inhibits phosphorylation of carbons 2 or 6. Naming of these phosphorylated compounds follows generally as PI(#P)P, PI(#P, #P)P, or PI(#P, #P, #P)P where #P refers to the number of the carbon where a phosphate is located. For example, PI(3)P refers to a phosphatidyl compound with a phosphate added to carbons 3 of the inositol ring, whereas PI(3,4,5)P is a phosphatidyl compound with a phosphate added to carbons 3,4,and 5.

Phosphatidylinositol-4,5- bisphosphate

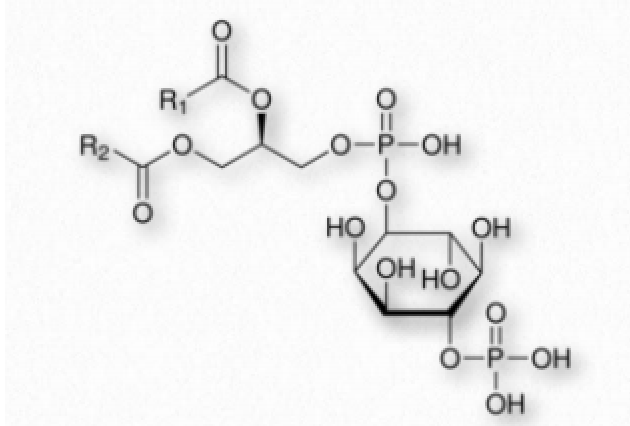


Figure 2.208 - Phosphatidylinositol-4- phosphate

Phosphatidylinositol-4,5-bisphosphate (PIP2 - Figure 2.207) is a phospholipid of plasma membranes that functions in the phospholipase C signaling cascade. In this signaling pathway, hydrolysis catalyzed by phospholipase C releases inositol-1,4,5-trisphosphate (IP3) and diacylglycerol. Synthesis of PIP2 begins with phosphatidylinositol, which is phosphorylated at position 4 followed by phosphorylation at position 5 by specific kinases.

PIP2 can be phosphorylated to form the signaling molecule known as phosphatidylinositol (3,4,5)-trisphosphate (PIP3). Along with PIP3, PIP2 serves as a docking phospholipid for the recruitment of proteins that play roles in signaling cascades. Binding of PIP2 is also required by inwardly directed potassium channels.

Phosphatidylinositol (3,4,5)- trisphosphate

Phosphatidylinositol (3,4,5)-trisphosphate (PIP3) is an important molecule for the activation of signaling proteins, such as AKT, which activates anabolic signaling pathways related to growth and survival. PIP3 can be dephosphorylated by phosphatase PTEN to yield PIP2 and can be synthesized from PIP2 by kinase action of Class I PI 3- kinases. Kinase activity to synthesize PIP3 results in movement of PIP3-binding proteins to the plasma membrane. They include Akt/ PKB, PDK1, Btk1, and ARNO and each is activated by binding to PIP3.

Plasmalogens

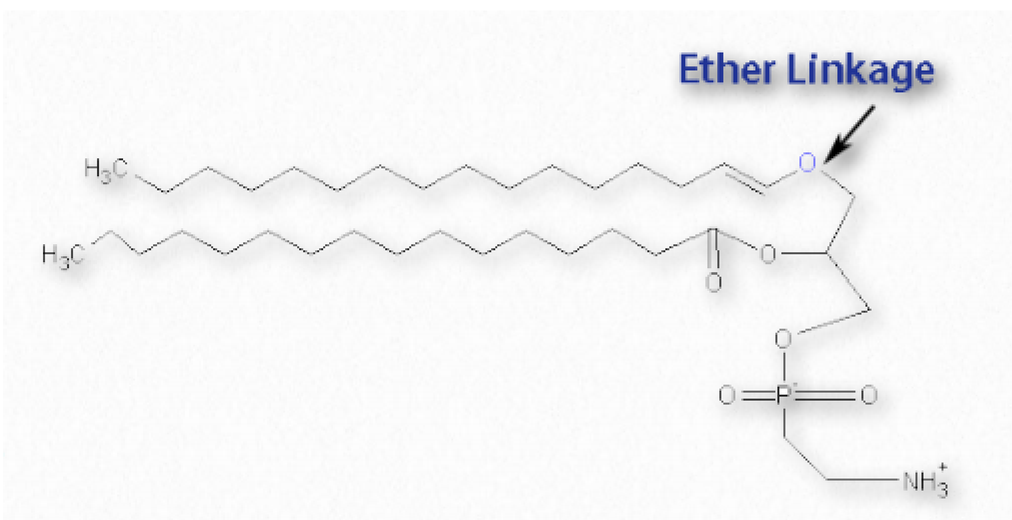


Figure 2.209 - Plasmalogen - A vinyl ether lipid Wikipedia

A special class of the glycerophospholipids are the plasmalogens (Figure 2.209). They differ in containing a vinyl ether linkage at position 1 of glycerol, in contrast to other glycerophopsholipids, which have an ester linkage at this position. Position 2 of each is an ester. The precursor for the ether linkage is typically a 16 or 18 carbon saturated alcohol or an 18 carbon unsaturated alcohol.

At the phosphate tail, the most commonly attached groups are ethanolamine or choline. Plasmalogens are found abundantly in

humans in heart (30-40% of choline phospholipids). 30% of the glycerophospholipids in brain are plasmalogens and 70% of the ethanolamine lipids of the myelin sheath of nerve cells are plasmalogens.

Though their function is not understood, it is believed that plasmalogens may provide some protection against reactive oxygen species and have roles in signaling.

Lecithin

Lecithin is a generic term for a combination of lipid substances that includes phosphoric acid, glycerol, glycolipids, triglycerides, and phospholipids. Lecithin is a wetting agent helpful with emulsification and encapsulation and is even used as an anti-sludge additive in motor lubricants. Lecithin is used in candy bars to keep cocoa and cocoa butter from separating. Though considered safe as a food ingredient, lecithin can be converted by gut bacteria to trimethylamine-N-oxide which may contribute to cholesterol deposition and atherosclerosis.

Sphingolipids

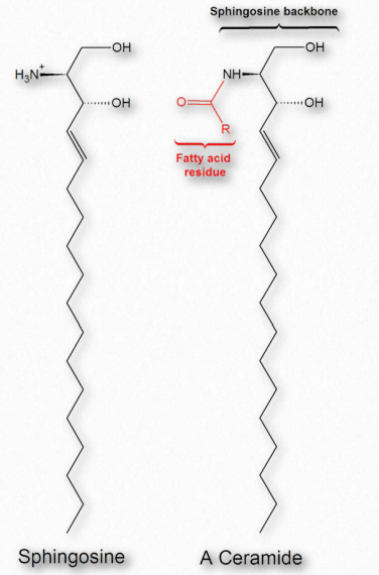


Figure 2.210 - Sphingosine and a ceramide made from it Wikipedia

Fatty acids are also components of a broad class of molecules called sphingolipids. Sphingolipids are structurally similar to glycerophospholipids, though they are synthesized completely independently of them starting with palmitic acid and the amino acid serine. Sphingolipids are named for the amino alcohol known as sphingosine (Figure 2.210), though they are not directly synthesized from it. Figure 2.211 shows the generalized structure of sphingolipids.

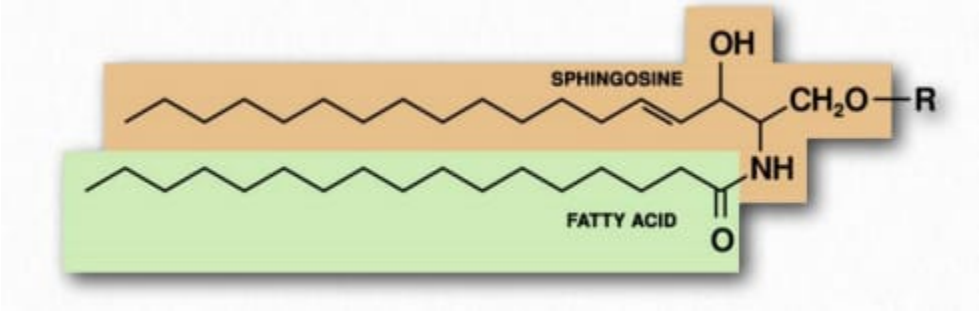


Figure 2.211 Schematic structure of a sphingolipid

If the R-group is a hydrogen, the molecule is called a ceramide. When the R-group is phosphoethanolamine the resulting molecule is sphingomyelin, an important component of the myelin sheath and lipid membranes. If a single, simple sugar is instead added, a cerebroside is created (Figure 2.212). Addition of a complex oligosaccharide creates a ganglioside.

Complex sphingolipids may play roles in cellular recognition and signaling. Sphingolipids are found most abundantly in plasma

membrane and are almost completely absent from mitochondrial and endoplasmic reticulum membranes. In animals, dietary sphingolipids have been linked to reduced colon cancer, reductions in LDLs, and increases in HDLs. Like the glycerophospholipids, sphingolipids are amphiphilic. Most sphingolipids except sphingomyelin do not contain phosphate.

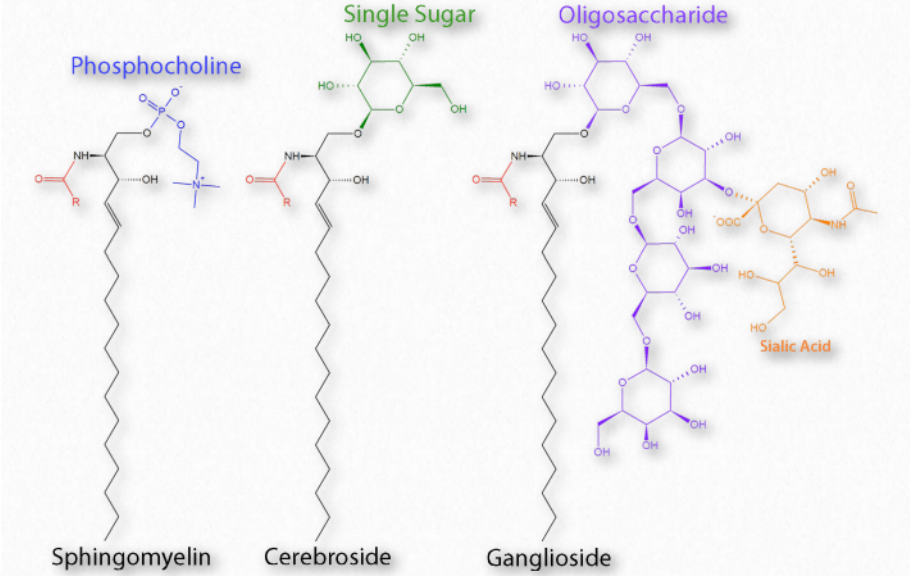


Figure 2.212 - Categories of Sphingolipid Wikipedia

Eicosanoids

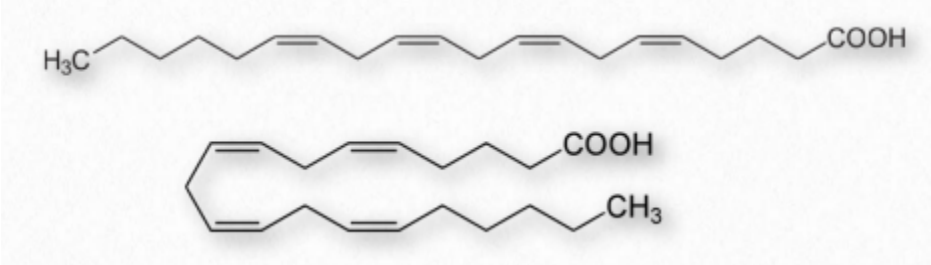


Figure 2.213 - Arachidonic acid drawn as straight (top) and bent (bottom)

Fatty acids made from omega-6 and omega-3 fatty acids include three important fatty acids containing 20 carbons. They include arachidonic acid (an ω-6 fatty acid with four double bonds (Δ-5,8,11,14) - Figure 2.213), eicosapentaenoic acid (an ω-3 fatty acid with five double bonds, and dihomo-γ-linolenic acid (an ω-6 fatty acid with three double bonds). The class of compounds known as eicosanoids is made by oxidation of these compounds. Subclasses include include prostaglandins, prostacyclins, thromboxanes, lipoxins, leukotrienes, and endocannabinoids (Figures 2.214-2.219). Eicosanoids play important roles affecting inflammation, immunity, mood, and behavior.

Prostaglandins

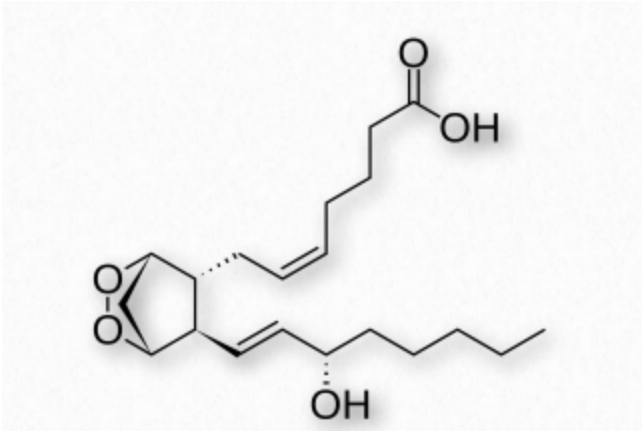


Figure 2.214 - Prostaglandin PGH_2

A collection of molecules acting like hormones, prostaglandins are derived from arachidonic acid and have many differing (even conflicting) physiological effects. These include constriction or dilation of vascular smooth muscle cells, induction of labor, regulation of inflammation, and action on the thermoregulatory center of the hypothalamus to induce fever, among others.

Prostaglandins are grouped with the thromboxanes (below) and prostacyclins (below), as prostanoids. The prostanoids, which all contain 20 carbons are a subclass of the eicosanoids. Prostaglandins are found in most tissues of higher organisms. They are autocrine or paracrine compounds produced from essential fatty acids. The primary precursor of the prostaglandins is the fatty acid known as arachidonic acid and the prostaglandin made from it is known as PGH2 (Figure 2.214), which, in turn is a precursor of other prostaglandins, as well as the prostacyclins and thromboxanes.

Interesting prostaglandins

PGD_2 - inhibits hair follicle growth, vasodilator, causes bronchial constriction, higher in lungs of asthmatics than others.

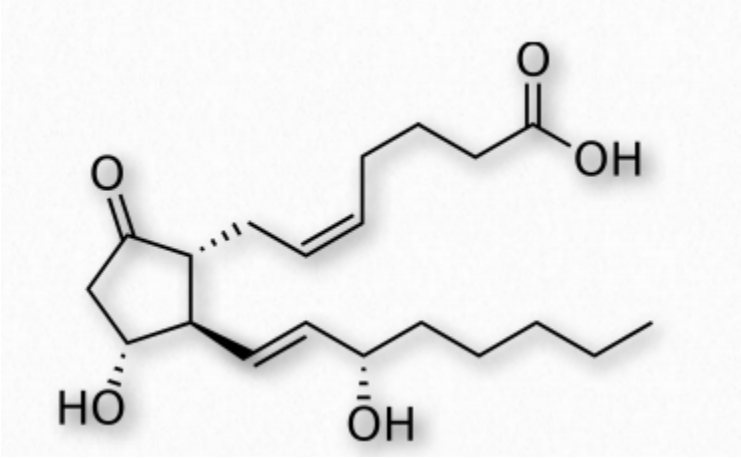


Figure 2.215 Prostaglandin E

PGE_2 (Figure 2.215) - exerts effects in labor (soften cervix, uterine contraction), stimulates bone resorption by osteoclasts, induces fever, suppresses T-cell receptor signaling, vasodilator, inhibits release of noradrenalin from sympathetic nerve terminals. It is a potent activator of the Wnt signaling pathway.

A prostaglandin can have opposite effects, depending on which receptor it binds to. Binding of PGE2 to the EP1 receptor causes bronchoconstriction and smooth muscle contraction, whereas binding of the same molecule to the EP2 receptor causes bronchodilation and smooth muscle relaxation.

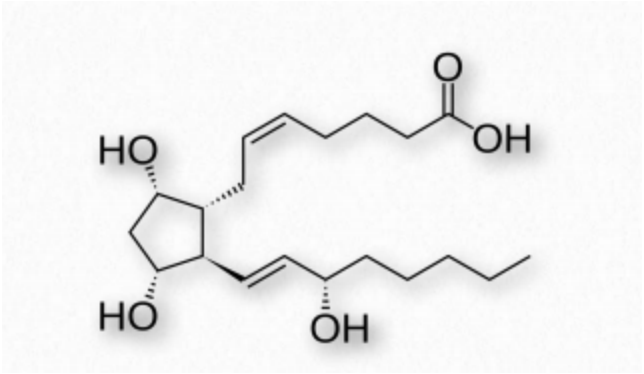


Figure 2.216 - Prostaglandin $F_{2\alpha}$

$PGF_{2\alpha}$ (Figure 2.216)- uterine contractions, induces labor, bronchoconstriction.

PGI_2 - vasodilation, bronchodilation, inhibition of platelet aggregation.

Thromboxanes

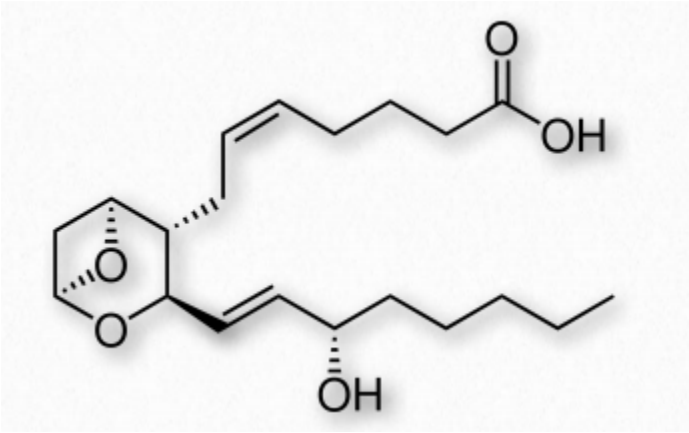


Figure 2.217 Thromboxane $A2_2$

Thromboxanes play roles in clot formation and named for their role in thrombosis. They are potent vasoconstrictors and facilitate platelet aggregation. They are synthesized in platelets, as well. The anti-clotting effects of aspirin have their roots in the inhibition of synthesis of PGH2, which is the precursor of the thromboxanes. The most common thromboxanes are A2 (Figure 2.217) and B2.

Prostacyclin

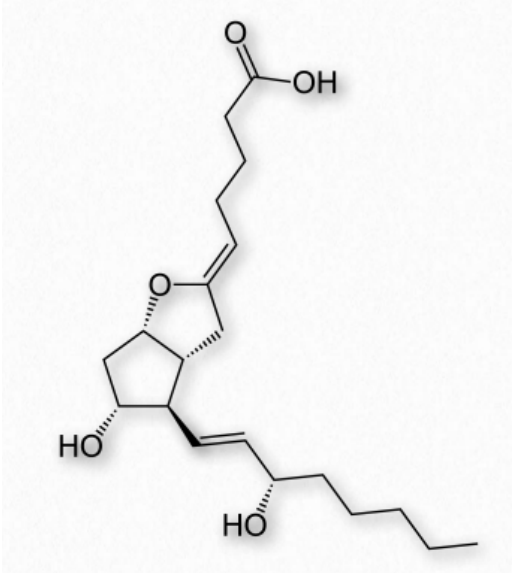


Figure 2.218 - Prostacyclin

Prostacyclin (also known as prostaglandin I2 or PGI2 - Figure 2.218) counters the effects of thromboxanes, inhibiting platelet activation and acting as vasodilators. It is produced from PGH2 by action of the enzyme prostacyclin synthase.

Leukotrienes

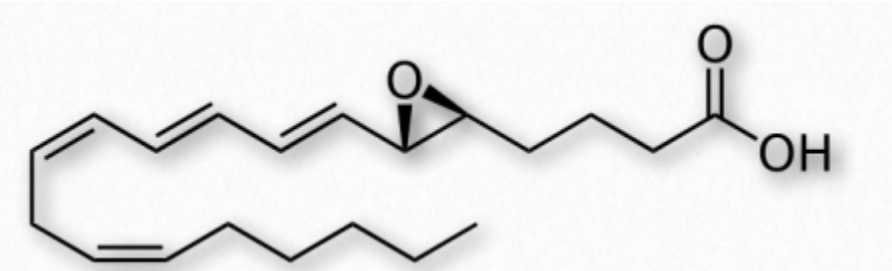


Figure 2.219 - Leukotriene A_4 (LTA_4)

Another group of eicosanoid compounds are the leukotrienes (Figure 2.219). Like prostaglandins, leukotrienes are made from arachidonic acid. The enzyme catalyzing their formation is a dioxygenase known as arachidonate 5-lipoxygenase. Leukotrienes are involved in regulating immune responses. They are found in leukocytes and other immunocompetent cells, such as neutrophils, monocytes, mast cells, eosinophils, and basophils. Leukotrienes are associated with production of histamines and prostaglandins, which act as mediators of inflammation. Leukotrienes also trigger contractions in the smooth muscles of the bronchioles. When overproduced, they may pay a role in asthma and allergic reactions. Some treatments for asthma aim at inhibiting production or action of leukotrienes.

Cholesterol

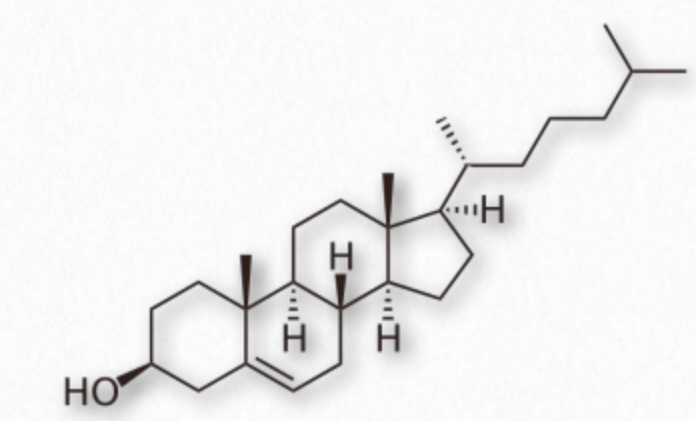


Figure 2.220 - Structure of cholesterol

Arguably, no single biomolecule has generated as much discussion and interest as has cholesterol (Figure 2.220). Certainly, from the perspective of the Nobel Prize committee, no small molecule even comes close, with 13 people having been awarded prizes for work on it. Evidence for cholesterol's importance comes from the study of brain tissue where it comprises 10-15% of the dry mass.

Membrane flexibility

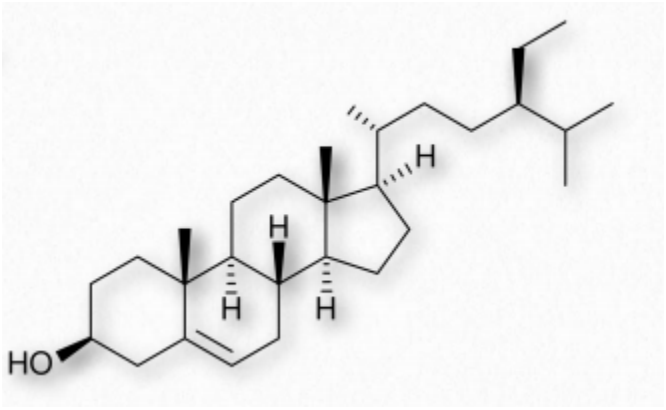


Figure 2.221 - Sitosterol - A phytosterol

In animal cells, cholesterol provides for membrane flexibility that allows for cellular movement that is in contrast to plant and bacterial cells with fixed structures. Cholesterol is made in many cells of the body, with the liver making the greatest amount. The anabolic pathway leading to synthesis of cholesterol is known as the isoprenoid pathway and branches of it lead to other molecules including other fat-soluble vitamins.



Figure 2.222 - Margarine - A common source of trans fat Wikipedia

Cholesterol is only rarely found in prokaryotes (Mycoplasma, which requires it for growth, is an exception) and is found in only trace amounts in plants. Instead, plants synthesize similar compounds called phytosterols (Figure 2.221). On average, the body of a 150 pound adult male makes about 1 gram of cholesterol per day, with a total content of about 35 grams.

Packaging

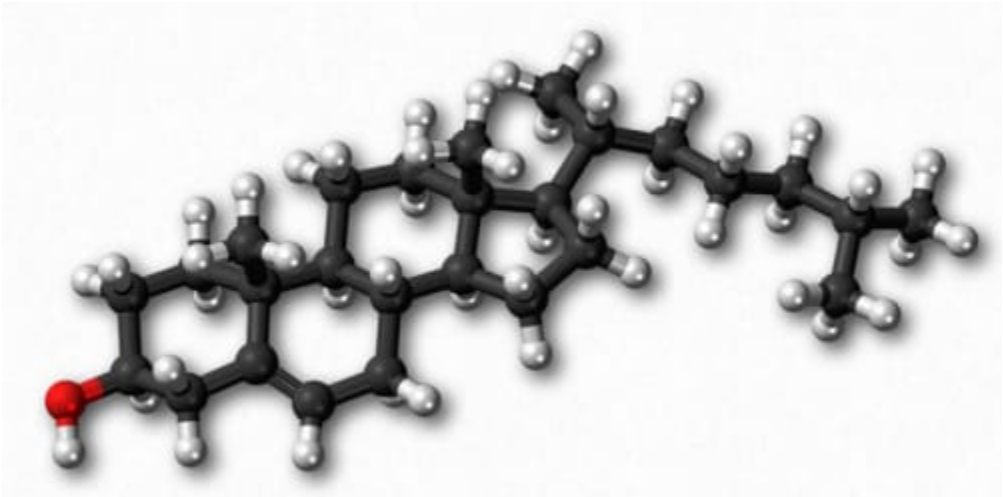


Figure 2.223 - Cholesterol - Ball and stick model

Cholesterol's (and other lipids') hydrophobicity requires special packaging into lipoprotein complexes (called chylomicrons, VLDLs, IDLs, LDLs, and HDLs) for movement in the lymph system and bloodstream. Though cholesterol can be made by cells, they also take it up from the blood supply by absorbing cholesterol-containing LDLs directly in a process called receptor-mediated endocytosis.

Oxidative damage to LDLs can lead to formation of atherosclerotic plaques and this is why cholesterol has gotten such a negative image in the public eye. The liver excretes cholesterol through the bile for elimination into the digestive system, but the compound is recycled there.

Reducing cholesterol levels

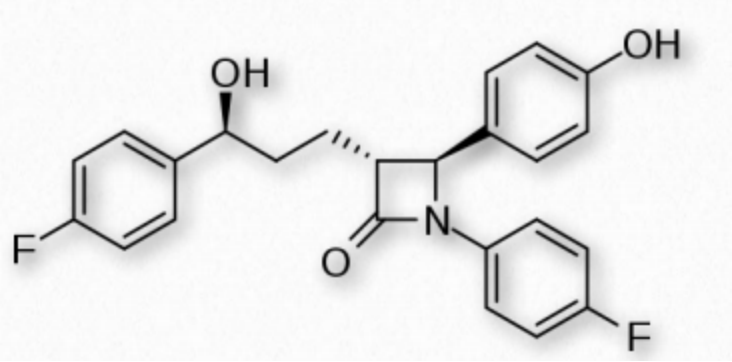


Figure 2.224 - Ezetimibe - An inhibitor of cholesterol absorption

Strategies for reducing cholesterol in the body focus primarily on three areas - reducing consumption, reducing endogenous synthesis, and reducing the recycling. Dietary considerations, such as saturated fat versus unsaturated fat consumption are currently debated. Dietary trans fats, though, correlate with incidence of coronary heart disease. Consumption of vegetables may provide some assistance with reducing levels of cholesterol recycled in the digestive system, because plant phytosterols compete with cholesterol for reabsorption and when this happens, a greater percentage of cholesterol exits the body in the feces. Drugs related to penicillin are also used to inhibit cholesterol recycling. One of these is ezetimibe, shown in Figure 2.224.

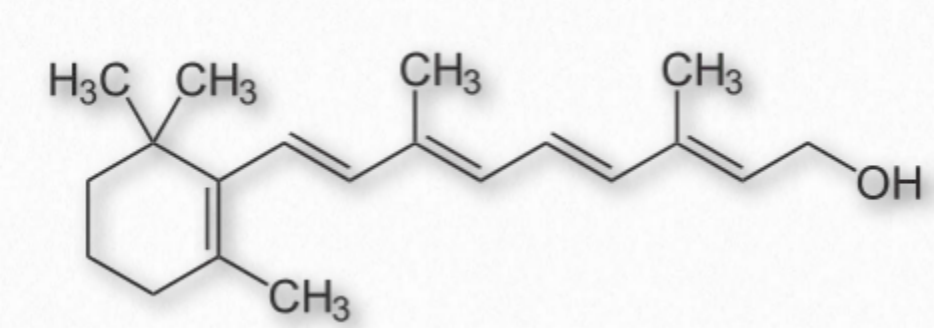


Figure 2.225 - All-trans retinol

Genetic defects in the cholesterol movement system are a cause of the rare disease known as familial hypercholesterolemia in which the blood of afflicted individuals contains dangerously high levels of LDLs. Left untreated, the disease is often fatal in the first 10-20 years of life. While LDLs have received (and deserve) a bad rap, another group of lipoprotein complexes known as the HDLs (high density lipoprotein complexes) are known as “good cholesterol” because their levels correlate with removal of debris (including cholesterol) from arteries and reduce inflammation.

Membrane function

In membranes, cholesterol is important as an insulator for the transmission of signals in nerve tissue and it helps to manage fluidity of membranes over a wide range of temperatures. Stacked in the lipid bilayer, cholesterol decreases a membrane’s fluidity and its permeability to neutral compounds, as well as protons and sodium ions. Cholesterol may play a role in signaling by helping with construction of lipid rafts within the cell membrane.

Vitamin A

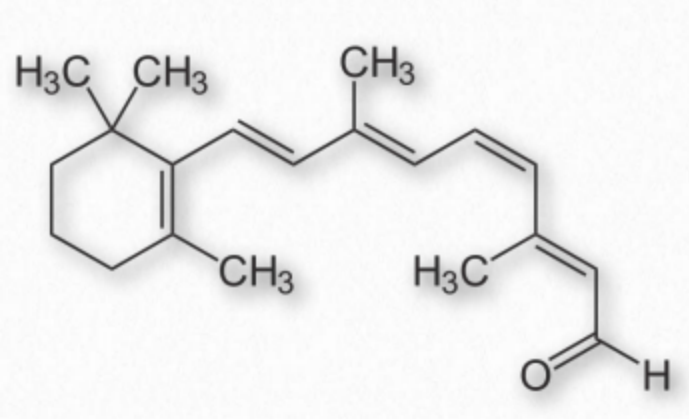


Figure 2.226 - 11-cis retinal

Vitamin A comes in three primary chemical forms, retinol (storage in liver - Figure 2.225), retinal (role in vision - Figure 2.226), and retinoic acid (roles in growth and development). All vitamin A forms are diterpenoids and differ only in the chemical form of the terminal group. Retinol is mostly used as the storage form of the vitamin.

Retinol is commonly esterified to a fatty acid and kept in the liver. In high levels, the compound is toxic. Retinol is obtained in the body by hydrolysis of the ester or by reduction of retinal. Importantly, neither retinal nor retinol can be made from retinoic acid. Retinoic acid is important for healthy skin and teeth, as well as bone growth. It acts in differentiation of stem cells through a specific cellular retinoic acid receptor.

Sources

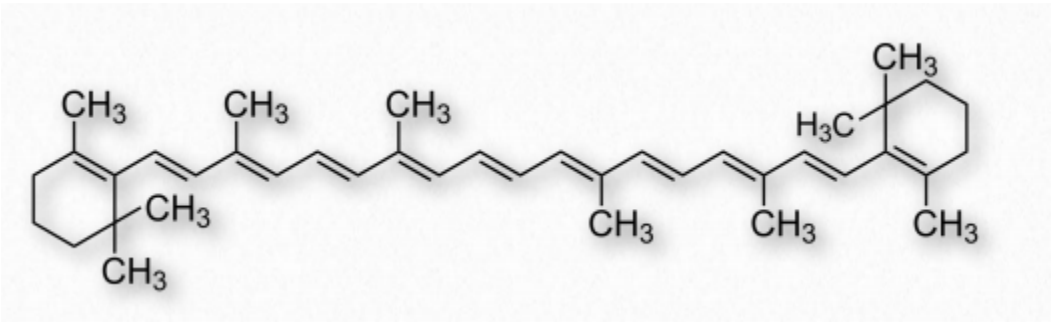


Figure 2.227 - β-Carotene

Good sources of vitamin A are liver and eggs, as well as many plants, including carrots, which have a precursor, β-carotene (Figure 2.227) from which retinol may be made by action of a dioxygenase.

Light sensitivity The conjugated double bond system in the side chain of vitamin A is sensitive to light and can flip between cis and trans forms on exposure to it. It is this response to light that makes it possible for retinal to have a role in vision in the rods and cones of the eyes. Here, the aldehyde form (retinal) is bound to the protein rhodopsin in the membranes of rod and cone cells.

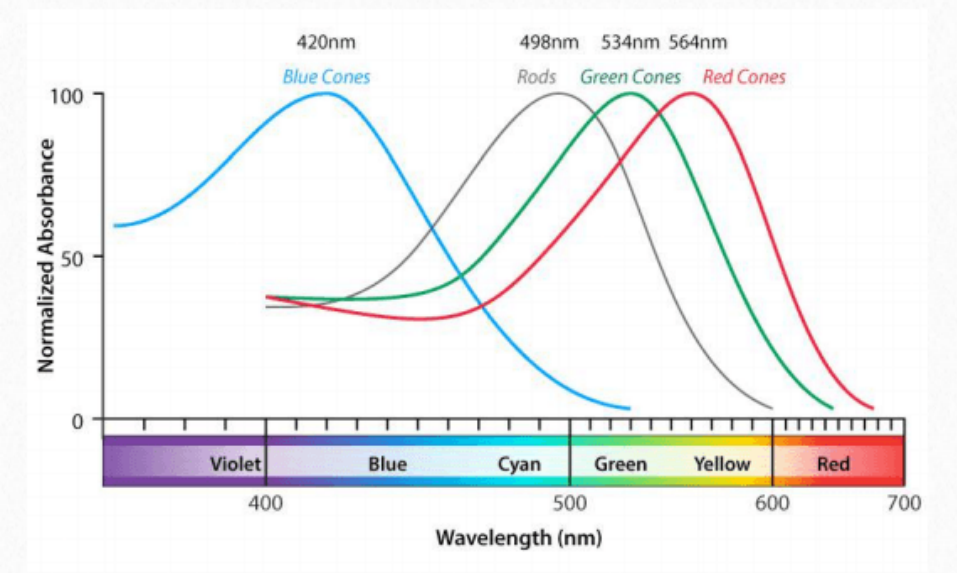


Figure 2.228 - Color sensitivity for cones and rods Image by Aleia Kim

When exposed to light of a particular wavelength, the “tail” of the retinal molecule will flip back and forth from cis to trans at the double bond at position 11 of the molecule. When this happens, a nerve signal is generated that signals the brain of exposure to light. Slightly different forms of rhodopsin have different maximum absorption maxima allowing the brain to perceive red, green and blue specifically and to assemble those into the images we see (Figure 2.228). Cones are the cells responsible for color vision, whereas rods are mostly involved in light detection in low light circumstances.

Deficiency and surplus

Deficiency of vitamin A is common in developing countries and was inspiration for the design and synthesis of the geneticallymodified golden rice, which is used as a source of vitamin A to help prevent blindness in children. Overdose of vitamin A, called hypervitaminosis A is dangerous and can be fatal. Excess vitamin A is also suspected to be linked to osteoporosis. In smokers, excess vitamin A is linked to an increased rate of lung cancer, but non-smokers have a reduced rate.

Vitamin D

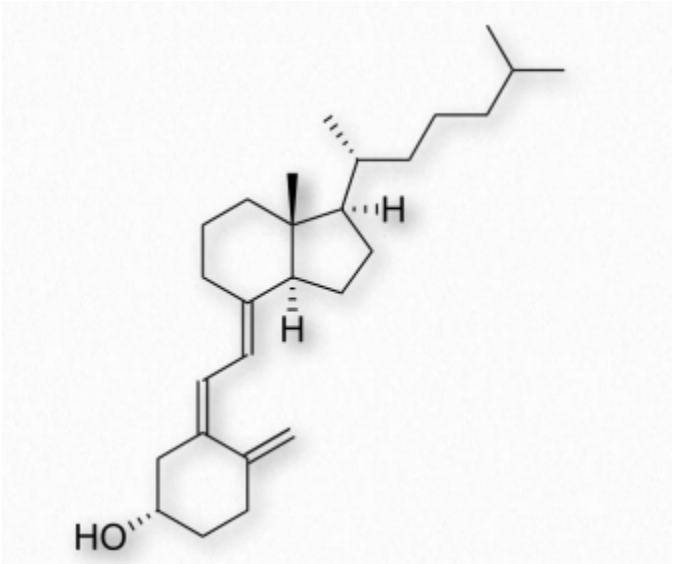


Figure 2.229 - Cholecalciferol - Vitamin D3

The active form of vitamin D plays important roles in the intestinal absorption of calcium and phosphate and thus in healthy bones. Technically, vitamin D isn’t even a vitamin, as it is a compound made by the body. Rather, it behaves more like a hormone.

Derived from ultimately from cholesterol, vitamin D can be made in a reaction catalyzed by ultraviolet light. In the reaction, the intermediate 7-dehydrocholesterol is converted to cholecalciferol (vitamin D3) by the uv light (Figure 2.229). The reaction occurs most readily in the bottom two layers of the skin shown in Figure 2.230.

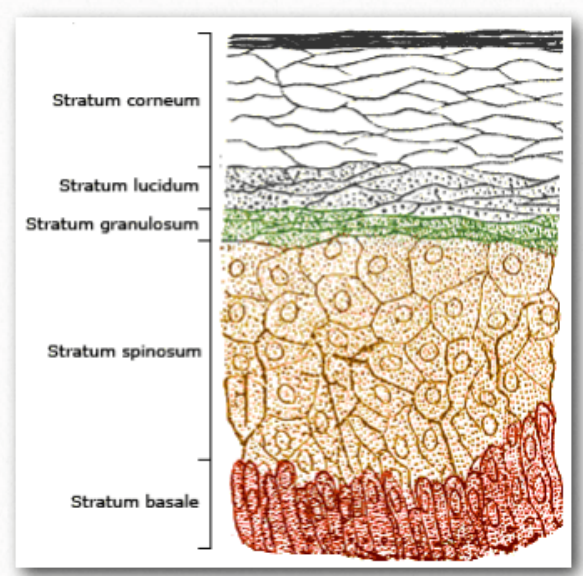


Figure 2.230 - Layers of the skin. Outside is at top.

Forms of vitamin D

Five different compounds are referred to as vitamin D. They are

Vitamin D1 - A mixture of ergocalciferol and lumisterol

Vitamin D2 - Ergocalciferol

Vitamin D3 - Cholecalciferol Vitamin

D4 - 22-Dihydroergocalciferol Vitamin

D5 - Sitocalciferol

Vitamin D3 is the most common form used in vitamin supplements and it and vitamin D2 are commonly obtained in the diet, as well. The active form of vitamin D, calcitriol (Figure 2.231), is made in the body in controlled amounts. This proceeds through two steps from cholecalciferol. First, a hydroxylation in the liver produces calcidiol and a second hydroxylation in the kidney produces calcitriol. Monocyte macrophages can also synthesize vitamin D and they use is as a cytokine to stimulate the innate immune system.

Mechanism of action

Calcitriol moves in the body bound to a vitamin D binding protein, which delivers it to target organs. Calcitriol inside of cells acts by binding a vitamin D receptor (VDR), which results in most of the vitamin's physiological effects. After binding calcitriol, the VDR migrates to the nucleus where it acts as a transcription factor to control levels of expression of calcium transport proteins (for example) in the intestine. Most tissues respond to VDR bound to calcitriol and the result is moderation of calcium and phosphate levels in cells.

Deficiency/excess

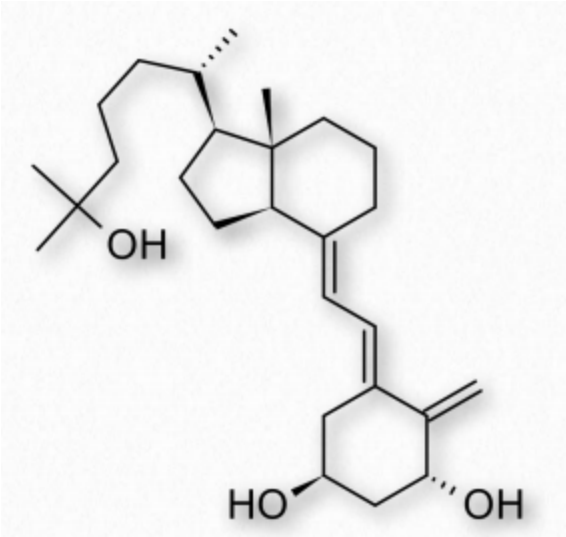


Figure 2.231 - Calcitriol - Active form of vitamin D

Deficiency of vitamin D is a cause of the disease known as rickets, which is characterized by soft, weak bones and most commonly is found in children. It is not common in the developed world, but elsewhere is of increasing concern.

Excess of vitamin D is rare, but has toxic effects, including hypercalcemia, which results in painful calcium deposits in major organs. Indications of vitamin D toxicity are increased urination and thirst. Vitamin D toxicity can lead to mental retardation and many other serious health problems.

Vitamin E

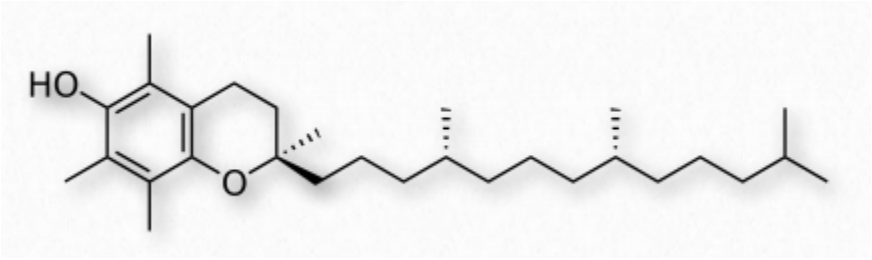


Figure 2.232 α-tocopherol - The most biologically active form of vitamin E Figure

Vitamin E comprises a group of two compounds (tocopherols and tocotrienols - Figure 2.232) and stereoisomers of each. It is commonly found in plant oils. The compounds act in cells as fat-soluble antioxidants. α-tocopherol (Figure 2.233), the most active form of the vitamin, works 1) through the glutathione peroxidase protective system and 2) in membranes to interrupt lipid peroxidation chain reactions. In both actions, vitamin E reduces levels of reactive oxygen species in cells.

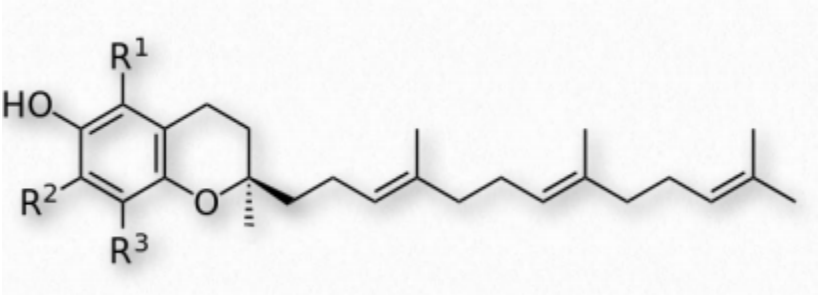


Figure 2.233 - Structure of tocotrienols

Action

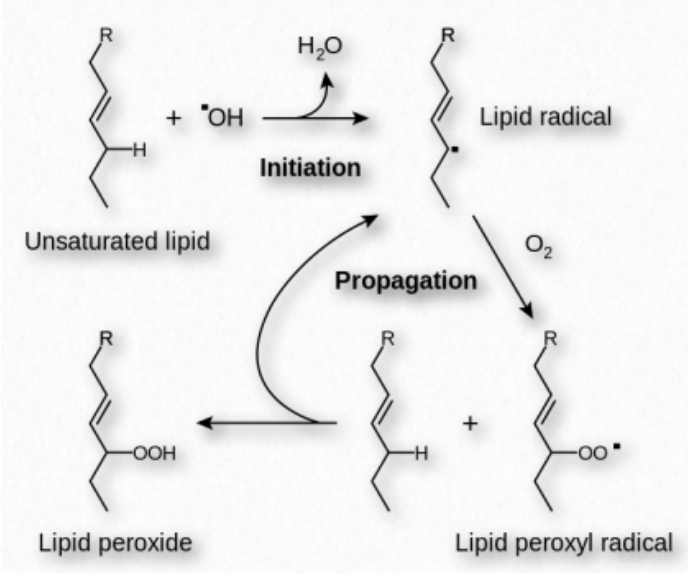


Figure 2.234 - Lipid peroxidation reactions

Vitamin E scavenges oxygen radicals (possessing unpaired electrons) by reacting with them to produce a tocopheryl radical. This vitamin E radical can be converted back to its original form by a hydrogen donor. Vitamin C is one such donor. Acting in this way, Vitamin E helps reduce oxidation of easily oxidized compounds in the lipid peroxidation reactions (Figure 2.234).

Vitamin E also can affect enzyme activity. The compound can inhibit action of protein kinase C in smooth muscle and simultaneously activate catalysis of protein phosphatase 2A to remove phosphates, stopping smooth muscle growth.

Deficiency/excess

Deficiency of vitamin E can lead to poor conduction of nerve signals and other issues arising from nerve problems. Low levels of the vitamin may be a factor in low birth weights and premature deliveries. Deficiency, however, is rare, and not usually associated with diet.

Excess Vitamin E reduces vitamin K levels, thus reducing the ability to clot blood. Hypervitaminosis of vitamin E in conjunction with aspirin can be life threatening. At lower levels, vitamin E may serve a preventative role with respect to atherosclerosis by reducing oxidation of LDLs, a step in plaque formation.

Vitamin K

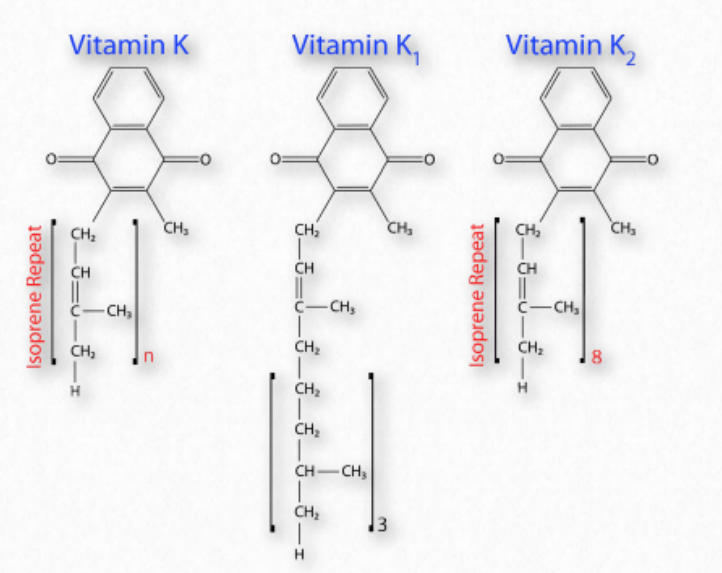


Figure 2.235 - Forms of vitamin K Image by Pehr Jacobson

Like the other fat-soluble vitamins, Vitamin K comes in multiple forms (Figure 2.235) and is stored in fat tissue in the body. There are two primary forms of the vitamin - K1 and K2 and the latter has multiple sub-forms . Vitamins K3, K4, and K5 are made synthetically, not biologically.

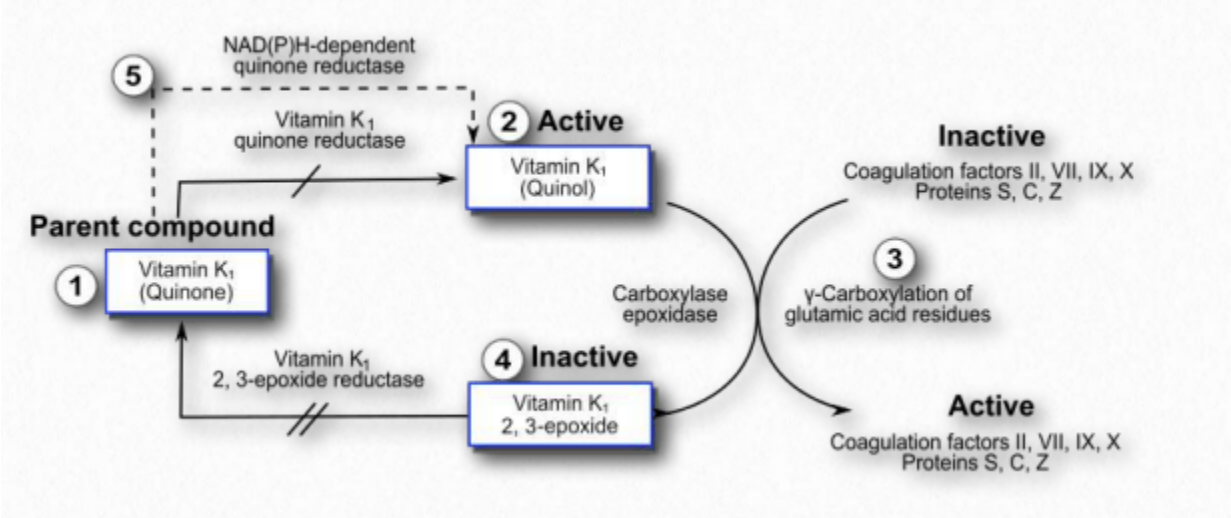


Figure 2.236 - Recycling of vitamin K Wikipedia

Action

Vitamin K is used as a co-factor for enzymes that add carboxyl groups to glutamate side chains of proteins to increase their affinity for calcium. Sixteen such proteins are known in humans. They include proteins involved in blood clotting (prothrombin (called Factor II), Factors VII, IX, and X), bone metabolism (osteocalcin, also called bone Gla protein (BGP), matrix Gla protein (MGP), and periostin) and others.

Modification of prothrombin is an important step in the process of blood clotting (see HERE). Reduced levels of vitamin K result in less blood clotting, a phenomenon sometimes referred to as blood thinning. Drugs that block recycling of vitamin K (Figure 2.236) by inhibiting the vitamin K epoxide reductase, produce lower levels of the vitamin and are employed in treatments for people prone to excessive clotting. Warfarin (coumadin) is one such compound that acts in this way and is used therapeutically. Individuals respond to the drug differentially, requiring them to periodically be tested for levels of clotting they possess, lest too much or too little occur.

Sources

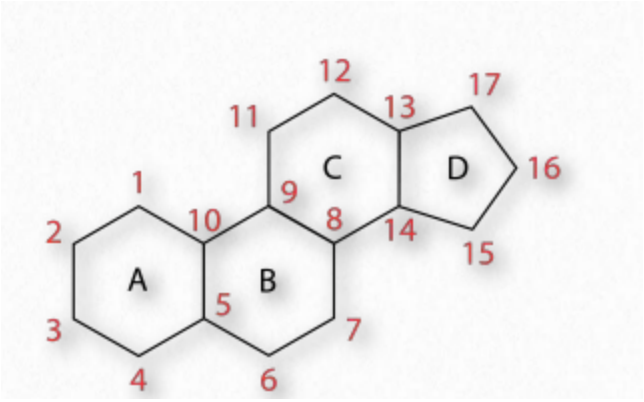


Figure 2.237 - Steroid numbering scheme Image by Pehr Jacobson

Vitamin K1 is a stereoisomer of the plant photosystem I electron receptor known as phylloquinone and is found abundantly in green, leafy vegetables. Phylloquinone is one source of vitamin K, but the compound binds tightly to thylakoid membranes and tends to have low bioavailability. Vitamin K2 is produced by microbes in the gut and is a primary source of the vitamin. Infants in the first few days before they establish their gut flora and people taking broad spectrum antibiotics may have reduced levels, as a result. Dietary deficiency is rare in the absence of damage to the small bowel. Others at risk of deficiency include people with chronic kidney disease and anyone suffering from a vitamin D deficiency. Deficiencies produce symptoms of easy bruising, heavy menstrual bleeding, anemia, and nosebleeds.

Steroids

Steroids, such as cholesterol are found in membranes and act as signaling hormones in traveling through the body.

Steroid hormones are all made from cholesterol and are grouped into five categories - mineralocorticoids (21 carbons), glucocorticoids (21 carbons), progestagens (21 carbons), androgens (19 carbons), and estrogens (18 carbons).

Mineralocorticoids

Mineralocorticoids are steroid hormones that influence water and electrolyte balances. Aldosterone (Figure 2.238) is the primary mineralocorticoid hormone, though other steroid hormones (including progesterone) have some functions like it. Aldosterone stimulates kidneys to reabsorb sodium, secrete potassium, and passively reabsorb water. These actions have the effect of increasing blood pressure and blood volume. Mineralocorticoids are produced by the zona glomerulosa of the cortex of the adrenal gland.

Glucocorticoids

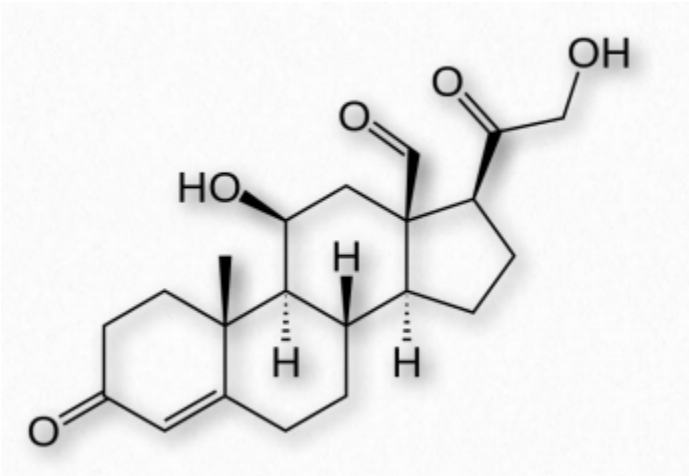


Figure 2.238 - Aldosterone - A mineralocorticoid

Glucocorticoids (GCs) bind to glucocorticoid receptors found in almost every vertebrate animal cell and act in a feedback

mechanism in the immune system to reduce its activity. GCs are used to treat diseases associated with overactive immune systems. These include allergies, asthma, and autoimmune dis- Figure 2.237 - Steroid numbering scheme Image by Pehr Jacobson eases. Cortisol (Figure 2.239) is an important glucocorticoid with cardiovascular, metabolic, and immunologic functions. The synthetic glucocorticoid known as dexamethasone has medical applications for treating rheumatoid arthritis, bronchospasms (in asthma), and inflammation due to its increased potency (25-fold) compared to cortisol. Glucocorticoids are produced primarily in the zona fasciculata of the adrenal cortex.

Progestagens

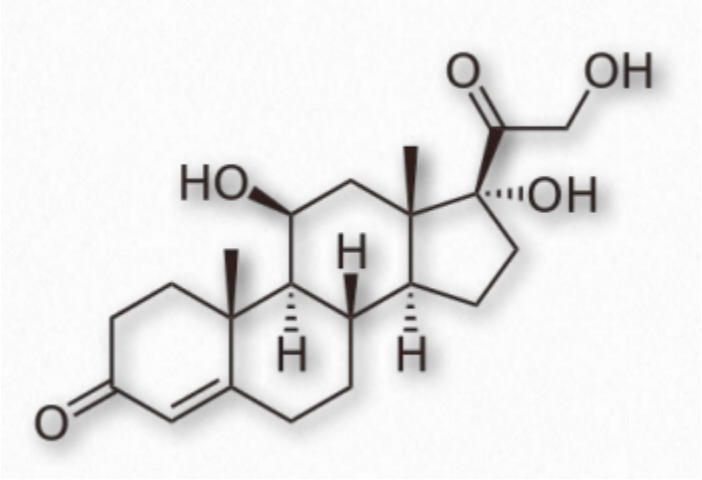


Figure 2.239 - Cortisol - A glucocorticoid

Progestagens (also called gestagens) are steroid hormones that work to activate the progesterone receptor upon binding to it. Synthetic progestagens are referred to as progestins. The most common progestagen is progesterone (also called P4 - Figure 2.240) and it has functions in maintaining pregnancy. Progesterone is produced primarily in the diestrus phase of the estrous cycle by the corpus luteum of mammalian ovaries. In pregnancy, the placenta takes over most progesterone production.

Androgens

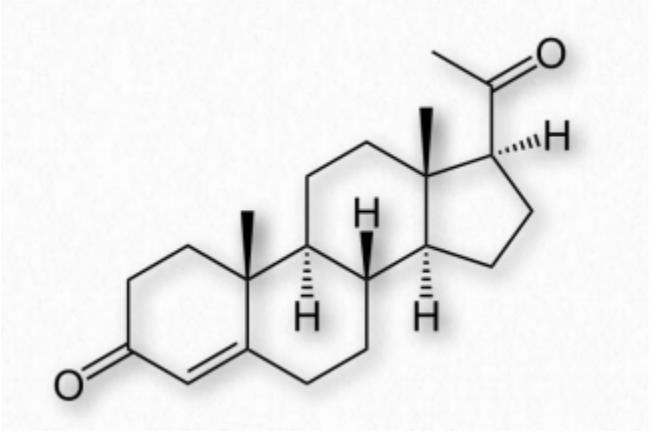


Figure 2.240 Progesterone - A progestagen

Androgens are steroid hormones that act by binding androgen receptors to stimulate development and maintenance of male characteristics in vertebrates. Androgens are precursors of estrogens (see below). The primary androgen is testosterone (Figure 2.241). Other important androgens include dihydrotestosterone (stimulates differentiation of penis, scrotum, and prostate in embryo) and androstenedione (common precursor of male and female hormones).

Estrogens

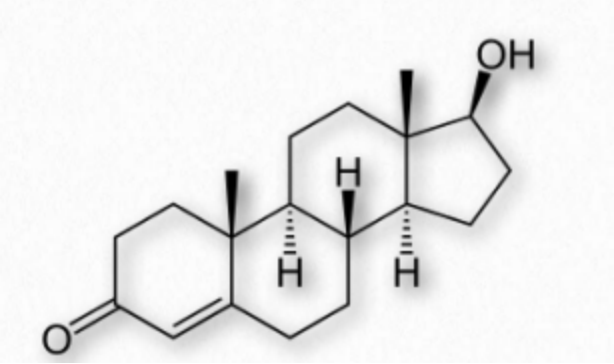


Figure 2.241 - Testosterone - An androgen

The estrogen steroid hormones are a class of compounds with important roles in menstrual and estrous cycles. They are the most important female sex hormones. Estrogens act by activating estrogen receptors inside of cells. These receptors, in turn, affect expression of many genes. The major estrogens in women include estrone (E1), estradiol (E2 - Figure 2.242), and estriol (E3). In the reproductive years, estradiol predominates. During pregnancy, estriol predominates and during menopause, estrone is the major estrogen.

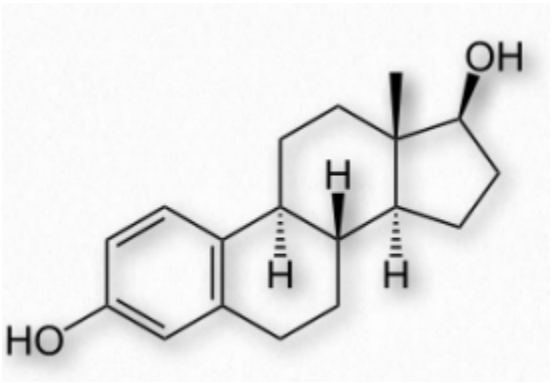


Figure 2.242 - Estradiol - An estrogen

Estrogens are made from the androgen hormones testosterone and androstenedione in a reaction catalyzed by the enzyme known as aromatase. Inhibition of this enzyme with aromatase inhibitors, such as exemestane, is a strategy for stopping estrogen production. This may be part of a chemotherapeutic treatment when estrogenresponsive tumors are present.

Cannabinoids

Cannabinoids are a group of chemicals that bind to and have effects on brain receptors (cannabinoid receptors), repressing neurotransmitter release. Classes of these compounds include endocannabinoids (made in the body), phytocannabinoids (made in plants, such as marijuana), and synthetic cannabinoids (man-made).

Endocannabinoids are natural molecules derived from arachidonic acid. Cannabinoid receptors are very abundant, comprising the largest number of G-protein- 247 Figure 2.243 - Tetrahydrocannabinol - Active ingredient in marijuana coupled receptors found in brain. The best known phytocannabinoid is Δ-9- tetrahydrocannabinol (THC), the primary psychoactive ingredient (of the 85 cannabinoids) of marijuana (Figure 2.243).

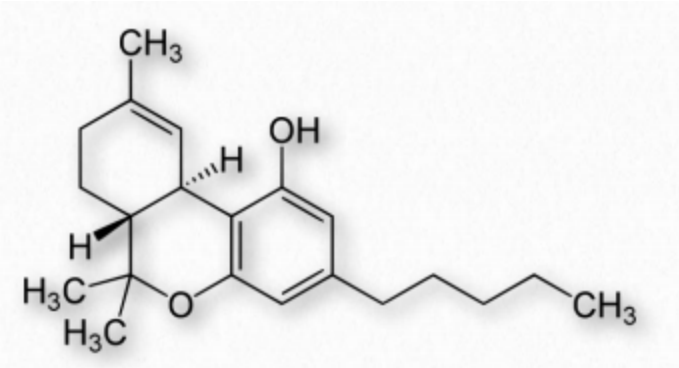


Figure 2.243 - Tetrahydrocannabinol - Active ingredient in marijuana

Anandamide

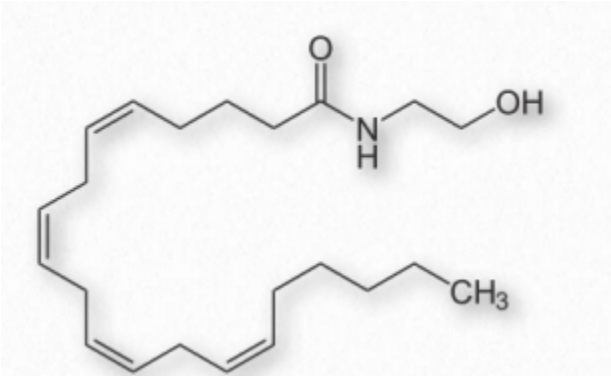


Figure 2.244 - Anandamide - An endocannabinoid

Anandamide (N-arachidonoylethanolamine - Figure 2.244) is an endocannabinoid neurotransmitter derived from arachidonic acid. It exerts its actions primarily through the CB1 and CB2 cannabinoid receptors, the same ones bound by the active ingredient in marijuana, Δ9-tetrahydrocannabinol. Anandamide has roles in stimulating eating/appetite and affecting motivation and pleasure. It has been proposed to play a role in "runners high," an analgesic effect experienced from exertion, especially among runners. Anandamide appears to impair memory function in rats.

Anandamide has been found in chocolate and two compounds that mimic its effects (N-oleoylethanolamine and Nlinoleoylethanolamine) are present as well. The enzyme fatty acid amide hydrolase (FAAH) breaks down anandamide into free arachidonic acid and ethanolamine.

Lipoxins

Lipoxins (Figure 2.245) are eicosanoid compounds involved in modulating immune responses and they have anti-inflammatory effects. When lipoxins appear in inflammation it begins the end of the process. Lipoxins act to attract macrophages to apoptotic cells at the site of inflammation and they are engulfed. Lipoxins further act to start the resolution phase of the inflammation process.

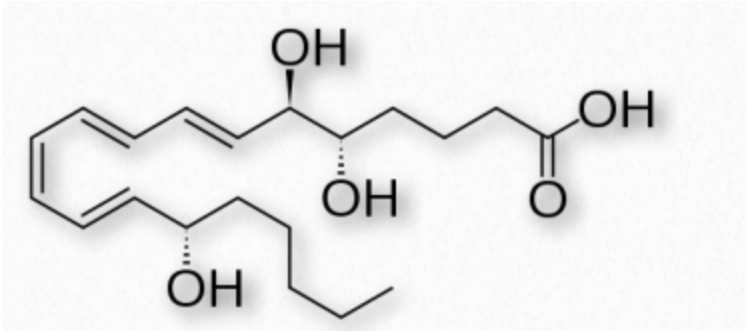


Figure 2.245 - Lipoxin A4

At least one lipoxin (aspirin-triggered LX4) has its synthesis stimulated by aspirin. This occurs as a byproduct of aspirin's acetylation of COX-2. When this occurs, the enzyme's catalytic activity is re-directed to synthesis of 15R-hydroxyeicosatetraenoic acid (HETE) instead of prostaglandins. 15R-HETE is a procursor of 15-epimer lipoxins, including aspirin-triggered LX4.

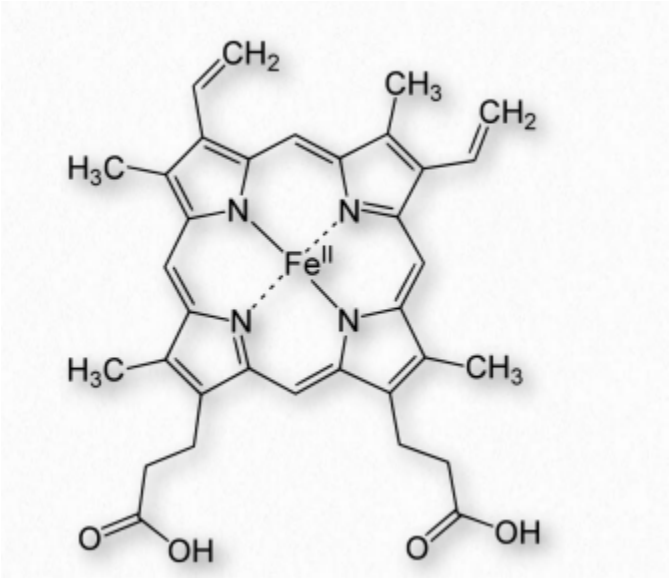


Figure 2.246 - Structure of heme B

Heme

Heme groups are a collection of protein/ enzyme cofactors containing a large heterocyclic aromatic ring known as a porphyrin ring with a ferrous (Fe++) ion in the middle. An example porphyrin ring with an iron (found in Heme B of hemoglobin), is shown in Figure 2.246. When contained in a protein, these are known collectively as hemoproteins (Figure 2.247).

Heme, of course, is a primary component of hemoglobin, but it is also found in other proteins, such as myoglobin, cytochromes, and the enzymes catalase and succinate dehydrogenase. Hemoproteins function in oxygen transport, catalysis, and electron transport. Heme is synthesized in the liver and bone marrow in a pathway that is conserved across a wide range of biology.

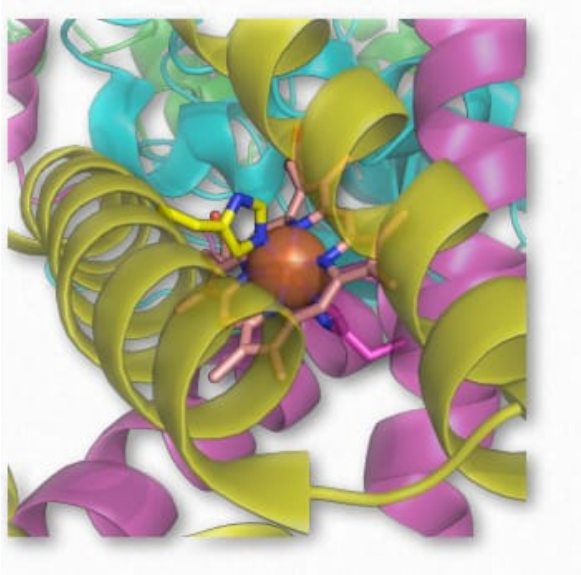


Figure 2.247 - Heme embedded in the succinate dehydrogenase hemoprotein Wikipedia

Porphobilinogen

Porphobilinogen (Figure 2.248) is a pyrrole molecule involved in porphyrin metabolism. It is produced from aminolevulinate by action of the enzyme known as ALA dehydratase. Porphobilinogen is acted upon by the enzyme porphobilinogen deaminase. Deficiency of the latter enzyme (and others in porphyrin metabolism) can result in a condition known as porphyria, which results in accumulation of porphobilinogen in the cytoplasm of cells.

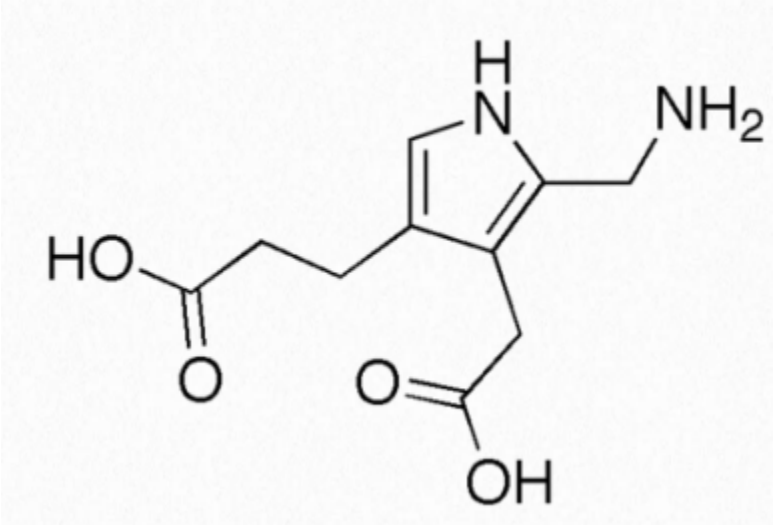


Figure 2.248 - Porphobilinogen

The disease can manifest itself with acute abdominal pain and numerous psychiatric issues. Both Vincent van Gogh and King ` George III are suspected to have suffered from porphyria, perhaps causing the “madness of King George III.” Porphyria is also considered by some to be the impetus for the legend of vampires seeking blood from victims, since the color of the skin in non-acute forms of the disease can be miscolored, leading some to perceive that as a deficiency of hemoglobin and hence the “thirst” for blood imagined for vampires.

Dolichols

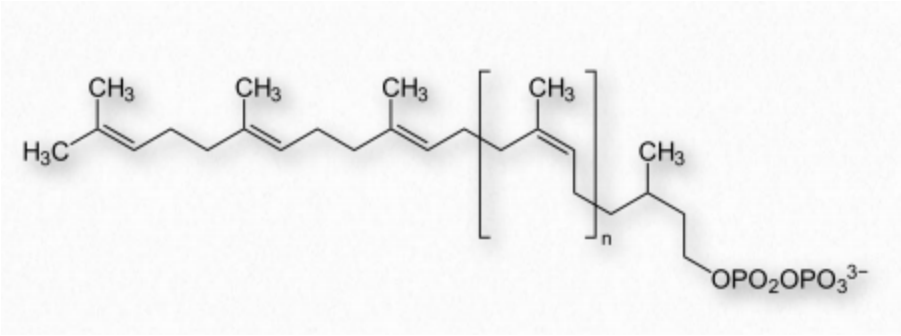


Figure 2.249 - Structure of dolichol pyrophosphate

Dolichol is a name for a group of non-polar molecules made by combining isoprene units together. Phosphorylated forms of dolichols play central roles in the N-glycosylation of proteins. This process, which occurs in the endoplasmic reticulum of eukaryotic cells, begins with a membrane-embedded dolichol pyrophosphate (Figure 2.249) to which an oligosaccharide is attached (also see HERE). This oligosaccharide contains three molecules of glucose, nine molecules of mannose and two molecules of N-acetylglucosamine.

Interestingly, the sugars are attached to the dolichol pyrophosphate with the pyrophosphate pointing outwards (away from) the endoplasmic reticulum, but after attachment, the dolichol complex flips so that the sugar portion is situated on the inside of the endoplasmic reticulum. There, the entire sugar complex is transferred to the amide of an asparagine side chain of a target protein.

The only asparagine side chains to which the attachment can be made are in proteins where the sequences Asn-X-Ser or Asn-X-Thr occur. Sugars can be removed/added after the transfer to the protein. Levels of dolichol in the human brain increase with age, but in neurodegenerative diseases, they decrease.

Terpenes



Figure 2.250 - Pine tree resin - A source of terpenes Wikipedia

Terpenes are members of a class of nonpolar molecules made from isoprene units. Terpenes are produced primarily by plants and by some insects. Terpenoids are a related group of molecules that contain functional groups lacking in terpenes.

Terpenes have a variety of functions. In plants, they often play a defensive role protecting from insects. The name of terpene comes from turpentine, which has an odor like some of the terpenes. Terpenes are common components of plant resins (think pine) and they are widely used in medicines and as fragrances. Hops, for example, gain some of their distinctive aroma and flavor from terpenes. Not all terpenes, however have significant odor.

Synthesis

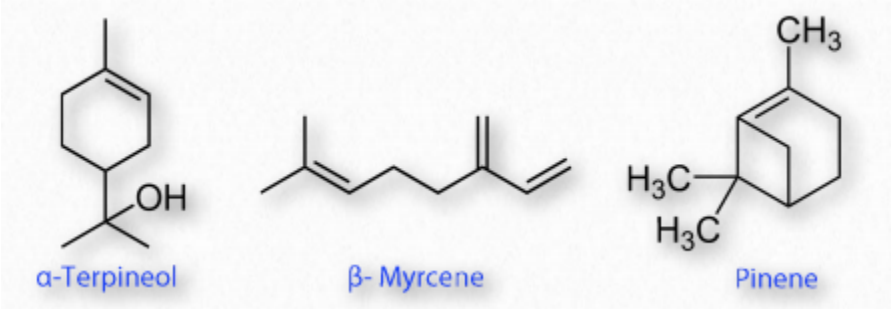


Figure 2.251 - Three monoterpenes

Terpenes, like steroids, are synthesized starting with simple building blocks known as isoprenes. There are two of them - dimethylallyl pyrophosphate and the related isopentenyl pyrophosphate and (Figures 2.252 and 2.253) which combine 1-2 units at a time to make higher order structures. Terpene synthesis overlaps and includes steroid synthesis.

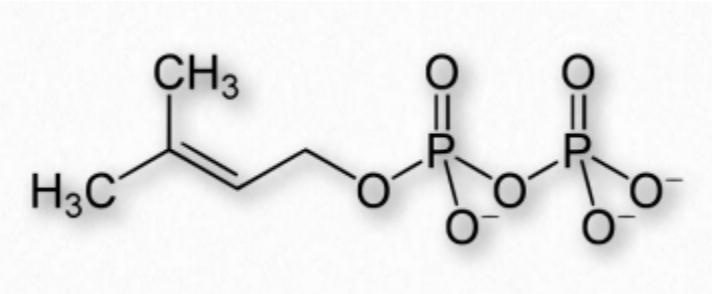


Figure 2.252 - Dimethylallyl pyrophosphate

Terpenes and terpenoids are classified according to how many isoprene units they contain. They include hemiterpenes (one unit), monoterpenes (two units), sesquiterpenes (three units), diterpenes (four units), sesterterpenes (five units), triterpenes (six units), sesquarterpenes (seven units), tetraterpenes (eight units), polyterpenes (many units). Another class of terpene-containing molecules, the norisoterpenoids arise from peroxidase-catalyzed reactions on terpene molecules.

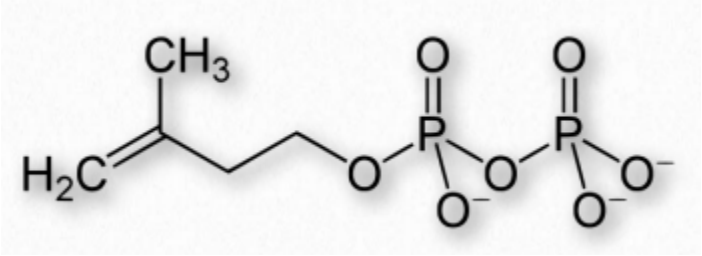


Figure 2.253 - Isopentenyl pyrophosphate

Examples

Common terpenes include monoterpenes of terpineol (lilacs), limonene (citrus), myrcene (hops), linalool (lavender), and pinene (pine). Higher order terpenes include taxadiene (diterpene precursor of taxol), lycopene (tetraterpenes), carotenes (tetraterpenes), and natural rubber (polyterpenes).

Steroid precursors geranyl pyrophosphate (monoterpene derivative), farnesyl pyrophosphate (sesquiterpene derivative), and squalene (triterpene) are all terpenes or derivatives of them. Vitamin A and phytol are derived from diterpenes.

Caffeine

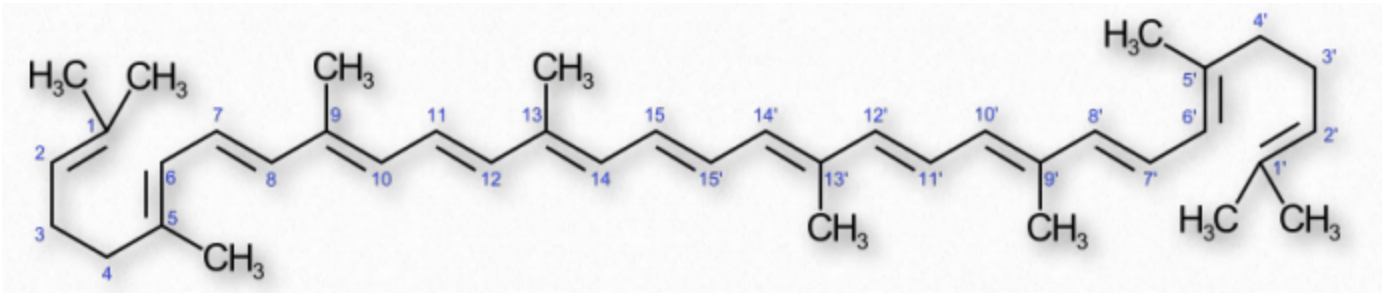


Figure 2.254 - Lycopene - A tetraterpene

Caffeine is the world's most actively consumed psychoactive drug (Figure 2.255). A methylxanthine alkaloid, caffeine is closely related to adenine and guanine and this is responsible for many effects on the body. Caffeine blocks the binding of adenosine on its receptor and consequently prevents the onset of drowsiness induced by adenosine. Caffeine readily crosses the blood-brain barrier and stimulates release of neurotransmitters. Caffeine stimulates portions of the autonomic nervous system and inhibits the activity of phosphodiesterase. The latter has the result of raising cAMP levels in cells, which activates protein kinase A and activates glycogen breakdown, inhibits TNF-α and leukotriene synthesis, which results in reduction of inflammation and innate immunity.

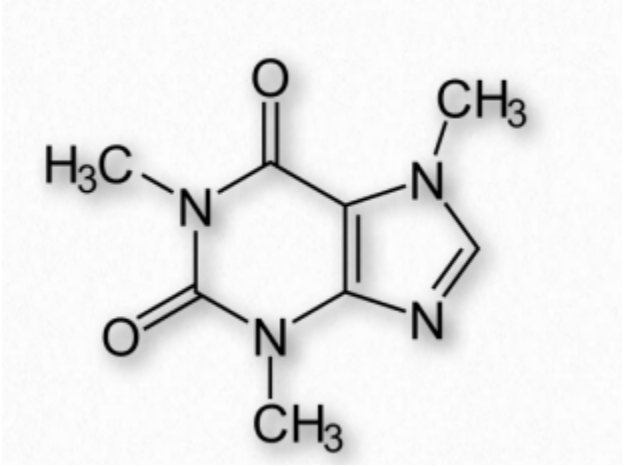


Figure 2.255 - Caffeine

Caffeine also has effects on the cholinergic system (acetylcholinesterase inhibitor), is an inositol triphosphate receptor 1 antagonist, and is a voltage independent activator of ryanodin receptors (a group of calcium channels found in skeletal muscle, smooth muscle, and heart muscle cells).

The half-life of caffeine in the body varies considerably. In healthy adults, it has a half-life of about 3-7 hours. Nicotine decreases the half-life and contraceptives and pregnancy can double it. The liver metabolizes caffeine, so the health of the liver is a factor in the halflife. CYP1A2 of the cytochrome P450 oxidase enzyme is primarily responsible. Caffeine is a natural pesticide in plants, paralyzing predator bugs.

Lipoprotein complexes and lipid movement in the body

Lipoprotein complexes are combinations of apolipoproteins and lipids bound to them that solubilize fats and other non-polar molecules, such as cholesterol, so they can travel in the bloodstream between various tissues of the body. The apolipoproteins provide the emulsification necessary for this. Lipoprotein complexes are formed in tiny "balls" with the water soluble apolipoproteins on the outside and non-polar lipids, such as fats, cholesteryl esters, and fat soluble vitamins on the inside.

They are categorized by their densities. These include (from highest density to the lowest) high density lipoproteins (HDLs), Low Density Lipoproteins (LDLs), Intermediate Density Lipoproteins (IDLs), Very Low Density Lipoproteins (VLDLs) and the chylomicrons. These particles are synthesized in the liver and small intestines.

Apolipoproteins

Name	Lipoprotein Complex(es)	Function
ApoA-I	HDL	Promotes Fat Movement to Liver
ApoA-II	HDL	Inhibit LCAT
ApoA-IV	Chylomicrons / HDL	Activate LCAT
AboB-48	Chylomicrons	Cholesterol Transport
AboB-100	VLDL / LDL	Bind LDL Receptor
ApoC-I	VLDL / LDL	Unknown
ApoC-II	All	Activate Lipoprotein Lipase
ApoC-III	All	Inhibit Lipoprotein Lipase
ApoD	HDL	Unknown
ApoE	VLDL / Chylomicrons / HDL	Clearance of Chylomicrons Remnants and VLDLs

Figure 2.256 - Apolipoproteins

Each lipoprotein complex contain a characteristic set of apolipoproteins, as shown in Figure 2.256. ApoC-II and ApoC-III are notable for their presence in all the lipoprotein complexes and the roles they play in activating (ApoC-II) or inactivating (ApoC-III) lipoprotein lipase. Lipoprotein lipase is a cellular enzyme that catalyzes the breakdown of fat from the complexes. ApoE (see below) is useful for helping the predict the likelihood of the occurrence of Alzheimer's disease in a patient.

Gene editing



Figure 2.257 - ApoA-I

ApoB-48 and ApoB-100 are interesting in being coded by the same gene, but a unique mRNA sequence editing event occurs that converts one into the other. ApoB-100 is made in the liver, but ApoB-48 is made in the small intestine. The small intestine contains an enzyme that deaminates the cytidine at nucleotide #2153 of the common mRNA. This changes it to a uridine and changes the codon it is in from CAA (codes for glutamine) to UAA (stop codon). The liver does not contain this enzyme and does not make the change in the mRNA. Consequently, a shorter protein is synthesized in the intestine (ApoB-48) than the one that is made in the liver (ApoB-100) using the same gene sequence in the DNA.

Movement

The movement of fats in the body is important because they are not stored in all cells. Only specialized cells called adipocytes

store fat. There are three relevant pathways in the body for moving lipids. As described below, they are 1) the exogenous pathway; 2) the endogenous pathway, and 3) the reverse transport pathway.

Exogenous pathway

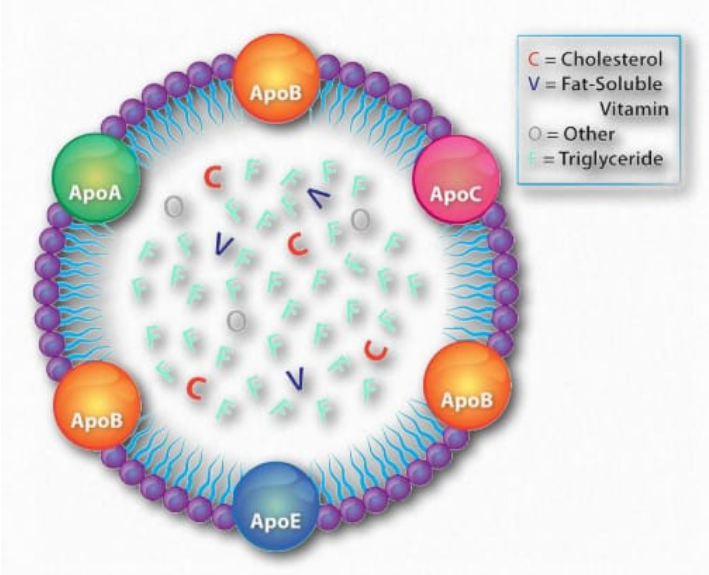


Figure 2.258 - Schematic diagram of a chylomicron Image by Aleia Kim

Dietary fat entering the body from the intestinal system must be transported, as appropriate, to places needing it or storing it. This is the function of the exogenous pathway of lipid movement in the body. All dietary lipids (fats, cholesterol, fat soluble vitamins, and other lipids) are moved by it. In the case of dietary fat, it begins its journey after ingestion first by being solubilized by bile acids in the intestinal tract. After passing through the stomach, pancreatic lipases clip two fatty acids from the fat, leaving a monoacyl glycerol. The fatty acids and monoacyl glycerol are absorbed by intestinal cells (enterocytes) and reassembled back into a fat, and then this is mixed with phospholipids, cholesterol esters, and apolipoprotein B-48 and processed to form chylomicrons (Figures 2.258 & 2.259) in the Golgi apparatus and smooth endoplasmic reticulum.

Exocytosis

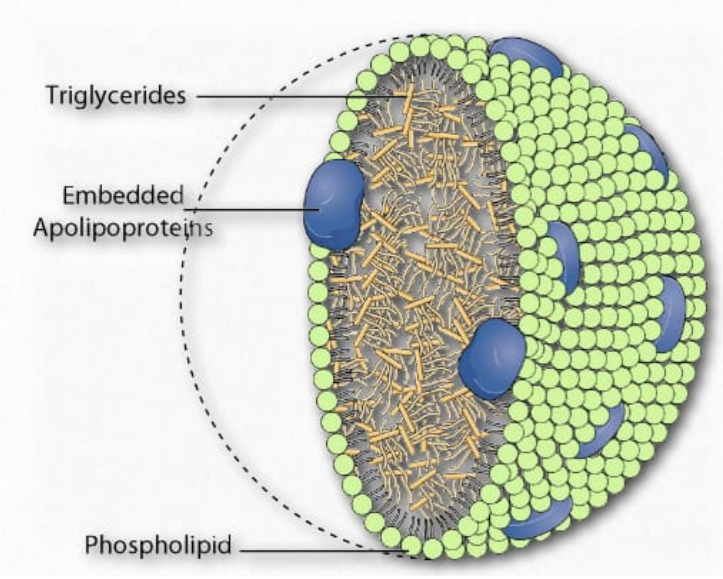


Figure 2.259 - Another perspective of a chylomicron WIkipedia

These are exocytosed from the cell into lymph capillaries called lacteals. The chylomicrons pass through the lacteals and enter the bloodstream via the left subclavian vein. Within the bloodstream, lipoprotein lipase breaks down the fats causing the chylomicron

to shrink and become what is known as a chylomicron remnant. It retains its cholesterol and other lipid molecules.

The chylomicron remnants travel to the liver where they are absorbed (Figure 2.260). This is accomplished by receptors in the liver that recognize and bind to the ApoE of the chylomicrons. The bound complexes are then internalized by endocytosis, degraded in the lysosomes, and the cholesterol is disbursed in liver cells.

Endogenous pathway

The liver plays a central role in managing the body's needs for lipids. When lipids are needed by the body or when the capacity of the liver to contain more lipids than is supplied by the diet, the liver packages up fats and cholesteryl esters into Very Low Density Lipoprotein (VLDL) complexes and exports them via the endogenous pathway. VLDL complexes contain ApoB-100, ApoC-I, ApoC-II, ApoC-III, and ApoE apolipoproteins. VLDLs enter the blood and travel to muscles and adipose tissue where lipoprotein lipase is activated by ApoC-II. In the muscle cells, the released fatty acids are taken up and oxidized. By contrast, in the adipoctyes, the fatty acids are taken up and reassembled back into triacylglycerides (fats) and stored in fat droplets. Removal of fat from the VLDLs causes them to shrink, first to Intermediate Density Lipoprotein (IDL) complexes (also called VLDL remnants) and then to Low Density Lipoprotein (LDL) complexes.

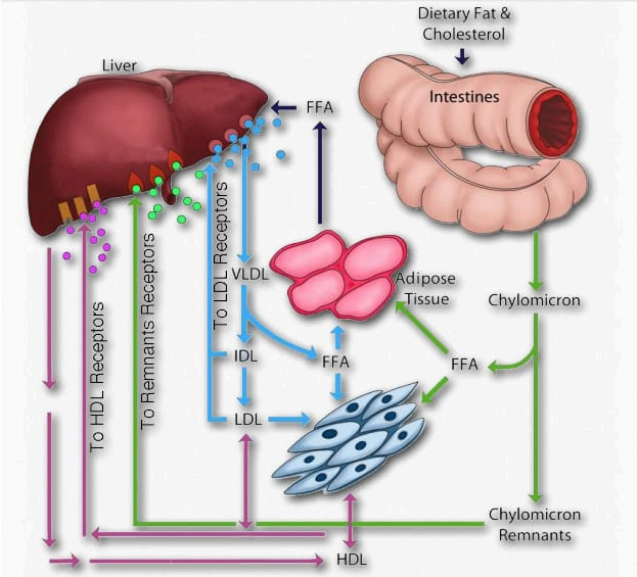


Figure 2.260 - Movement of lipids in the body - Green = exogenous pathway; Blue = endogenous pathway; Purple = reverse transport pathway
Image by Aleia Kim

Shrinking of VLDLs is accompanied by loss of apolipoproteins so that LDLs are comprised primarily of ApoB-100. This lipoprotein complex is important because cells have receptors for it to bind and internalize it by receptor-mediated endocytosis (Figure 2.261). Up until this point, cholesterol and cholesteryl esters have traveled in chylomicrons, VLDLs, and IDLs as fat has been stripped stripped away. For cholesterol compounds to get into the cell from the lipoprotein complexes, they must be internalized by cells and that is the job of receptormediated endocytosis.

Reverse transport pathway

Another important consideration of the movement of lipids in the body is the reverse transport pathway (Figure 2.260). It is also called the reverse cholesterol transport pathway, since cholesterol is the primary molecule involved. This pathway involves the last class of lipoprotein complexes known as the High Density Lipoproteins (HDLs). In contrast to the LDLs, which are commonly referred to as "bad cholesterol" (see below also), the HDLs are known as "good cholesterol."

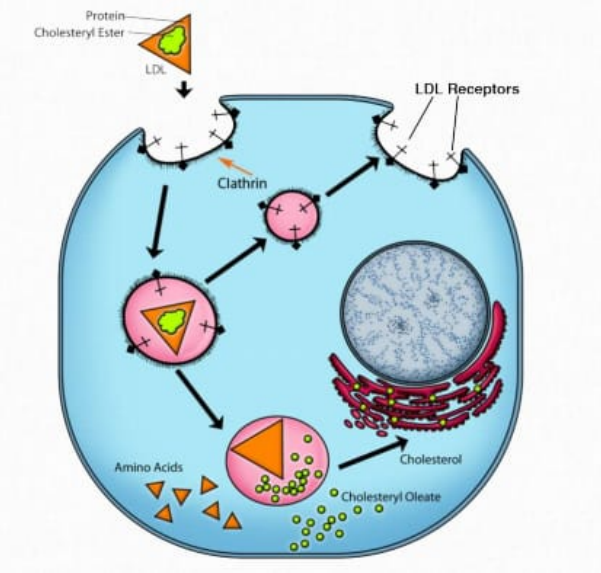


Figure 2.261 - The process of receptor-mediated endocytosis Image by Aleia Kim

HDLs are synthesized in the liver and small intestine. They contain little or no lipid when made (called depleted HDLs), but serve the role of “scavenger” for cholesterol in the blood and from remnants of other (damaged) lipoprotein complexes in the blood. To perform its task, HDLs carry the enzyme known as lecithincholesterol acyl transferase (LCAT), which they use to form cholesteryl esters using fatty acids from lecithin (phosphatidylcholine) and then they internalize them.

The cholesterol used for this purpose comes from the bloodstream, from macrophages, and from foam cells (macrophage-LDL complexes - Figure 2.262). Addition of cholesteryl esters causes the HDL to swell and Figure 2.261 - The process of receptor-mediated endocytosis Image by Aleia Kim when it is mature, it returns its load of cholesterol back to the liver or, alternatively, to LDL molecules for endocytosis. HDLs have the effect of lowering levels of cholesterol and it is for that reason they are described as “good cholesterol.”

Regulation of lipid transfer

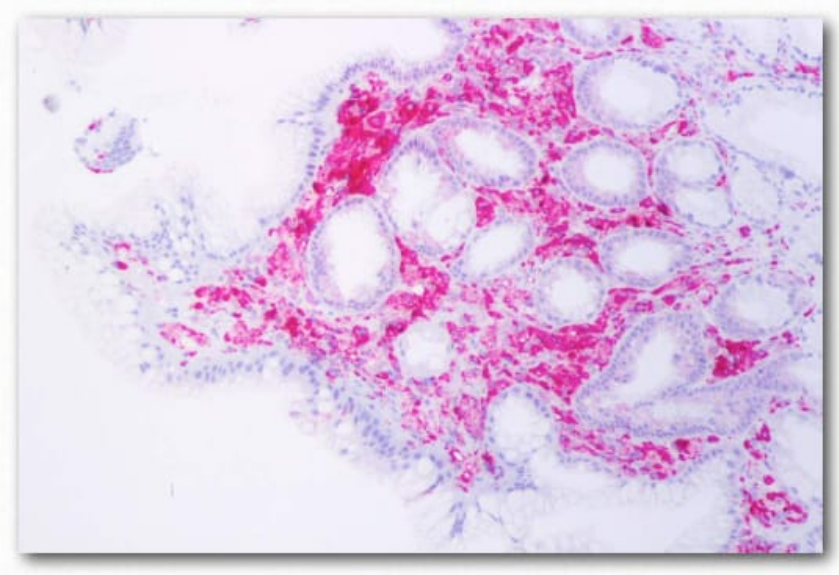


Figure 2.262 - Foam cell aggregate Wikipedia

It is important that cells get food when they need it so some control of the movement of nutrients is critical. The liver, which plays the central role in modulating blood glucose levels, is also important for performing the same role for lipids. It accomplishes this task the use of specialized LDL receptors on its surface. Liver LDL receptors bind LDLs that were not taken up by other cells in their path through the bloodstream. High levels of LDLs are a signal to the liver to reduce the creation of VLDLs for release.

People with the genetic disease known as familial hypercholesterolemia, which manifests with dangerously high levels of LDLs, lack properly functioning LDL receptors on their liver cells.Figure 2.263 - Progression of atherosclerosis Wikipedia Figure 2.263 - Progression of atherosclerosis Wikipedia Figure 2.263 - Progression of atherosclerosis Wikipedia

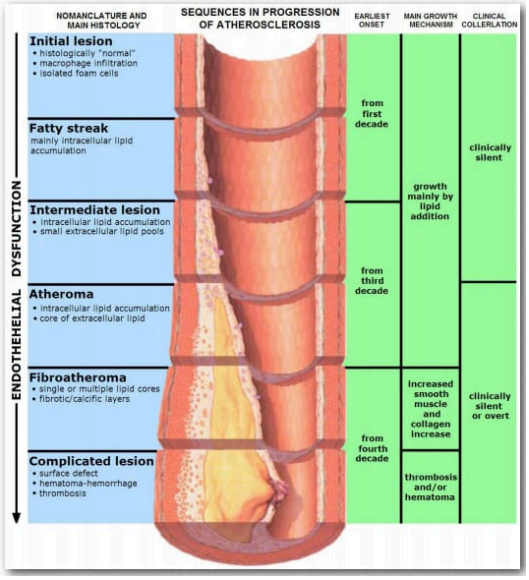


Figure 2.263 - Progression of atherosclerosis Wikipedia

In sufferers of this disease, the liver never gets the signal that the LDL levels are high. In fact, to the liver, it appears that all VLDLs and LDLs are being taken up by peripheral tissues, so it creates more VLDLs to attempt to boost levels. Untreated, the disease used to be fatal early, but newer drugs like the statins have significantly increased life spans of patients. Cellular needs for the contents of LDLs are directly linked to the levels of synthesis of LDL receptors on their membranes. As cells are needing more cholesterol, their synthesis of components for receptors goes up and it decreases as need diminishes.

Good cholesterol / bad cholesterol

It is commonly accepted that “high cholesterol” levels are not healthy. This is due, at least indirectly, to the primary carriers of cholesterol, the LDLs. A primary function of the LDLs is to deliver cholesterol and other lipids directly into cells by receptor mediated endocytosis (Figure 2.237). High levels of LDLs, though, are correlated with formation of atherosclerotic plaques (Figure 2.263 & 2.264) and incidence of atherosclerosis, leading to the description of them as “bad cholesterol.” This is because when LDL levels are very high, plaque formation begins. It is thought that reactive oxygen species (higher in the blood of smokers) causes partial oxidation of fatty acid groups in the LDLs. When levels are high, they tend to accumulate in the extracellular matrix of the epithelial cells on the inside of the arteries. Macrophages of the immune system take up the damaged LDLs (including the cholesterol).

Since macrophages can’t control the amount of cholesterol they take up, cholesterol begins to accumulate in them and they take on appearance that leads to their being described as “foam cells.” With too much cholesterol, the foam cells, however, are doomed to die by the process of programmed cell death (apoptosis). Accumulation of these, along with scar tissue from inflammation result in formation of a plaque. Plaques can grow and block the flow of blood or pieces of them can break loose and plug smaller openings in the blood supply, ultimately leading to heart attack or stroke.

Good cholesterol

On the other hand, high levels of HDL are inversely correlated with atherosclerosis and arterial disease. Depleted HDLs are able to remove cholesterol from foam cells. This occurs as a result of contact between the ApoA-I protein of the HDL and a transport protein on the foam cell (ABC-G1). Another transport protein in the foam cell, ABCA-1 transports extra cholesterol from inside the cell to the plasma membrane where it is taken up into the HDL and returned to the liver or to LDLs by the reverse transport cholesterol pathway.



Figure 2.264 - Actual carotid artery plaque Wikipedia

Deficiency of the ABCA-1 gene leads to Tangier disease. In this condition, HDLs are almost totally absent because they remain empty as a result of not being able to take up cholesterol from foam cells, so they are destroyed by the body.

ApoE and Alzheimer's disease

ApoE is a component of the chylomicrons and is also found in brain, macrophages, kidneys, and the spleen. In humans, it is found in three different alleles, E2, E3, and E4. The E4 allele (present at about 14% of the population) is associated with increased likelihood of contracting Alzheimer's disease. People heterozygous for the allele are 3 times as likely to contract the disease and those homozygous for it are 15 times as likely to do so. It is not known why this gene or allele is linked to the disease. The three alleles differ only slightly in amino acid sequence, but the changes do cause notable structural differences. The E4 allele is associated with increased calcium ion levels and apoptosis after injury. Alzheimer's disease is associated with accumulation of aggregates of the β- amyloid peptide. ApoE does enhance the proteolytic breakdown of it and the E4 isoform is not as efficient in these reactions as the other isoforms.

This page titled 2.8: Structure and Function - Lipids and Membranes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.2: Structure & Function - Amino Acids

Source: BiochemFFA_2_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

All of the proteins on the face of the earth are made up of the same 20 amino acids. Linked together in long chains called polypeptides, amino acids are the building blocks for the vast assortment of proteins found in all living cells.



"*It is one of the more striking generalizations of biochemistry ...that the twenty amino acids and the four bases, are, with minor reservations, the same throughout Nature.*" - Francis Crick

All amino acids have the same basic structure, which is shown in Figure 2.1. At the “center” of each amino acid is a carbon called the α carbon and attached to it are four groups - a hydrogen, an α- carboxyl group, an α-amine group, and an R-group, sometimes referred to as a side chain. The α carbon, carboxyl, and amino groups are common to all amino acids, so the R-group is the only unique feature in each amino acid. (A minor exception to this structure is that of proline, in which the end of the R-group is attached to the α-amine.) With the exception of glycine, which has an R-group consisting of a hydrogen atom, all of the amino acids in proteins have four different groups attached to them and consequently can exist in two mirror image forms, L and D. With only very minor exceptions, every amino acid found in cells and in proteins is in the L configuration.

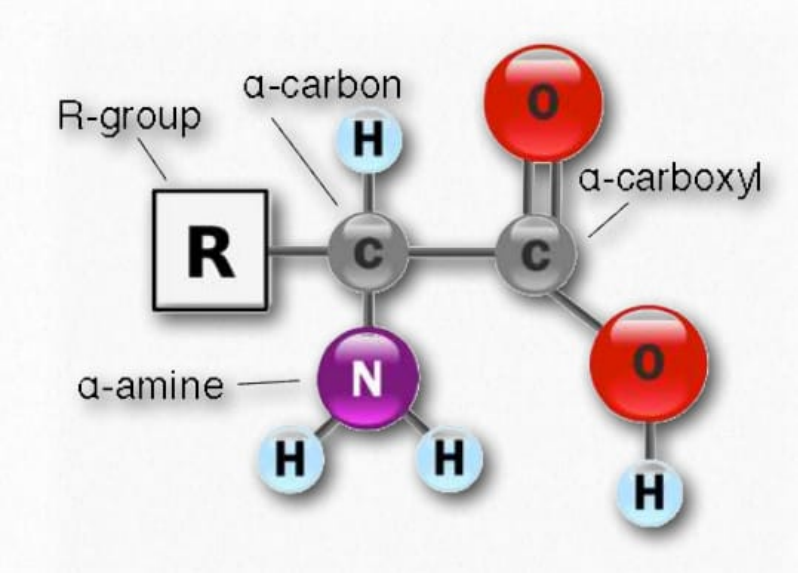


Figure 2.1 - General amino acid structure

There are 22 amino acids that are found in proteins and of these, only 20 are specified by the universal genetic code. The others, selenocysteine and pyrrolysine use tRNAs that are able to base pair with stop codons in the mRNA during translation. When this happens, these unusual amino acids can be incorporated into proteins. Enzymes containing selenocysteine, for example, include glutathione peroxidases, tetraiodothyronine 5' deiodinases, thioredoxin reductases, formate dehydrogenases, glycine reductases, and selenophosphate synthetase. Pyrrolysine-containing proteins are much rarer and are mostly confined to archaea.

Essential and non-essential

Nutritionists divide amino acids into two groups - essential amino acids (must be in the diet because cells can't synthesize them) and non-essential amino acids (can be made by cells). This classification of amino acids has little to do with the structure of amino acids. Essential amino acids vary considerable from one organism to another and even differ in humans, depending on whether they are adults or children. Table 2.1 shows essential and non-essential amino acids in humans.

Some amino acids that are normally nonessential, may need to be obtained from the diet in certain cases. Individuals who do not synthesize sufficient amounts of arginine, cysteine, glutamine, proline, selenocysteine, serine, and tyrosine, due to illness, for example, may need dietary supplements containing these amino acids.

Essential	Non-Essential
Histidine	Alanine
Isoleucine	Arginine
Leucine	Asparagine
Lysine	Aspartic acid
Methionine	Cysteine
Phenylalanine	Glutamic acid
Threonine	Glutamine
Tryptophan	Glycine
Valine	Proline
	Selenocysteine
	Serine
	Tyrosine

Table 2.1 - Essential and non-essential amino acids

Table 2.1 - Essential and non-essential amino acids

Non-protein amino acids

There are also α-amino acids found in cells that are not incorporated into proteins. Common ones include ornithine and citrulline. Both of these compounds are intermediates in the urea cycle. Ornithine is a metabolic precursor of arginine and citrulline can be produced by the breakdown of arginine. The latter reaction produces nitric oxide, an important signaling molecule. Citrulline is the metabolic byproduct. It is sometimes used as a dietary supplement to reduce muscle fatigue.

R-group chemistry

Non-Polar	Carboxyl	Amine	Aromatic	Hydroxyl	Other
Alanine	Aspartic Acid	Arginine	Phenylalanine	Serine	Asparagine
Glycine	Glutamic Acid	Histidine	Tryptophan	Threonine	Cysteine
Isoleucine		Lysine	Tyrosine	Tyrosine	Glutamine
Leucine					Selenocysteine
Methionine					Pyrrolysine
Proline					
Valine					

Table 2.2 - Amino acid categories (based on R-group properties)

We separate the amino acids into categories based on the chemistry of their R-groups. If you compare groupings of amino acids in different textbooks, you will see different names for the categories and (sometimes) the same amino acid being categorized differently by different authors. Indeed, we categorize tyrosine both as an aromatic amino acid and as a hydroxyl amino acid. It is useful to classify amino acids based on their R-groups, because it is these side chains that give each amino acid its characteristic properties. Thus, amino acids with (chemically) similar side groups can be expected to function in similar ways, for example, during protein folding.

Non-polar amino acids

- Alanine (Ala/A) is one of the most abundant amino acids found in proteins, ranking second only to leucine in occurrence. A D-form of the amino acid is also found in bacterial cell walls. Alanine is non-essential, being readily synthesized from pyruvate. It is coded for by GCU, GCC, GCA, and GCG.
- Glycine (Gly/G) is the amino acid with the shortest side chain, having an R-group consistent only of a single hydrogen. As a result, glycine is the only amino acid that is not chiral. Its small side chain allows it to readily fit into both hydrophobic and hydrophilic environments.

Amino acid	Short	Abbrev.	Side chain	Hydro-phobic	pKa	Polar	pH	Small	Tiny	Aromatic or Aliphatic	van der Waals volume
Alanine	A	Ala	$-CH_3$	X	-	-	-	X	X	-	67
Cysteine	C	Cys	$-CH_2SH$	-	8.18	-	acidic	X	X	-	86
Aspartic acid	D	Asp	$-CH_2COOH$	-	3.90	X	acidic	X	-	-	91
Glutamic acid	E	Glu	$-CH_2CH_2COOH$	-	4.07	X	acidic	-	-	-	109
Phenylalanine	F	Phe	$-CH_2C_6H_5$	X	-	-	-	-	-	Aromatic	135
Glycine	G	Gly	-H	X	-	-	-	X	X	-	48
Histidine	H	His	$-CH_2-C_3H_3N_2$	-	6.04	X	weak basic	-	-	Aromatic	118
Isoleucine	I	Ile	$-CH(CH_3)CH_2CH_3$	X	-	-	-	-	-	Aliphatic	124
Lysine	K	Lys	$-(CH_2)_4NH_2$	-	10.54	X	basic	-	-	-	135
Leucine	L	Leu	$-CH_2CH(CH_3)_2$	X	-	-	-	-	-	Aliphatic	124
Methionine	M	Met	$-CH_2CH_2SCH_3$	X	-	-	-	-	-	-	124
Asparagine	N	Asn	$-CH_2CONH_2$	-	-	X	-	X	-	-	96
Pyrrolysine	O	Pyl	$-(CH_2)_4NHCOC_4H_5NCH_3$	-	-	X	weak basic	-	-	-	
Proline	P	Pro	$-CH_2CH_2CH_2-$	X	-	-	-	X	-	-	90
Glutamine	Q	Gln	$-CH_2CH_2CONH_2$	-	-	X	weak basic	-	-	-	114
Arginine	R	Arg	$-(CH_2)_3NH-C(NH)NH_2$	-	12.48	X	strongly basic	-	-	-	148
Serine	S	Ser	$-CH_2OH$	-	5.68	X	weak acidic	X	X	-	73
Threonine	T	Thr	$-CH(OH)CH_3$	-	5.53	X	weak acidic	X	-	-	93
Selenocysteine	U	Sec	$-CH_2SeH$	-	5.73	-	acidic	X	X	-	
Valine	V	Val	$-CH(CH_3)_2$	X	-	-	-	X	-	Aliphatic	105
Tryptophan	W	Trp	$-CH_2C_8H_6N$	-	5.885	X	weak basic	-	-	Aromatic	163
Tyrosine	Y	Tyr	$-CH_2-C_6H_4OH$	-	10.46	X	weak acidic	-	-	Aromatic	141

Figure 2.2 - Amino acid side chain properties Wikipedia

- Glycine is specified in the genetic code by GGU, GGC, GGA, and GGG. It is nonessential to humans.
- Isoleucine (Ile/I) is an essential amino acid encoded by AUU, AUC, and AUA. It has a hydrophobic side chain and is also chiral in its side chain.
- Leucine (Leu/L) is a branched-chain amino acid that is hydrophobic and essential. Leucine is the only dietary amino acid reported to directly stimulate protein synthesis in muscle, but caution is in order, as 1) there are conflicting studies and 2) leucine toxicity is dangerous, resulting in "the four D's": diarrhea, dermatitis, dementia and death . Leucine is encoded by six codons: UUA,UUG, CUU, CUC, CUA, CUG.
- Methionine (Met/M) is an essential amino acid that is one of two sulfurcontaining amino acids - cysteine is the other. Methionine is non-polar and encoded solely by the AUG codon. It is the "initiator" amino acid in protein synthesis, being the first one incorporated into protein chains. In prokaryotic cells, the first methionine in a protein is formylated.

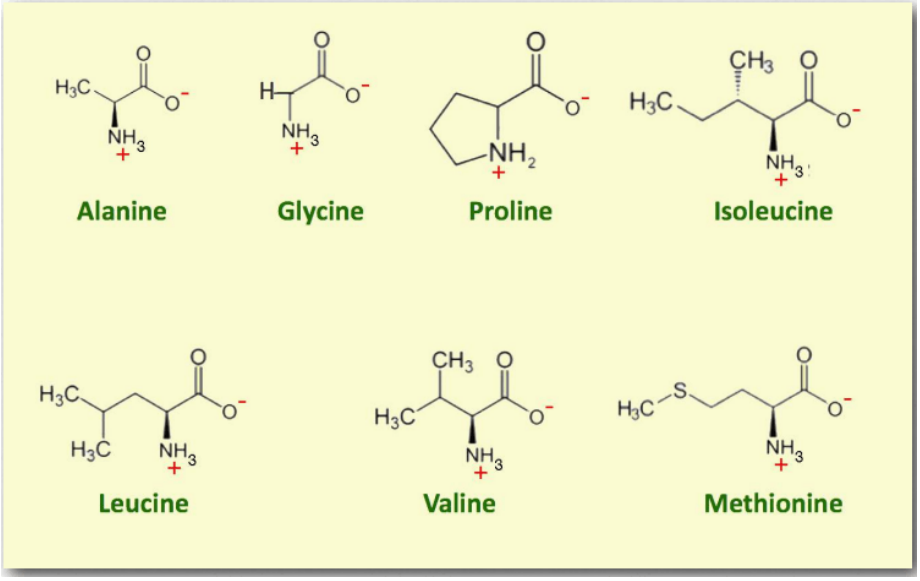


Figure 2.3 - Non-polar amino acids

- Proline (Pro/P) is the only amino acid found in proteins with an R-group that joins with its own α-amino group, making a secondary amine and a ring. Proline is a non-essential amino acid and is coded by CCU, CCC, CCA, and CCG. It is the least flexible of the protein amino acids and thus gives conformational rigidity when present in a protein. Proline's presence in a protein affects its secondary structure. It is a disrupter of α-helices and β-strands. Proline is often hydroxylated in collagen (the reaction requires Vitamin C - ascorbate) and this has the effect of increasing the protein's conformational stability. Proline hydroxylation of hypoxia-inducible factor (HIF) serves as a sensor of oxygen levels and targets HIF for destruction

when oxygen is plentiful.
- Valine (Val/V) is an essential, non-polar amino acid synthesized in plants. It is noteworthy in hemoglobin, for when it replaces glutamic acid at position number six, it causes hemoglobin to aggregate abnormally under low oxygen conditions, resulting in sickle cell disease. Valine is coded in the genetic code by GUU, GUC, GUA, and GUG.

Carboxyl Amino Acids

- Aspartic acid (Asp/D) is a non-essential amino acid with a carboxyl group in its Rgroup. It is readily produced by transamination of oxaloacetate. With a pKa of 3.9, aspartic acid's side chain is negatively charged at physiological pH. Aspartic acid is specified in the genetic code by the codons GAU and GAC.
- Glutamic acid (Glu/E), which is coded by GAA and GAG, is a non-essential amino acid readily made by transamination of α- ketoglutarate. It is a neurotransmitter and has an R-group with a carboxyl group that readily ionizes (pKa = 4.1) at physiological pH.

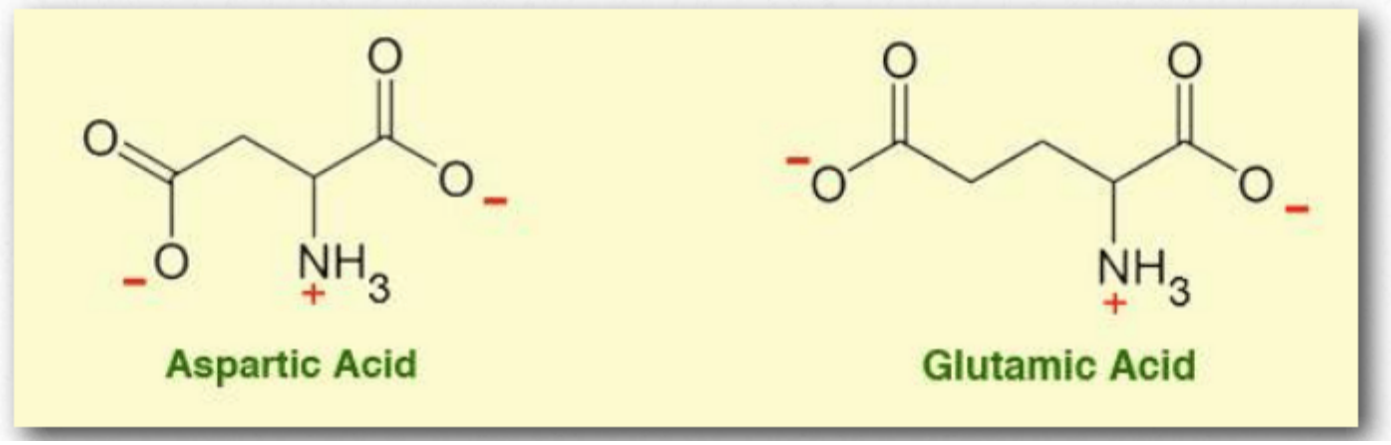


Figure 2.4 - Carboxyl amino acids

Amine amino acids

- Arginine (Arg/R) is an amino acid that is, in some cases, essential, but non-essential in others. Premature infants cannot synthesize arginine. In addition, surgical trauma, sepsis, and burns increase demand for arginine. Most people, however, do not need arginine supplements. Arginine's side chain contains a complex guanidinium group with a pKa of over 12, making it positively charged at cellular pH. It is coded for by six codons - CGU, CGC, CGA, CGG, AGA, and AGG.
- Histidine (His/H) is the only one of the proteinaceous amino acids to contain an imidazole functional group. It is an essential amino acid in humans and other mammals. With a side chain pKa of 6, it can easily have its charge changed by a slight change in pH. Protonation of the ring results in two NH structures which can be drawn as two equally important resonant structures.



Figure 2.5 - Amine amino acids

- Lysine (Lys/K) is an essential amino acid encoded by AAA and AAG. It has an Rgroup that can readily ionize with a charge of +1 at physiological pH and can be posttranslationally modified to form acetyllysine, hydroxylysine, and methyllysine. It can also be ubiquitinated, sumoylated, neddylated, biotinylated, carboxylated, and pupylated, and. O-Glycosylation of hydroxylysine is used to flag proteins for export from the cell. Lysine is often added to animal feed because it is a limiting amino acid and is necessary for optimizing growth of pigs and chickens.

Aromatic amino acids

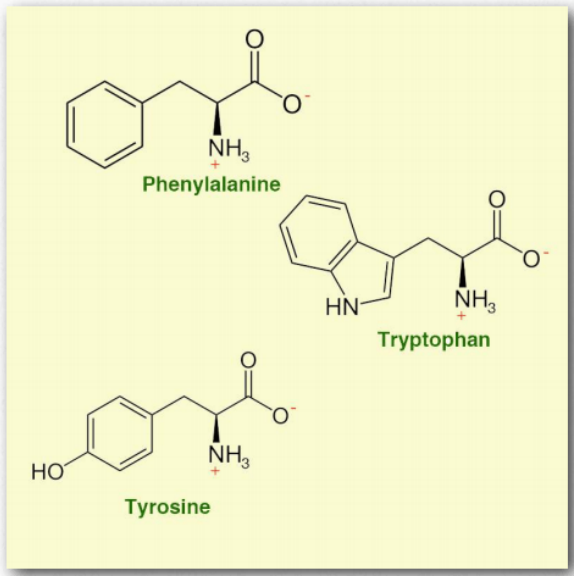


Figure 2.6 - Aromatic amino acids

- Phenylalanine (Phe/ F) is a non-polar, essential amino acid coded by UUU and UUC. It is a metabolic precursor of tyrosine. Inability to metabolize phenylalanine arises from the genetic disorder known as phenylketonuria. Phenylalanine is a component of the aspartame artificial sweetener.
- Tryptophan (Trp/W) is an essential amino acid containing an indole functional group. It is a metabolic precursor of serotonin, niacin, and (in plants) the auxin phytohormone. Though reputed to serve as a sleep aid, there are no clear research results indicating this.
- Tyrosine (Tyr/Y) is a non-essential amino acid coded by UAC and UAU. It is a target for phosphorylation in proteins by tyrosine protein kinases and plays a role in signaling processes. In dopaminergic cells of the brain, tyrosine hydroxylase converts tyrosine to l-dopa, an immediate precursor of dopamine. Dopamine, in turn, is a precursor of norepinephrine and epinephrine. Tyrosine is also a precursor of thyroid hormones and melanin.

Hydroxyl amino acids

- Serine (Ser/S) is one of three amino acids having an R-group with a hydroxyl in it (threonine and tyrosine are the others). It is coded by UCU, UCC, UCA, UGC, AGU, and AGC. Being able to hydrogen bond with water, it is classified as a polar amino acid. It is not essential for humans. Serine is precursor of many important cellular compounds, including purines, pyrimidines, sphingolipids, folate, and of the amino acids glycine, cysteine, and tryptophan. The hydroxyl group of serine in proteins is a target for phosphorylation by certain protein kinases. Serine is also a part of the catalytic triad of serine proteases.

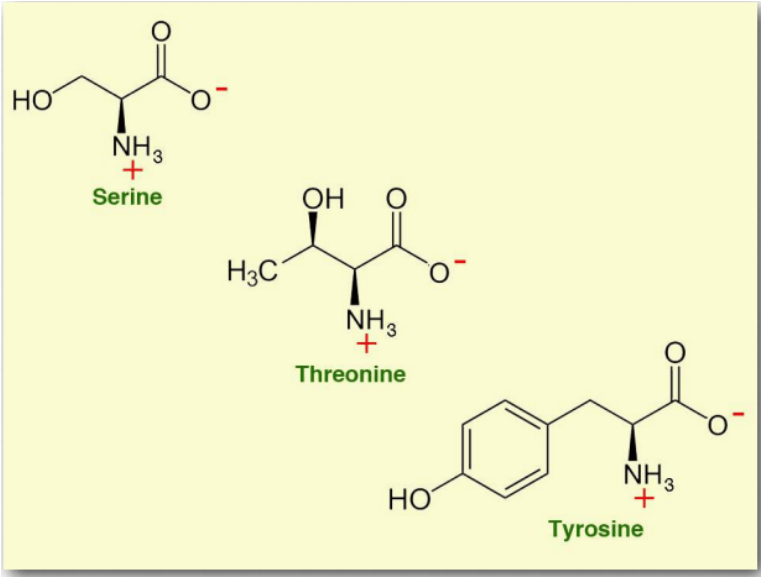


Figure 2.7 - Hydroxyl amino acids

- Threonine (Thr/T) is a polar amino acid that is essential. It is one of three amino acids bearing a hydroxyl group (serine and tyrosine are the others) and, as such, is a target for phosphorylation in proteins. It is also a target for Oglycosylation of proteins. Threonine proteases use the hydroxyl group of the amino acid in their catalysis and it is a precursor in one biosynthetic pathway for making glycine. In some applications, it is used as a pro-drug to increase brain glycine levels. Threonine is encoded in the genetic code by ACU, ACC, ACA, and ACG.

Tyrosine - see HERE.

Amino acid	Short	Abbrev.	Avg. mass (Da)	pI	pK_1 (α-COOH)	pK_2 (α-$^+NH_3$)
Alanine	A	Ala	89.09404	6.01	2.35	9.87
Cysteine	C	Cys	121.15404	5.05	1.92	10.70
Aspartic acid	D	Asp	133.10384	2.85	1.99	9.90
Glutamic acid	E	Glu	147.13074	3.15	2.10	9.47
Phenylalanine	F	Phe	165.19184	5.49	2.20	9.31
Glycine	G	Gly	75.06714	6.06	2.35	9.78
Histidine	H	His	155.15634	7.60	1.80	9.33
Isoleucine	I	Ile	131.17464	6.05	2.32	9.76
Lysine	K	Lys	146.18934	9.60	2.16	9.06
Leucine	L	Leu	131.17464	6.01	2.33	9.74
Methionine	M	Met	149.20784	5.74	2.13	9.28
Asparagine	N	Asn	132.11904	5.41	2.14	8.72
Pyrrolysine	O	Pyl	255.31			
Proline	P	Pro	115.13194	6.30	1.95	10.64
Glutamine	Q	Gln	146.14594	5.65	2.17	9.13
Arginine	R	Arg	174.20274	10.76	1.82	8.99
Serine	S	Ser	105.09344	5.68	2.19	9.21
Threonine	T	Thr	119.12034	5.60	2.09	9.10
Selenocysteine	U	Sec	168.053	5.47		
Valine	V	Val	117.14784	6.00	2.39	9.74
Tryptophan	W	Trp	204.22844	5.89	2.46	9.41
Tyrosine	Y	Tyr	181.19124	5.64	2.20	9.21

Figure 2.8 - Amino acid properties Wikipedia

Other amino acids

- Asparagine (Asn/N) is a non-essential amino acid coded by AAU and AAC. Its carboxyamide in the R-group gives it polarity. Asparagine is implicated in formation of acrylamide in foods cooked at high temperatures (deep frying) when it reacts with carbonyl groups. Asparagine can be made in the body from aspartate by an amidation reaction with an amine from glutamine. Breakdown of asparagine produces malate, which can be oxidized in the citric acid cycle.
- Cysteine (Cys/C) is the only amino acid with a sulfhydryl group in its side chain. It is nonessential for most humans, but may be essential in infants, the elderly and individuals who suffer from certain metabolic diseases. Cysteine's sulfhydryl group

is readily oxidized to a disulfide when reacted with another one. In addition to being found in proteins, cysteine is also a component of the tripeptide, glutathione. Cysteine is specified by the codons UGU and UGC.

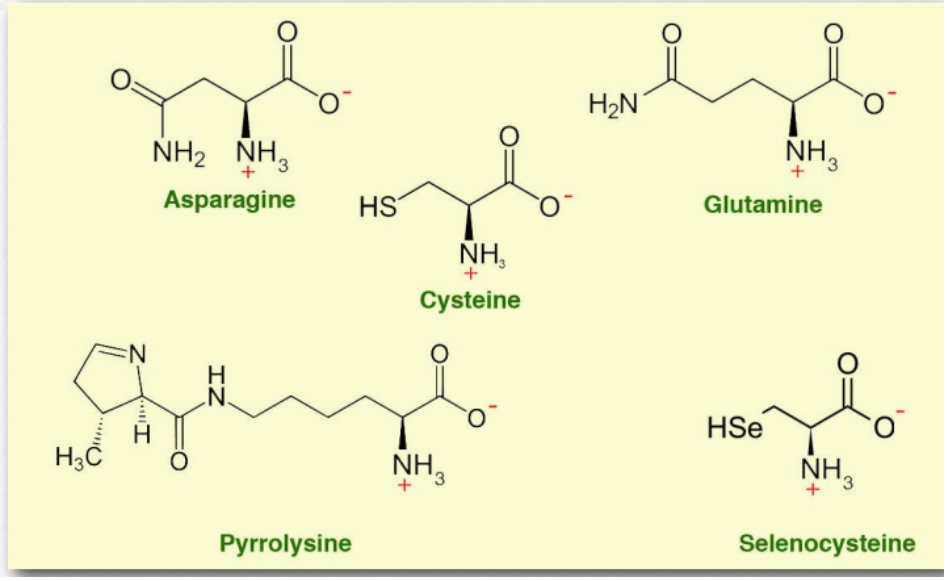


Figure 2.9 - Other amino acids

- Glutamine (Gln/Q) is an amino acid that is not normally essential in humans, but may be in individuals undergoing intensive athletic training or with gastrointestinal disorders. It has a carboxyamide side chain which does not normally ionize under physiological pHs, but which gives polarity to the side chain. Glutamine is coded for by CAA and CAG and is readily made by amidation of glutamate. Glutamine is the most abundant amino acid in circulating blood and is one of only a few amino acids that can cross the blood-brain barrier.
- Selenocysteine (Sec/U) is a component of selenoproteins found in all kingdoms of life. It is a component in several enzymes, including glutathione peroxidases and thioredoxin reductases. Selenocysteine is incorporated into proteins in an unusual scheme involving the stop codon UGA. Cells grown in the absence of selenium terminate protein synthesis at UGAs. However, when selenium is present, certain mRNAs which contain a selenocysteine insertion sequence (SECIS), insert selenocysteine when UGA is encountered. The SECIS element has characteristic nucleotide sequences and secondary structure base-pairing patterns. Twenty five human proteins contain selenocysteine.
- Pyrrolysine (Pyl/O) is a twenty second amino acid, but is rarely found in proteins. Like selenocysteine, it is not coded for in the genetic code and must be incorporated by unusual means. This occurs at UAG stop codons. Pyrrolysine is found in methanogenic archaean organisms and at least one methane-producing bacterium. Pyrrolysine is a component of methane-producing enzymes.

Ionizing groups

pKa values for amino acid side chains are very dependent upon the chemical environment in which they are present. For example, the R-group carboxyl found in aspartic acid has a pKa value of 3.9 when free in solution, but can be as high as 14 when in certain environments inside of proteins, though that is unusual and extreme. Each amino acid has at least one ionizable amine group (α-amine) and one ionizable carboxyl group (α- carboxyl). When these are bound in a peptide bond, they no longer ionize. Some, but not all amino acids have R-groups that can ionize. The charge of a protein then arises from the charges of the α-amine group, the α- carboxyl group. and the sum of the charges of the ionized R-groups. Titration/ionization of aspartic acid is depicted in Figure 2.10. Ionization (or deionization) within a protein's structure can have significant effect on the overall conformation of the protein and, since structure is related to function, a major impact on the activity of a protein.

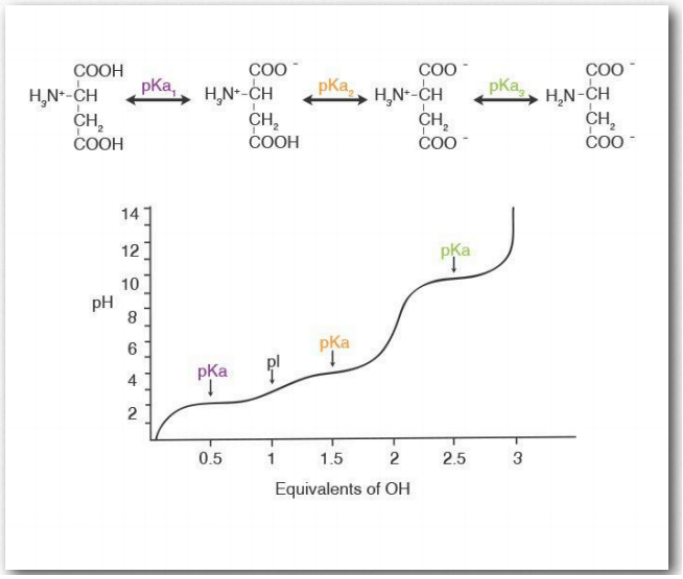


Figure 2.10 - Titration curve for aspartic acid Image by Penelope Irving

Most proteins have relatively narrow ranges of optimal activity that typically correspond to the environments in which they are found (Figure 2.11). It is worth noting that formation of peptide bonds between amino acids removes ionizable hydrogens from both the α- amine and α- carboxyl groups of amino acids. Thus, ionization/ deionization in a protein arises only from 1) the amino terminus; 2) carboxyl terminus; 3) R-groups; or 4) other functional groups (such as sulfates or phosphates) added to amino acids post-translationally - see below.

Carnitine

Not all amino acids in a cell are found in proteins. The most common examples include ornithine (arginine metabolism), citrulline (urea cycle), and carnitine (Figure 2.12). When fatty acids destined for oxidation are moved into the mitochondrion for that purpose, they travel across the inner membrane attached to carnitine. Of the two stereoisomeric forms, the L form is the active one. The molecule is synthesized in the liver from lysine and methionine.

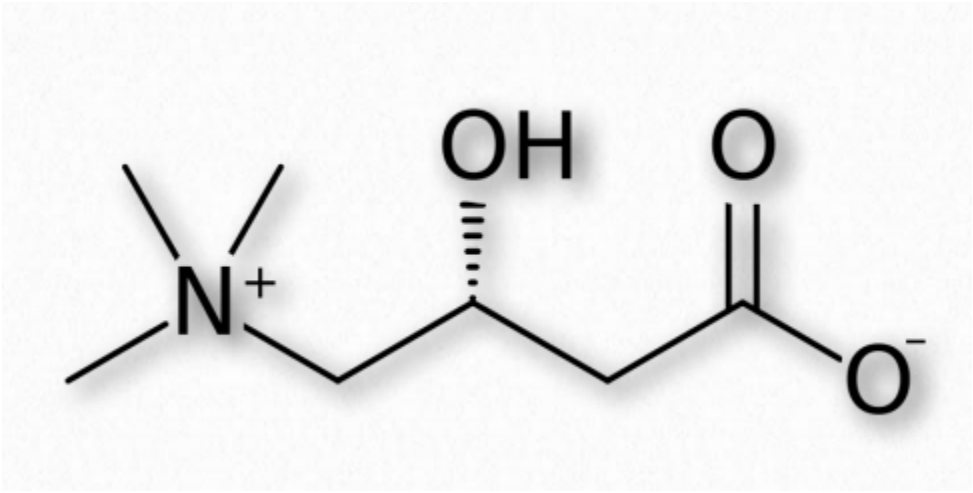


Figure 2.12 - L-Carnitine

From exogenous sources, fatty acids must be activated upon entry into the cytoplasm by being joined to coenzyme A. The CoA portion of the molecule is replaced by carnitine in the intermembrane space of the mitochondrion in a reaction catalyzed by carnitine acyltransferase I. The resulting acylcarnitine molecule is transferred across the inner mitochondrial membrane by the carnitineacylcarnitine translocase and then in the matrix of the mitochondrion, carnitine acyltransferase II replaces the carnitine with coenzyme A (Figure 6.88).

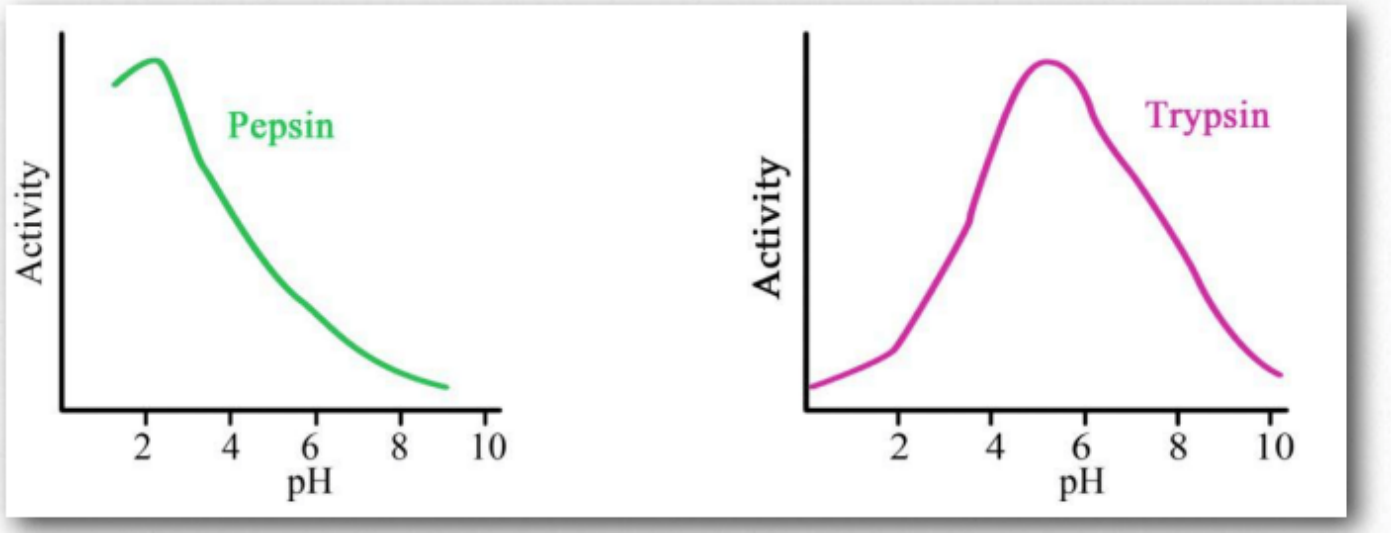


Figure 2.11 - Enzyme activity changes as pH changes Image by Aleia Kim

Catabolism of amino acids

We categorize amino acids as essential or non-essential based on whether or not an organism can synthesize them. All of the amino acids, however, can be broken down by all organisms. They are, in fact, a source of energy for cells, particularly during times of starvation or for people on diets containing very low amounts of carbohydrate. From a perspective of breakdown (catabolism), amino acids are categorized as glucogenic if they produce intermediates that can be made into glucose or ketogenic if their intermediates are made into acetyl-CoA. Figure 2.13 shows the metabolic fates of catabolism of each of the amino acids. Note that some amino acids are both glucogenic and ketogenic.

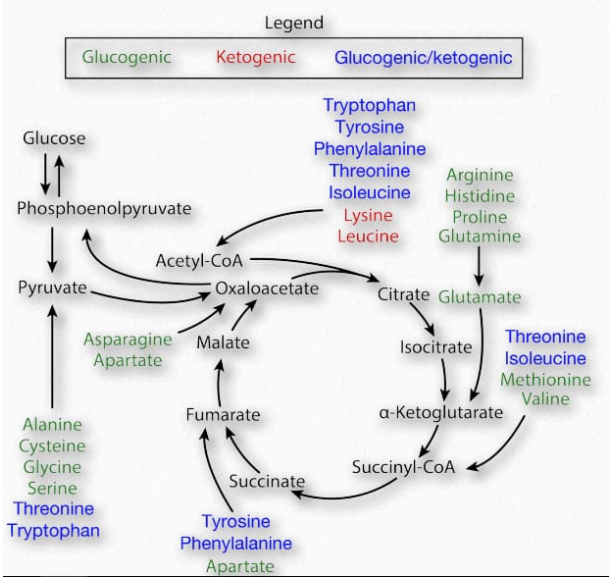


Figure 2.13 - Catabolism of amino acids. Some have more than one path. Image by Pehr Jacobson

Post-translational modifications

After a protein is synthesized, amino acid side chains within it can be chemically modified, giving rise to more diversity of structure and function (Figure 2.14). Common alterations include phosphorylation of hydroxyl groups of serine, threonine, or tyrosine. Lysine, proline, and histidine can have hydroxyls added to amines in their R-groups. Other modifications to amino acids in proteins include addition of fatty acids (myristic acid or palmitic acid), isoprenoid groups, acetyl groups, methyl groups, iodine, carboxyl groups, or sulfates. These can have the effects of ionization (addition of phosphates/sulfates), deionization (addition of acetyl group to the R-group amine of lysine), or have no effect on charge at all. In addition, N-linked- and O-linkedglycoproteins

have carbohydrates covalently attached to side chains of asparagine and threonine or serine, respectively.

Some amino acids are precursors of important compounds in the body. Examples include epinephrine, thyroid hormones, Ldopa, and dopamine (all from tyrosine), serotonin (from tryptophan), and histamine (from histidine).

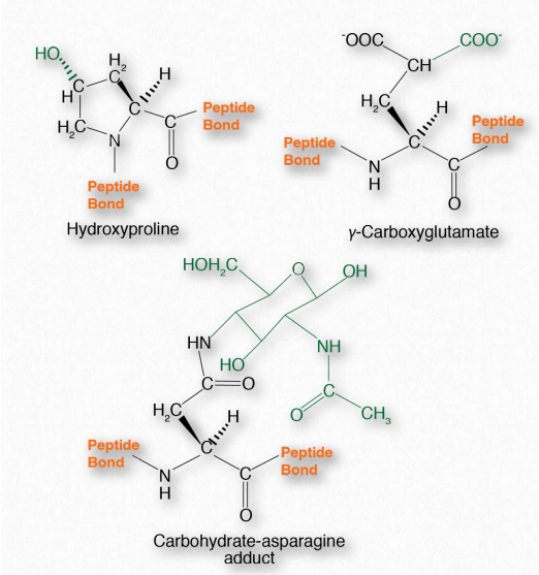


Figure 2.14 - Post-translationally modified amino acids. Modifications shown in green. Image by Penelope Irving

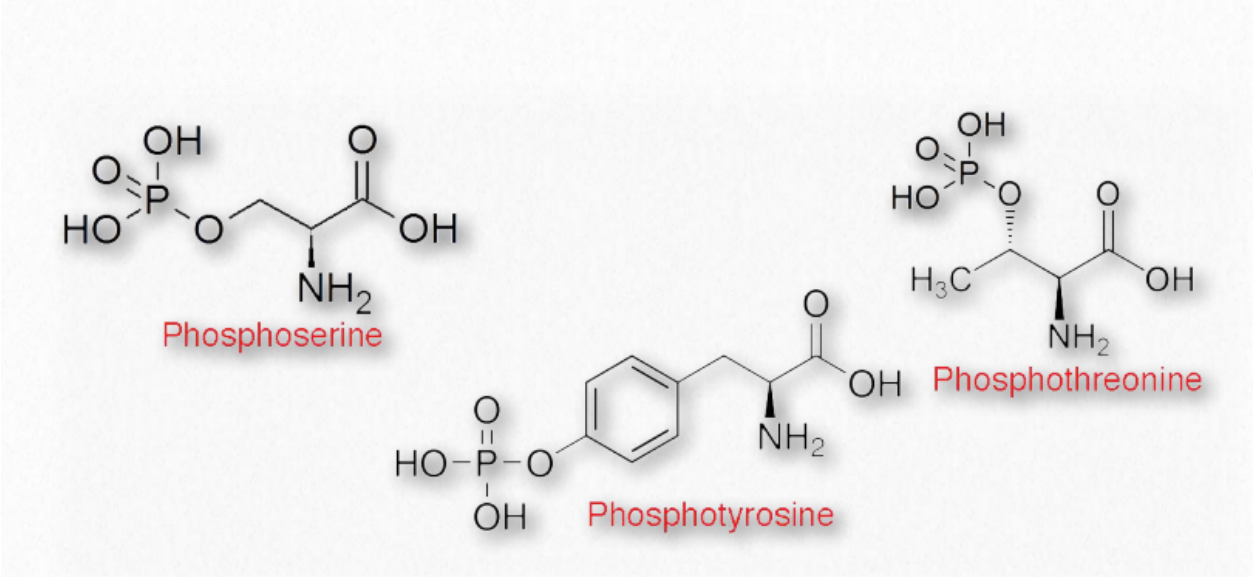


Figure 2.15 - Phosphorylated amino acids

Building Polypeptides

Although amino acids serve other functions in cells, their most important role is as constituents of proteins. Proteins, as we noted earlier, are polymers of amino acids.

Amino acids are linked to each other by peptide bonds, in which the carboxyl group of one amino acid is joined to the amino group of the next, with the loss of a molecule of water. Additional amino acids are added in the same way, by formation of peptide bonds between the free carboxyl on the end of the growing chain and the amino group of the next amino acid in the sequence. A chain made up of just a few amino acids linked together is called an oligopeptide (oligo=few) while a typical protein, which is made up of many amino acids is called a polypeptide (poly=many). The end of the peptide that has a free amino group is called the N-terminus (for NH2), while the end with the free carboxyl is termed the C-terminus (for carboxyl).

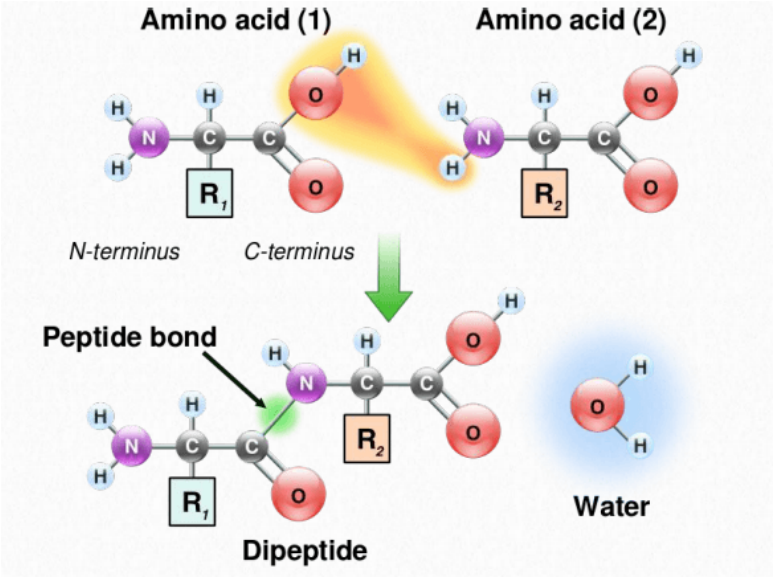


Figure 2.16 Formation of a peptide bond

As we've noted before, function is dependent on structure, and the string of amino acids must fold into a specific 3-D shape, or conformation, in order to make a functional protein. The folding of polypeptides into their functional forms is the topic of the next section.

This page titled 2.2: Structure & Function - Amino Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

2.3: Structure & Function- Proteins I

Source: BiochemFFA_2_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Proteins are the workhorses of the cell. Virtually everything that goes on inside of cells happens as a result of the actions of proteins. Among other things, protein enzymes catalyze the vast majority of cellular reactions, mediate signaling, give structure both to cells and to multicellular organisms, and exert control over the expression of genes. Life, as we know it, would not exist if there were no proteins. The versatility of proteins arises because of their varied structures.

Proteins are made by linking together amino acids, with each protein having a characteristic and unique amino acid sequence. To get a sense for the diversity of proteins that can be made using 20 different amino acids, consider that the number of different combinations possible with 20 amino acids is 20^n, where n=the number of amino acids in the chain. It becomes apparent that even a dipeptide made of just two amino acids joined together gives us 20^2 = 400 different combinations. If we do the calculation for a short peptide of 10 amino acids, we arrive at an enormous 10,240,000,000,000 combinations. Most proteins are much larger than this, making the possible number of proteins with unique amino acid sequences unimaginably huge.

Levels of Structure

The significance of the unique sequence, or order, of amino acids, known as the protein's primary structure, is that it dictates the 3-D conformation the folded protein will have. This conformation, in turn, will determine the function of the protein. We shall examine protein structure at four distinct levels (Figure 2.17) - 1) how sequence of the amino acids in a protein (primary structure) gives identity and characteristics to a protein (Figure 2.18); 2) how local interactions between one part of the polypeptide backbone and another affect protein shape (secondary structure); 3) how the polypeptide chain of a protein can fold to allow amino acids to interact with each other that are not close in primary structure (tertiary structure); and 4) how different polypeptide chains interact with each other within a multi-subunit protein (quaternary structure).

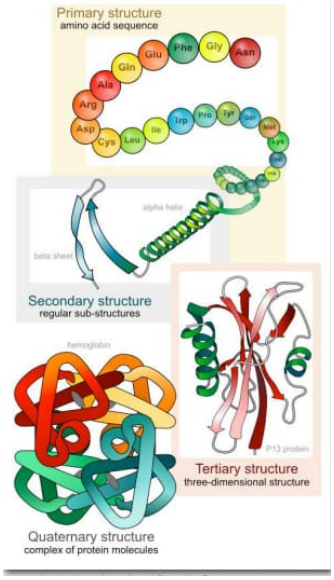


Figure 2.17 - Four Levels of Protein Structure

At this point, we should provide a couple of definitions. We use the term polypeptide to refer to a single polymer of amino acids. It may or may not have folded into its final, functional form. The term protein is sometimes used interchangeably with polypeptide, as in "protein synthesis". It is generally used, however, to refer to a folded, functional molecule that may have one or more subunits (made up of individual polypeptides). Thus, when we use the term protein, we are usually referring to a functional, folded polypeptide or peptides. Structure is essential for function. If you alter the structure, you alter the function - usually, but not always, this means you lose all function. For many proteins, it is not difficult to alter the structure.

Proteins are flexible, not rigidly fixed in structure. As we shall see, it is the flexibility of proteins that allows them to be amazing catalysts and allows them to adapt to, respond to, and pass on signals upon binding of other molecules or proteins. However, proteins are not infinitely flexible. There are constraints on the conformations that proteins can adopt and these constraints govern the conformations that proteins display.

Subtle changes

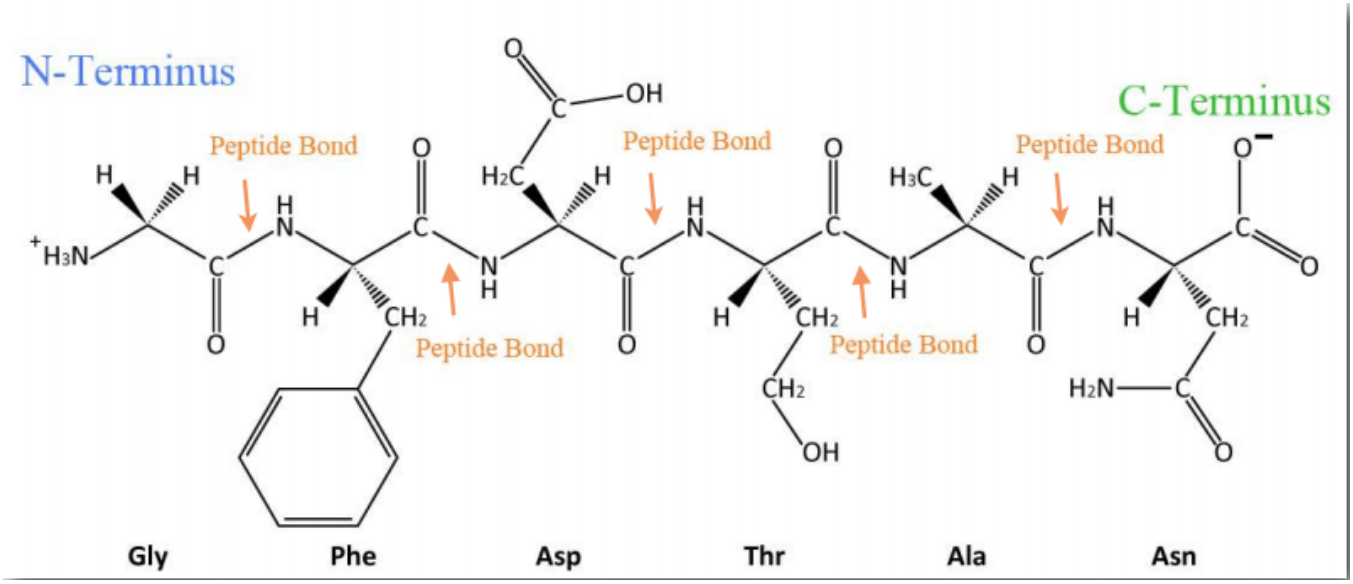


Figure 2.18 - Sequence of a simple polypeptide Wikipedia

Even very tiny, subtle changes in protein structure can give rise to big changes in the behavior of proteins. Hemoglobin, for example, undergoes an incredibly small structural change upon binding of one oxygen molecule, and that simple change causes the remainder of the protein to gain a considerably greater affinity for oxygen that the protein didn't have before the structural change.

Sequence, structure and function

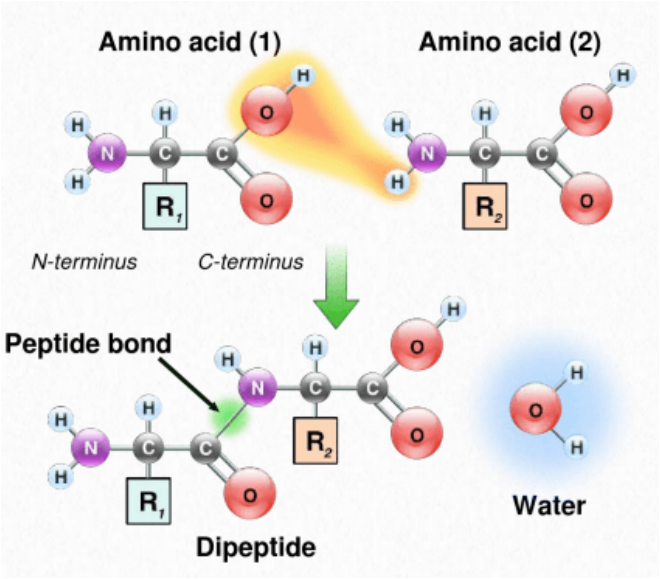


Figure 2.19 Linking of amino acids through peptide bond formation

As discussed earlier, the number of different amino acid sequences possible, even for short peptides, is very large. No two proteins with different amino acid sequences (primary structure) have identical overall structure. The unique amino acid sequence of a protein is reflected in its unique folded structure. This structure, in turn, determines the protein's function. This is why

mutations that alter amino acid sequence can affect the function of a protein.

Protein Synthesis

Synthesis of proteins occurs in the ribosomes and proceeds by joining the carboxyl terminus of the first amino acid to the amino terminus of the next one (Figure 2.19). The end of the protein that has the free α-amino group is referred to as the amino terminus or N-terminus. The other end is called the carboxyl terminus or C-terminus , since it contains the only free α-carboxyl group. All of the other α-amino groups and α-carboxyl groups are tied up in forming peptide Figure 2.19 Linking of amino acids through peptide bond formation bonds that join adjacent amino acids together. Proteins are synthesized starting with the amino terminus and ending at the carboxyl terminus.

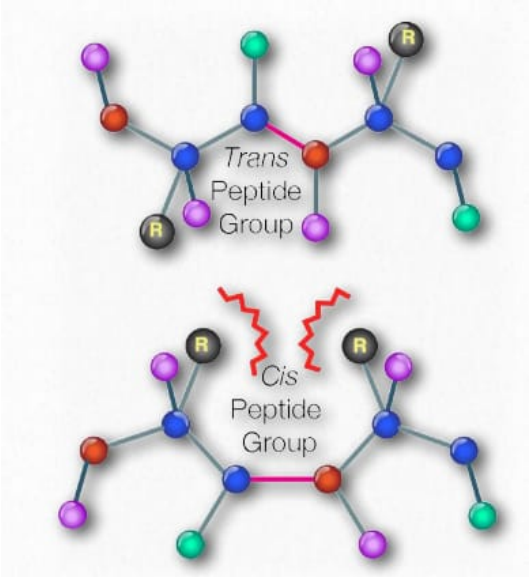


Figure 2.20 - Cis vs trans orientation of R-groups around peptide bond Image by Aleia Kim

Schematically, in Figure 2.18, we can see how sequential R-groups of a protein are arranged in an alternating orientation on either side of the polypeptide chain. Organization of R-groups in this fashion is not random. Steric hindrance can occur when consecutive R-groups are oriented on the same side of a peptide backbone (Figure 2.20)

Primary Structure

Primary structure is the ultimate determinant of the overall conformation of a protein. The primary structure of any protein arrived at its current state as a result of mutation and selection over evolutionary time. Primary structure of proteins is mandated by the sequence of DNA coding for it in the genome. Regions of DNA specifying proteins are known as coding regions (or genes).

The base sequences of these regions directly specify the sequence of amino acids in proteins, with a one-to-one correspondence between the codons (groups of three consecutive bases) in the DNA and the amino acids in the encoded protein. The sequence of codons in DNA, copied into messenger RNA, specifies a sequence of amino acids in a protein. (Figure 2.21).

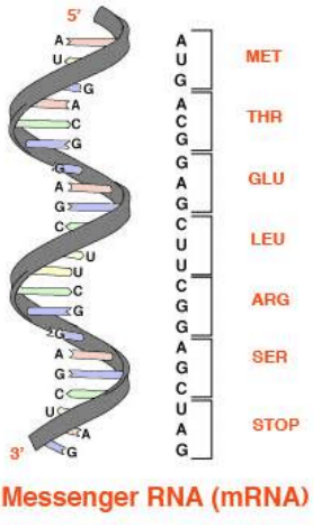


Figure 2.21 - From RNA to amino acids - the genetic code Wikipedia

The order in which the amino acids are joined together in protein synthesis starts defining a set of interactions between amino acids even as the synthesis is occurring. That is, a polypeptide can fold even as it is being made. The order of the R-group structures and resulting interactions are very important because early interactions affect later interactions. This is because interactions start establishing structures - secondary and tertiary. If a helical structure (secondary structure), for example, starts to form, the possibilities for interaction of a particular amino acid Rgroup may be different than if the helix had not formed (Figure 2.22). R-group interactions can also cause bends in a polypeptide sequence (tertiary structure) and these bends can create (in some cases) opportunities for interactions that wouldn't have been possible without the bend or prevent (in other cases) similar interaction possibilities.

Secondary Structure

As protein synthesis progresses, interactions between amino acids close to each other begin to occur, giving rise to local patterns called secondary structure. These secondary structures include the well known α- helix and β-strands. Both were predicted by Linus Pauling, Robert Corey, and Herman Branson in 1951. Each structure has unique features.

α-helix

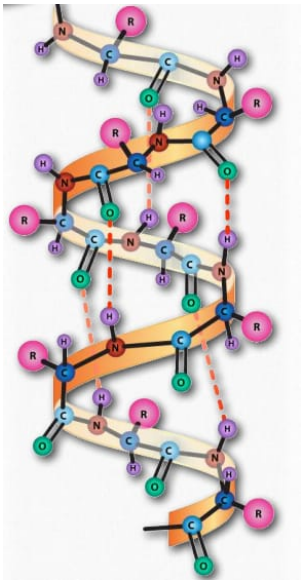


Figure 2.22 - The α-helix. Hydrogen bonds (dotted lines) between the carbonyl oxygen and the amine hydrogen stabilize the structure. Image by Aleia Kim

The α-helix has a coiled structure, with 3.6 amino acids per turn of the helix (5 helical turns = 18 amino acids). Helices are predominantly right handed - only in rare cases, such as in sequences with many glycines can left handed α- helices form. In the α-helix, hydrogen bonds form between C=O groups and N-H groups in the polypeptide backbone that are four amino acids distant. These hydrogen bonds are the primary forces stabilizing the α-helix.

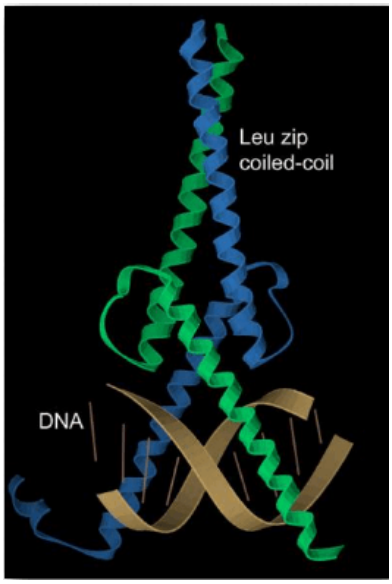


Figure 2.23 - α-helices in a protein with a leucine zipper structural domain. The α-helices are shown in blue and green and are bound to a DNA double helix in brown.

We use the terms rise, repeat, and pitch to describe the parameters of any helix. The repeat is the number of residues in a helix before it begins to repeat itself. For an α-helix, the repeat is 3.6 amino acids per turn of the helix. The rise is the distance the helix elevates with addition of each residue. For an α-helix, this is 0.15 nm per amino acid. The pitch is the distance between complete turns of the helix. For an α-helix, this is 0.54 nm. The stability of an α-helix is enhanced by the presence of the amino acid aspartate.



Figure 2.24 - α-helix sculpture outside Linus Pauling's boyhood home Wikipedia

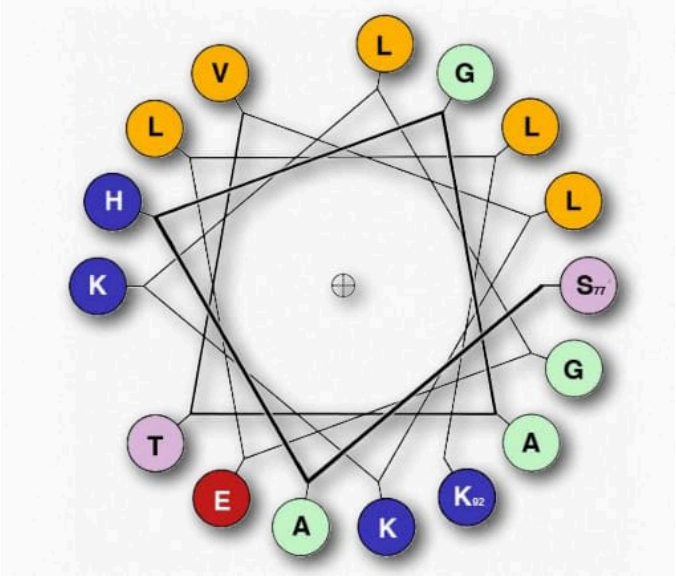


Figure 2.25 - Helical Wheel Representation of an α-Helix. The one letter genetic code is used. The helix starts at Serine #77 at the right and ends at lysine #92 in the lower right. Hydrophobic amino acids are shown in yellow and ionizing amino acids are shown in blue. Hydrophobic amino acids tend to interact with each other and not with ionizing amino acids. Wikipedia

β strand/sheet

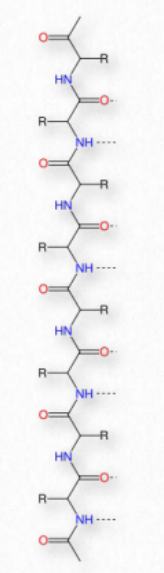


Figure 2.26 - β strand

A helix is, of course, a three-dimensional object. A flattened form of helix in two dimensions is a common description for a β-strand. Rather than coils, β-strands have bends and these are sometimes referred to as pleats, like the pleats in a curtain. β-strands can be organized to form elaborately organized structures, such as sheets, barrels, and other arrangements.

Higher order β-strand structures are sometimes called supersecondary structures), since they involve interactions between amino acids not close in primary sequence. These structures, too, are stabilized by hydrogen bonds between carbonyl oxygen atoms and hydrogens of amine groups in the polypeptide backbone (Figure 2.28). In a higher order structure, strands can be arranged parallel (amino to carboxyl orientations the same) or anti-parallel (amino to carboxyl orientations opposite of each other (in Figure 2.27, the direction of the strand is shown by the arrowhead in the ribbon diagrams).

Turns

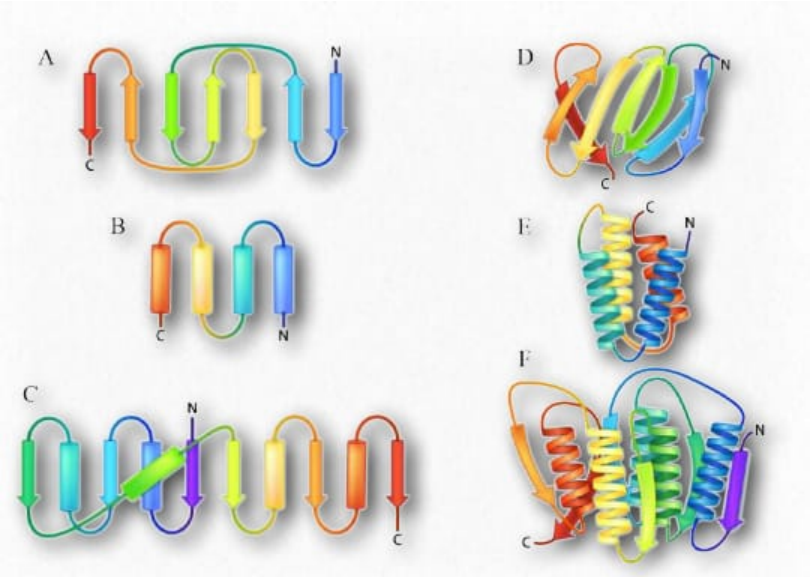


Figure 2.27 - Ribbon depictions of supersecondary β-sheets (A-D) and α-helix arrangements (E-F) Image by Aleia Kim

Turns (sometimes called reverse turns) are a type of secondary structure that, as the name suggests, causes a turn in the structure of a polypeptide chain. Turns give rise to tertiary structure ultimately, causing interruptions in the secondary structures (α- helices and β-strands) and often serve as connecting regions between two regions of secondary structure in a protein. Proline and glycine play common roles in turns, providing less flexibility (starting the turn) and greater flexibility (facilitating the turn), respectively.

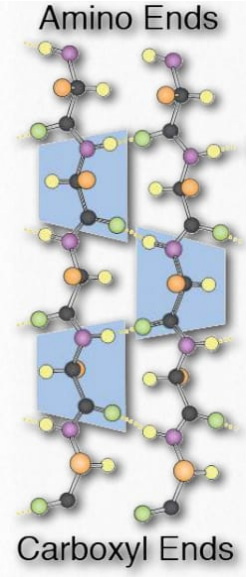


Figure 2.28 - Components of a β-sheet in a parallel arrangement. H-bonds in yellow. Image by Aleia Kim

There are at least five types of turns, with numerous variations of each giving rise to many different turns. The five types of turns are

• δ-turns - end amino acids are separated by one peptide bond

• γ-turns - separation by two peptide bonds

•β-turns - separation by three peptide bonds

•α-turns - separation by four peptide bonds

•π-turns - separation by five bonds

Of these, the β-turns are the most common form and the δ-turns are theoretical, but unlikely, due to steric limitations. Figure 2.29 depicts a β- turn.

3_{10} helices

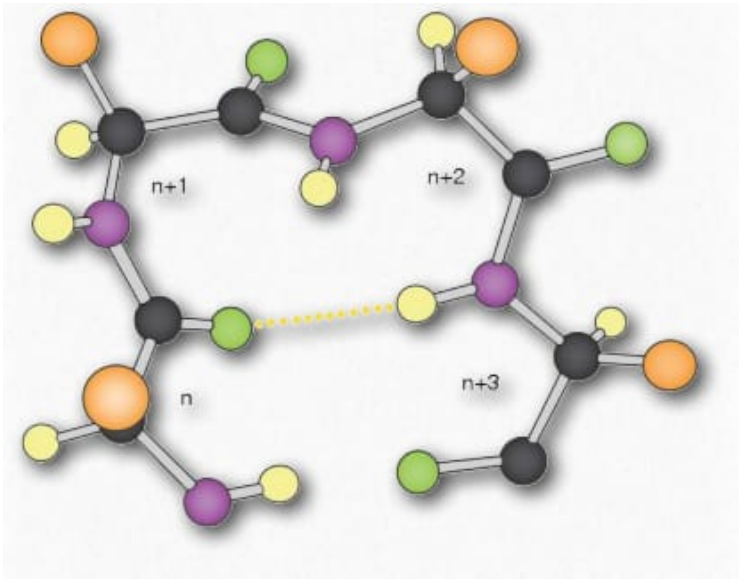


Figure 2.29 - β-turn. R-groups are shown in orange, hydrogens in yellow, carbons in charcoal, nitrogens in purple, and oxygens in green. A stabilizing hydrogen bond is indicated with the dotted line. Image by Aleia Kim

In addition to the α-helix, β-strands, and various turns, other regular, repeating structures are seen in proteins, but occur much less commonly. The 3_{10} helix is the fourth most abundant secondary structure in proteins, constituting about 10-15% of all helices. The helix derives its name from the fact that it contains 10 amino acids in 3 turns. It is right-handed. Hydrogen bonds form between amino acids that are three residues apart. Most commonly, the 3_{10} helix appears at the amine or carboxyl end of an α-helix. Like the α-helix, the 3_{10} helix is stabilized by the presence of aspartate in its sequence.

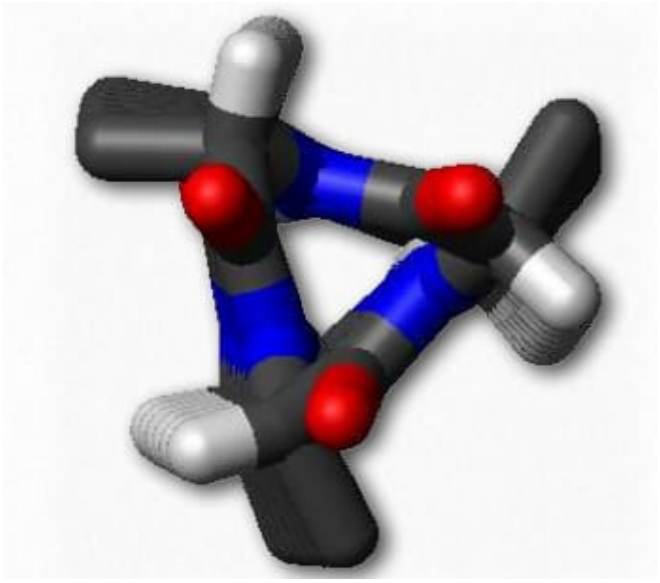


Figure 2.30 - Top view of a 310 Helix. Carbonyl groups are in red and pointed upwards. Note the almost perfect 3-fold symmetry Wikipedia

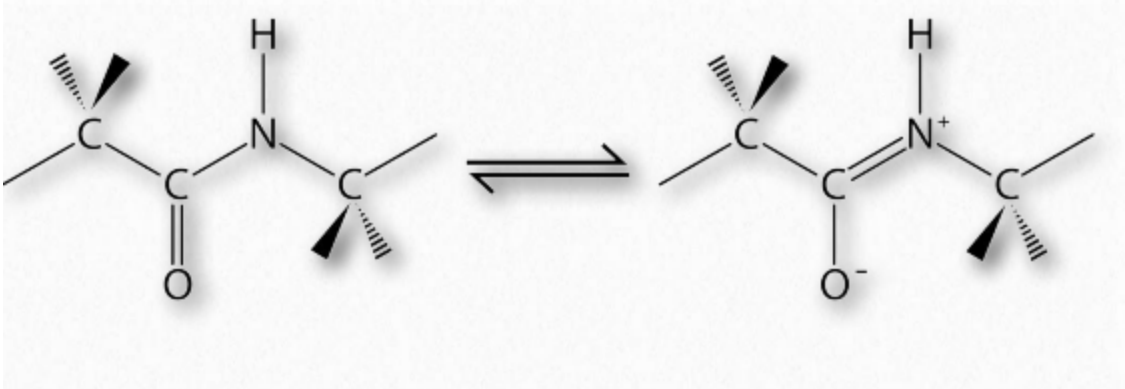


Figure 2.31 - Resonance of the peptide bond Wikipedia

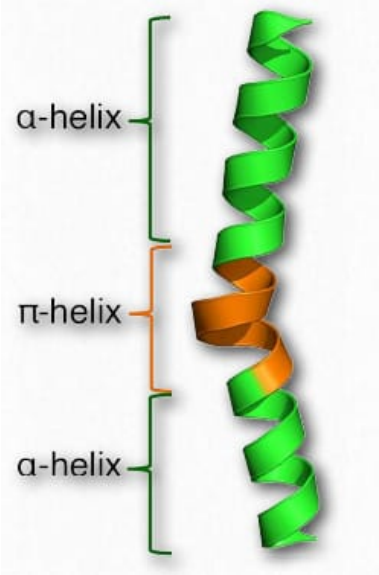


Figure 2.32 - π helix Wikipedia

π-helices

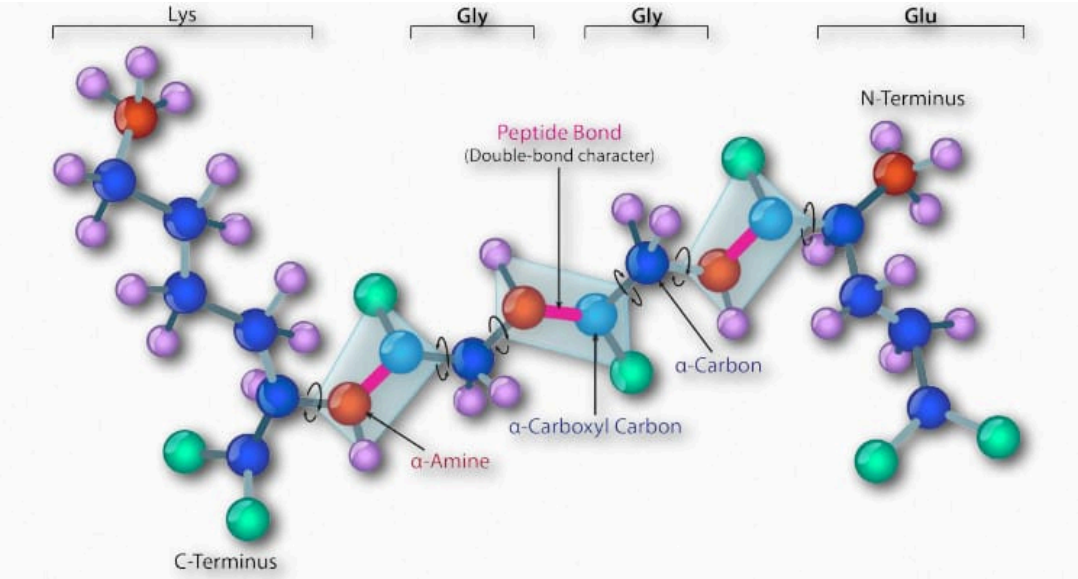


Figure 2.33 - Planes (light blue) defined by the double-bonded character of the peptide bond Image by Aleia Kim

A π-helix may be thought of as a special type of α- helix. Some sources describe it as an α-helix with an extra amino acid stuck in the middle of it (Figure 2.32). π-helices are not exactly rare, occurring at least once in as many as 15% of all proteins. Like the α-helix, the π-helix is right-handed, but where the α-helix has 18 amino acids in 5 turns, the π-helix has 22 amino acids in 5 turns. π-helices typically do not stretch for very long distances. Most are only about 7 amino acids long and the sequence almost always occurs in the middle of an α-helical region.

Ramachandran plots

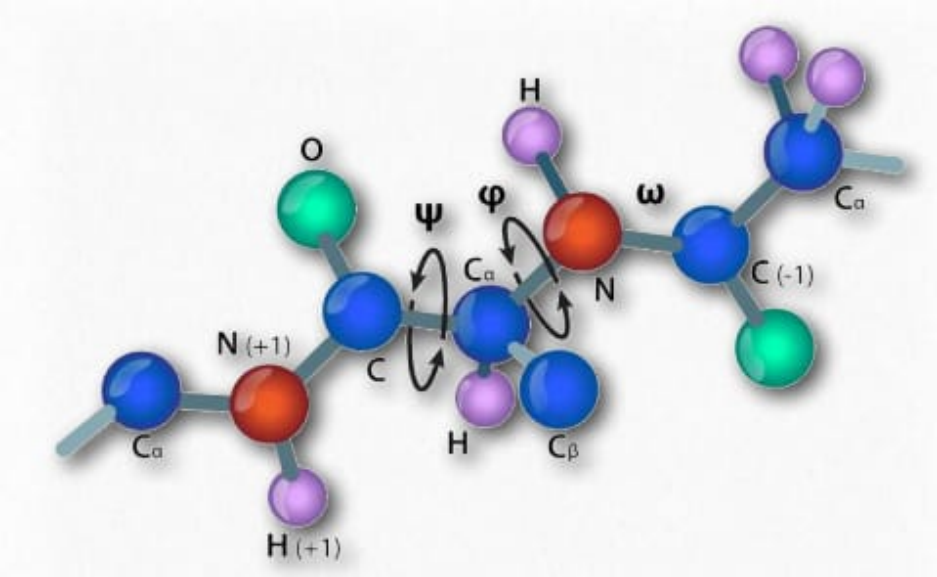


Figure 2.34 - ω, ψ, and φ rotational angles in a peptide Image by Aleia Kim

In 1963, G.N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan described a novel way to describe protein structure. If one considers the backbone of a polypeptide chain, it consists of a repeating set of three bonds. Sequentially (in the amino to carboxyl direction) they are 1) a rotatable bond (ψ) between α-carbon and α-carboxyl preceding the peptide bond (see HERE), 2) a non-rotatable peptide bond (ω) between the α-carboxyl and α-amine groups), and 3) a rotatable bond (φ) between the α-amine and α-carbon following the peptide bond (see HERE). Note in Figures 2.33 and 2.34 that the amino to carboxyl direction is right to left.

The presence of the carbonyl oxygen on the α-carboxyl group allows the peptide bond to exist as a resonant structure, meaning that it behaves some of the time as a double bond. Double bonds cannot, of course, rotate, but the bonds on either side of it have some freedom of rotation. The φ and ψ angles are restricted to certain values, because some angles will result in steric hindrance. In addition, each type of secondary structure has a characteristic range of values for φ and ψ.

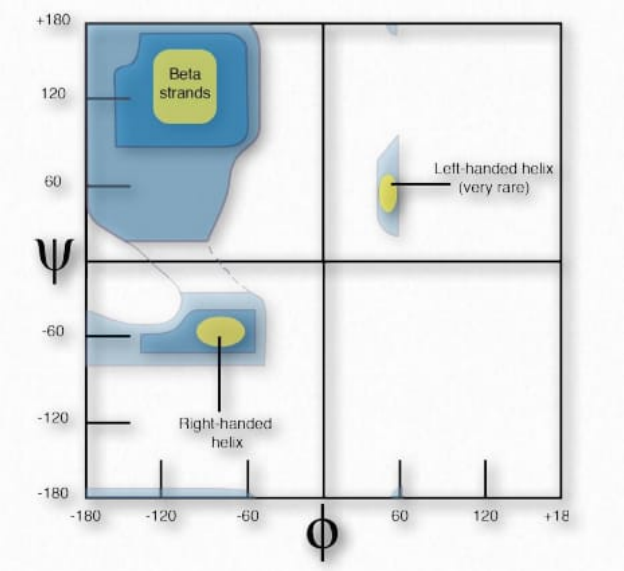


Figure 2.35 - Theoretical Ramachandran plot Image by Penelope Irving

Ramachandran and colleagues made theoretical calculations of the energetic stability of all possible angles from 0° to 360° for each of the φ and ψ angles and plotted the results on a Ramachandran Plot (also called a φ-ψ plot), delineating regions of angles that were theoretically the most stable (Figure 2.35).

Three primary regions of stability were identified, corresponding to φ-ψ angles of β-strands (top left), right handed α- helices (bottom left), and lefthanded α-helices (upper right). The plots of predicted stability are remarkably accurate when compared to φ-ψ angles of actual proteins.

Secondary structure prediction

Amino Acid	α helix	Reverse turn	β sheet
Ala	1.41	0.82	0.72
Arg	1.21	0.90	0.84
Asn	0.76	1.34	0.48
Asp	0.99	1.24	0.39
Cys	0.66	0.54	1.40
Gln	1.27	0.84	0.98
Glu	1.59	1.01	0.52
Gly	0.43	1.77	0.58
His	1.05	0.81	0.80
Ile	1.09	0.47	1.67
Leu	1.34	0.57	1.22
Lys	1.23	1.07	0.69
Met	1.30	0.52	1.14
Phe	1.16	0.59	1.33
Pro	0.34	1.32	0.31
Ser	0.57	1.22	0.96
Thr	0.76	0.96	1.17
Trp	1.02	0.65	1.35
Tyr	0.74	0.76	1.45
Val	0.90	0.41	1.87

Table 2.3 - Relative tendencies of each amino acid to be in a secondary structure. Higher values indicate greater tendency Image by Penelope Irving

By comparing primary structure (amino acid sequences) to known 3D protein structures, one can tally each time an amino acid is found in an α-helix, β-strand/sheet, or a turn. Computer analysis of thousands of these sequences allows one to assign a likelihood of any given amino acid appearing in each of these structures. Using these tendencies, one can, with up to 80% accuracy, predict regions of secondary structure in a protein based solely on amino acid sequence.

This is seen in Table 2.3. Occurrence in primary sequence of three consecutive amino acids with relative tendencies higher than one is an indicator that that region of the polypeptide is in the corresponding secondary structure. An online resource for predicting secondary structures called PSIPRED is available HERE.

Hydrophobicity

Amino Acid Hydropathy Scores

Amino Acid	One Letter Code	Hydropathy Score
Isoleucine	I	4.5
Valine	V	4.2
Leucine	L	3.8
Phenylalanine	F	2.8
Cysteine	C	2.5
Methionine	M	1.9
Alanine	A	1.8
Glycine	G	-0.4
Threonine	T	-0.7
Tryptophan	W	-0.9
Serine	S	-0.8
Tyrosine	Y	-1.3
Proline	P	-1.6
Histidine	H	-3.2
Glutamic acid	E	-3.5
Glutamine	Q	-3.5
Aspartic acid	D	-3.5
Asparagine	N	-3.5
Lysine	K	-3.9
Arginine	R	-4.5

Table 2.4 - Hydropathy Scores

The chemistry of amino acid Rgroups affects the structures they are most commonly found in. Subsets of their chemical properties can give clues to structure and, sometimes, cellular location. A prime example is the hydrophobicity (wateravoiding tendencies) of some Rgroups. Given the aqueous environment of the cell, such R-groups are not likely to be on the outside surface of a folded protein.

However, this rule does not hold for regions of protein that may be embedded within the lipid bilayers of cellular/ organelle membranes. This is because the region of such proteins that form the transmembrane domains are are buried in the hydrophobic environment in the middle of the lipid bilayer.

Not surprisingly, scanning primary sequences for specifically sized/spaced stretches of hydrophobic amino acids can help to identify proteins found in membranes. Table 2.4 shows hydrophobicity values for R-groups of the amino acids. In this set, the scale runs from positive values (hydrophobic) to negative values (hydrophilic). A KyteDoolittle Hydropathy plot for the RET protooncogene membrane protein is shown in Figure 2.36. Two regions of the protein are very hydrophobic as can be seen from the peaks near amino acids 5-10 and 630-640. Such regions might be reasonably expected to be situated either within the interior of the folded protein or to be part of transmembrane domains.

Random coils

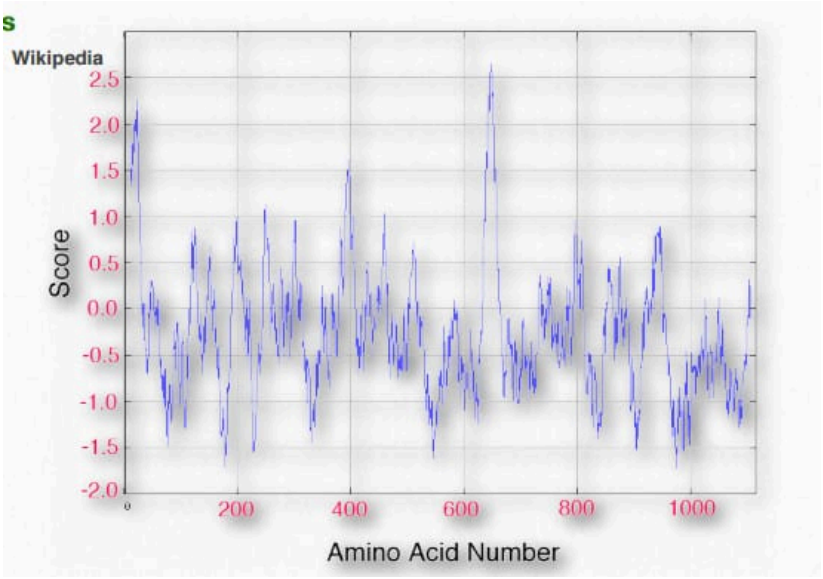


Figure 2.36 Kyte-Doolittle hydropathy plot for the RET protooncogene Wikipedia

Some sections of a protein assume no regular, discernible structure and are sometimes said to lack secondary structure, though they may have hydrogen bonds. Such segments are described as being in random coils and may have fluidity to their structure that results in them having multiple stable forms. Random coils are identifiable with spectroscopic methods, such as circular dichroism Wikipedia and nuclear magnetic resonance (NMR) in which distinctive signals are observed. See also metamorphic proteins (HERE) and intrinsically disordered proteins (HERE).

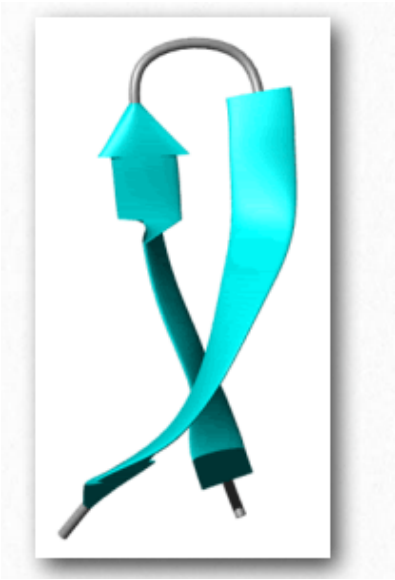


Figure 2.37 - Ribbon depiction of a β-hairpin. Shown are two β strands in turquoise interacting with each other.

Supersecondary structure

Another element of protein structure is harder to categorize because it incorporates elements of secondary and tertiary structure. Dubbed supersecondary structure (or structural motifs), these structures contain multiple nearby secondary structure components arranged in a specific way and that appear in multiple proteins. Since there are many ways of making secondary structures from different primary structures, so too can similar motifs arise from different primary sequences. An example of a structural motif is shown in Figure 2.37.

Tertiary structure

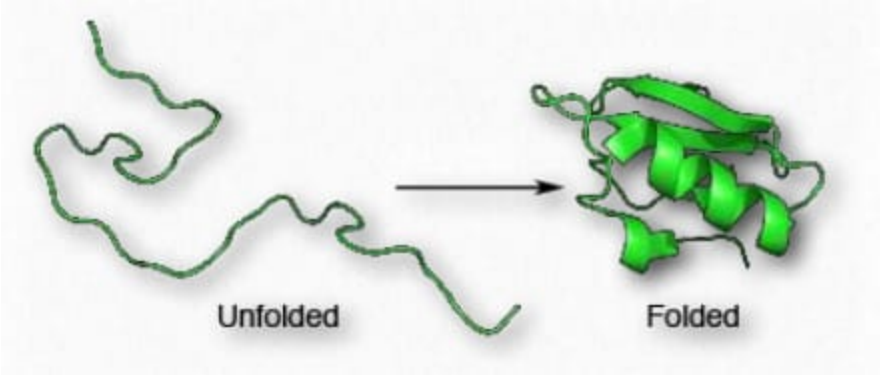


Figure 2.38 - Folding of a polypeptide chain

Proteins are distinguished from each other by the sequence of amino acids comprising them. The sequence of amino acids of a protein determines protein shape, since the chemical properties of each amino acid are forces that give rise to intermolecular interactions to begin to create secondary structures, such as α-helices and β-strands. The sequence also defines turns and random coils that play important roles in the process of protein folding.

Since shape is essential for protein function, the sequence of amino acids gives rise to all of the properties a protein has. As protein synthesis proceeds, individual components of secondary structure start to interact with each other, giving rise to folds that bring amino acids close together that are not near each other in primary structure (Figure 2.38). At the tertiary level of structure, interactions among the R-groups of the amino acids in the protein, as well as between the polypeptide backbone and amino acid side groups play a role in folding.

Globular proteins

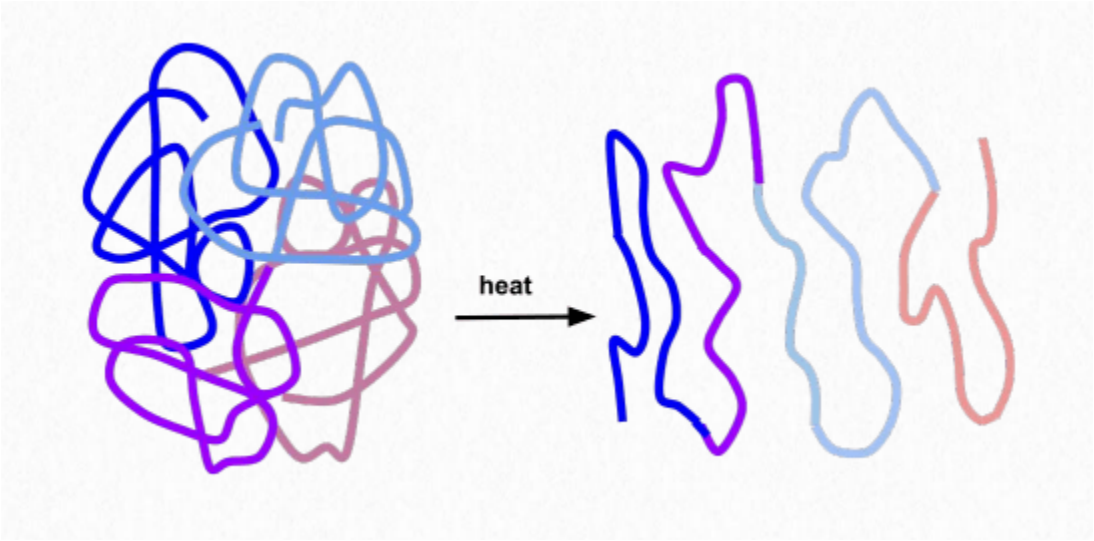


Figure 2.39 - Unfolding (denaturation) of a protein Wikipedia

Folding gives rise to distinct 3-D shapes in proteins that are non-fibrous. These proteins are called globular. A globular protein is stabilized by the same forces that drive its formation. These include ionic interactions, hydrogen bonding, hydrophobic forces, ionic bonds, disulfide bonds and metallic bonds. Treatments such as heat, pH changes, detergents, urea and mercaptoethanol overpower the stabilizing forces and cause a protein to unfold, losing its structure and (usually) its function (Figure 2.39). The ability of heat and detergents to denature proteins is why we cook our food and wash our hands before eating - such treatments denature the proteins in the microorganisms on our hands. Organisms that live in environments of high temperature (over 50°C) have proteins with changes in stabilizing forces - additional hydrogen bonds, additional salt bridges (ionic interactions), and compactness may all play roles in keeping these proteins from unfolding.

Protein stabilizing forces

Before considering the folding process, let us consider some of the forces that help to stabilize proteins.

Hydrogen bonds

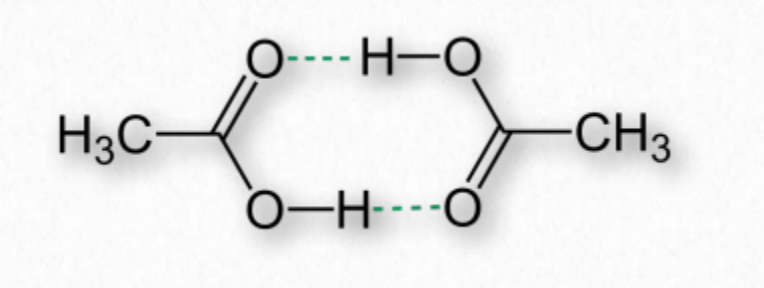


Figure 2.40 - Hydrogen bonds (dotted lines) between two molecules of acetic acid

Hydrogen bonds arise as a result of partially charged hydrogens found in covalent bonds. This occurs when the atom the hydrogen is bonded to has a greater electronegativity than hydrogen itself does, resulting in hydrogen having a partial positive charge because it is not able to hold electrons close to itself (Figure 2.40).

Hydrogen partially charged in this way is attracted to atoms, such as oxygen and nitrogen that have partial negative charges, due to having greater electronegativities and thus holding electrons closer to themselves. The partially positively charged hydrogens are called donors, whereas the partially negative atoms they are attracted to are called acceptors. (See Figure 1.30).

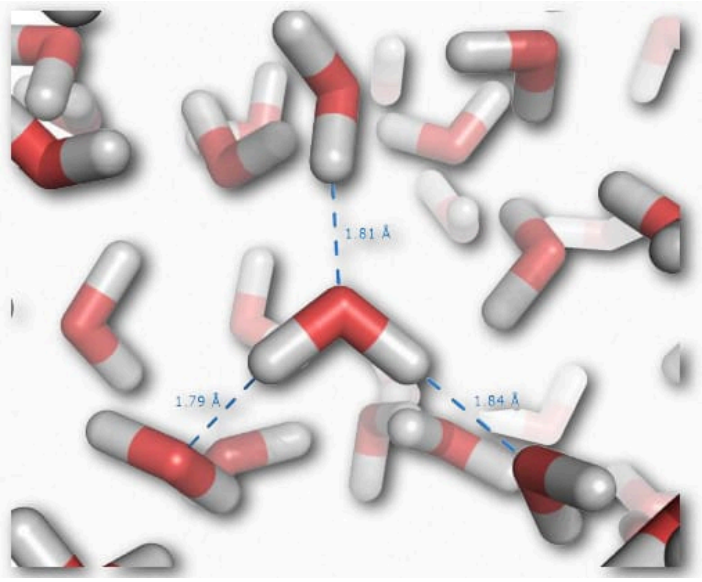


Figure 2.41 - Hydrogen bonding in liquid water Wikipedia

Individual hydrogen bonds are much weaker than a covalent bond, but collectively, they can exert strong forces. Consider liquid water, which contains enormous numbers of hydrogen bonds (Figure 2.41). These forces help water to remain liquid at room temperature. Other molecules lacking hydrogen bonds of equal or greater molecular weight than water, such as methane or carbon dioxide, are gases at the same temperature. Thus, the intermolecular interactions between water molecules help to “hold” water together and remain a liquid. Notably, only by raising the temperature of water to boiling are the forces of hydrogen bonding overcome, allowing water to become fully gaseous.

Hydrogen bonds are important forces in biopolymers that include DNA, proteins, and cellulose. All of these polymers lose their native structures upon boiling. Hydrogen bonds between amino acids that are close to each other in primary structure can give rise to regular repeating structures, such as helices or pleats, in proteins (secondary structure).

Ionic interactions

Ionic interactions are important forces stabilizing protein structure that arise from ionization of R-groups in the amino acids comprising a protein. These include the carboxyl amino acids (HERE), the amine amino acids as well as the sulfhydryl of cysteine and sometimes the hydroxyl of tyrosine.

Hydrophobic forces

Hydrophobic forces stabilize protein structure as a result of interactions that favor the exclusion of water. Non-polar amino acids (commonly found in the interior of proteins) favor associating with each other and this has the effect of excluding water. The excluded water has a higher entropy than water interacting with the hydrophobic side chains. This is because water aligns itself very regularly and in a distinct pattern when interacting with hydrophobic molecules.

When water is prevented from having these kinds of interactions, it is much more disordered that it would be if it could associate with the hydrophobic regions. It is partly for this reason that hydrophobic amino acids are found in protein interiors - so they can exclude water and increase entropy.

Disulfide bonds

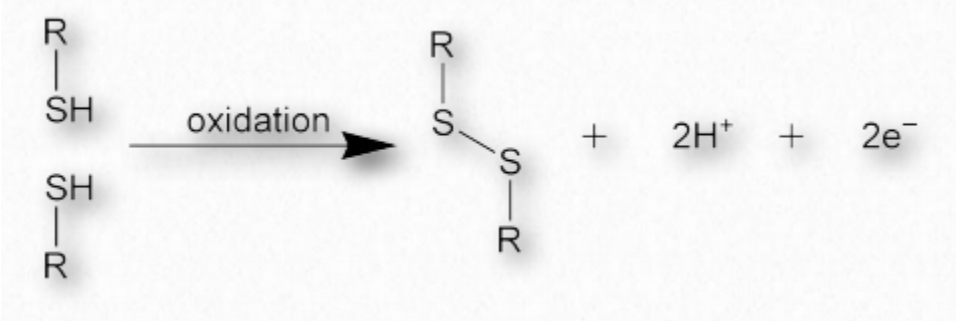


Figure 2.42 - Formation of a disulfide bond

Disulfide bonds, which are made when two sulfhydryl side-chains of cysteine are brought into close proximity, covalently join together different protein regions and can give great strength to the overall structure (Figures 2.42 & 2.43). An Ode to Protein Structure by Kevin Ahern The twenty wee amino A's Define a protein many ways Their order in a peptide chain Determines forms that proteins gain And when they coil, it leaves me merry Cuz that makes structures secondary It's tertiary, I am told That happens when a protein folds But folded chains are downright scary When put together quaternary They're nature's wonders, that's for sure Creating problems, making cures A fool can fashion peptide poems But proteins come from ribosoems These joined residues of cysteine are sometimes referred to as cystine. Disulfide bonds are the strongest of the forces stabilizing protein structure.

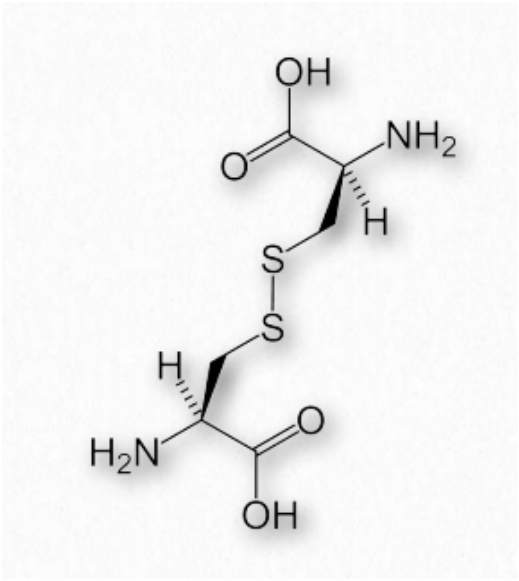


Figure 2.43 - Cystine - Two cysteines joined by a disulfide bond

van der Waals forces

van der Waals forces is a term used to describe various weak interactions, including those caused by attraction between a polar molecule and a transient dipole, or between two temporary dipoles. van der Waals forces are dynamic because of the fluctuating

nature of the attraction, and are generally weak in comparison to covalent bonds, but can, over very short distances, be significant.

Post-translational modifications

Post-translational modifications can result in formation of covalent bonds stabilizing proteins as well. Hydroxylation of lysine and proline in strands of collagen can result in cross-linking of these groups and the resulting covalent bonds help to strengthen and stabilize the collagen.

Folding models

Two popular models of protein folding are currently under investigation. In the first (diffusion collision model), a nucleation event begins the process, followed by secondary structure formation. Collisions between the secondary structures (as in the β-hairpin in Figure 2.37) allow for folding to begin. By contrast, in the nucleation-condensation model, the secondary and tertiary structures form together.

Folding in proteins occurs fairly rapidly (0.1 to 1000 seconds) and can occur during synthesis - the amino terminus of a protein can start to fold before the carboxyl terminus is even made, though that is not always the case.

Folding process

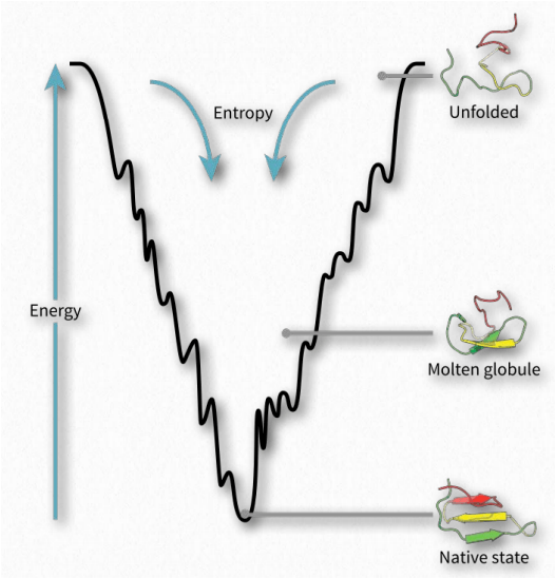


Figure 2.44 Folding funnel energy model of folding Wikipedia

Protein folding is hypothesized to occur in a “folding funnel” energy landscape in which a folded protein’s native state corresponds to the minimal free energy possible in conditions of the medium (usually aqueous solvent) in which the protein is dissolved. As seen in the diagram (Figure 2.44), the energy funnel has numerous local minima (dips) in which a folding protein can become trapped as it moves down the energy plot. Other factors, such as temperature, electric/magnetic fields, and spacial considerations likely play roles.

If external forces affect local energy minima during folding, the process and end-product can be influenced. As the speed of a car going down a road will affect the safety of the journey, so too do energy considerations influence and guide the folding process, resulting in fully functional, properly folded proteins in some cases and misfolded “mistakes” in others.

Getting stuck

As the folding process proceeds towards an energy minimum (bottom of the funnel in Figure 2.44), a protein can get “stuck” in any of the local minima and not reach the final folded state. Though the folded state is, in general, more organized and therefore has reduced entropy than the unfolded state, there are two forces that overcome the entropy decrease and drive the process forward.

The first is the magnitude of the decrease in energy as shown in the graph. Since $\Delta G = \Delta H - T\Delta S$, a decrease in ΔH can

overcome a negative ΔS to make ΔG negative and push the folding process forward. Favorable (decreased) energy conditions arise with formation of ionic bonds, hydrogen bonds, disulfide bonds, and metallic bonds during the folding process. In addition, the hydrophobic effect increases entropy by allowing hydrophobic amino acids in the interior of a folded protein to exclude water, thus countering the impact of the ordering of the protein structure by making the ΔS less negative.

Structure prediction

Computer programs are very good at predicting secondary structure solely based on amino acid sequence, but struggle with determining tertiary structure using the same information. This is partly due to the fact that secondary structures have repeating points of stabilization based on geometry and any regular secondary structure (e.g., α-helix) varies very little from one to another. Folded structures, though, have an enormous number of possible structures as shown by Levinthal's Paradox.

Spectroscopy

Because of our inability to accurately predict tertiary structure based on amino acid sequence, proteins structures are actually determined using techniques of spectroscopy. In these approaches, proteins are subjected to varied forms of electromagnetic radiation and the ways they interact with the radiation allows researchers to determine atomic coordinates at Angstrom resolution from electron densities (see X-ray crystallography) and how nuclei spins interact (see NMR).

Levinthal's paradox

In the late 1960s, Cyrus Levinthal outlined the magnitude of the complexity of the protein folding problem. He pointed out that for a protein with 100 amino acids, it would have 99 peptide bonds and 198 considerations for φ and ψ angles. If each of these had only three conformations, that would result in 3198 different possible foldings or 2.95x1094.

Even allowing a reasonable amount of time (one nanosecond) for each possible fold to occur, it would take longer than the age of the universe to sample all of them, meaning clearly that the process of folding is not occurring by a sequential random sampling and that attempts to determine protein structure by random sampling were doomed to fail. Levinthal, therefore, proposed that folding occurs by a sequential process that begins with a nucleation event that guides the process rapidly and is not unlike the funnel process depicted in Figure 2.44.

Diseases of protein misfolding

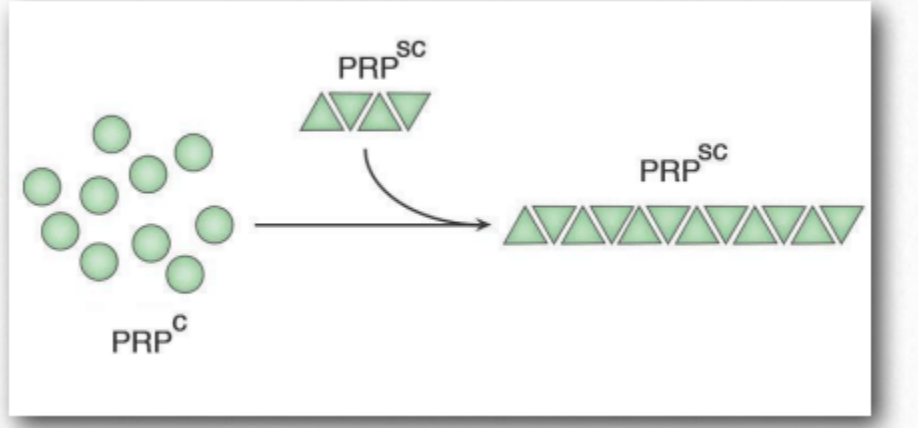


Figure 2.45 - Misfolding of the normal PRPc protein induced by PRPsc Image by Penelope Irving

The proper folding of proteins is essential to their function. It follows then that misfolding of proteins (also called proteopathy) might have consequences. In some cases, this might simply result in an inactive protein. Protein misfolding also plays a role in numerous diseases, such as Mad Cow Disease, Alzheimers, Parkinson's Disease, and CreutzfeldJakob disease. Many, but not all, misfolding diseases affect brain tissue.



Figure 2.46 - Cows with Mad Cow Disease lose their ability to stand

Insoluble deposits

Misfolded proteins will commonly form aggregates called amyloids that are harmful to tissues containing them because they change from being soluble to insoluble in water and form deposits. The process by which misfolding (Figure 2.45) occurs is not completely clear, but in many cases, it has been demonstrated that a "seed" protein which is misfolded can induce the same misfolding in other copies of the same protein. These seed proteins are known as prions and they act as infectious agents, resulting in the spread of disease. The list of human diseases linked to protein misfolding is long and continues to grow. A Wikipedia link is HERE.

Prions

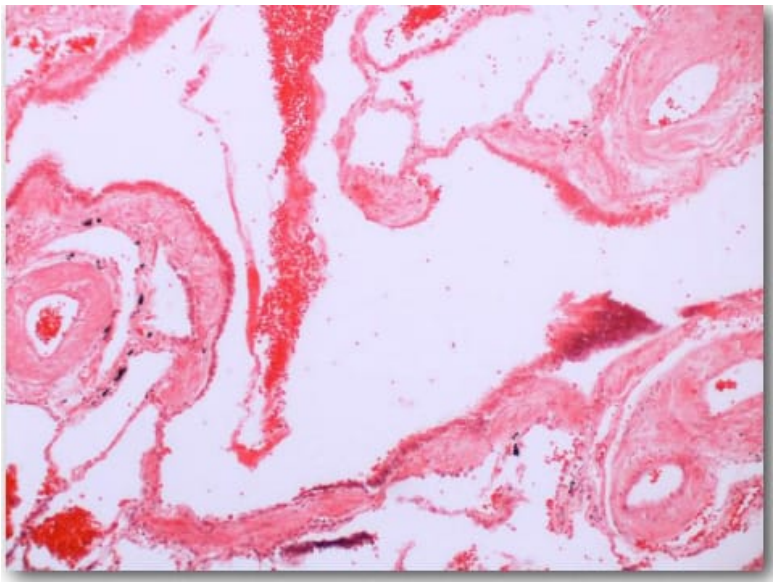


Figure 2.47 - Diffuse amyloidosis in a blood vessel (red dots) Wikipedia

Prions are infectious protein particles that cause transmissible spongiform encephalopathies (TSEs), the best known of which is Mad Cow disease. Other manifestations include the disease, scrapie, in sheep, and human diseases, such as CreutzfeldtJakob disease (CJD), Fatal Familial Insomnia, and kuru. The protein involved in these diseases is a membrane protein called PrP. PrP is encoded in the genome of many organisms and is found in most cells of the body. PrPc is the name given to the structure of PrP that is normal and not associated with disease. PrPSc is the name given to a misfolded form of the same protein, that is associated with the development of disease symptoms (Figure 2.45).

Misfolded

The misfolded PrPSc is associated with the TSE diseases and acts as an infectious particle. A third form of PrP, called PrPres can be found in TSEs, but is not infectious. The 'res' of PrPres indicates it is protease resistant. It is worth noting that all three forms of PrP have the same amino acid sequence and differ from each other only in the ways in which the polypeptide chains are folded. The most dangerously misfolded form of PrP is PrPSc, because of its ability to act like an infectious agent - a seed protein that can induce misfolding of PrPc , thus converting it into PrPSc.

Function

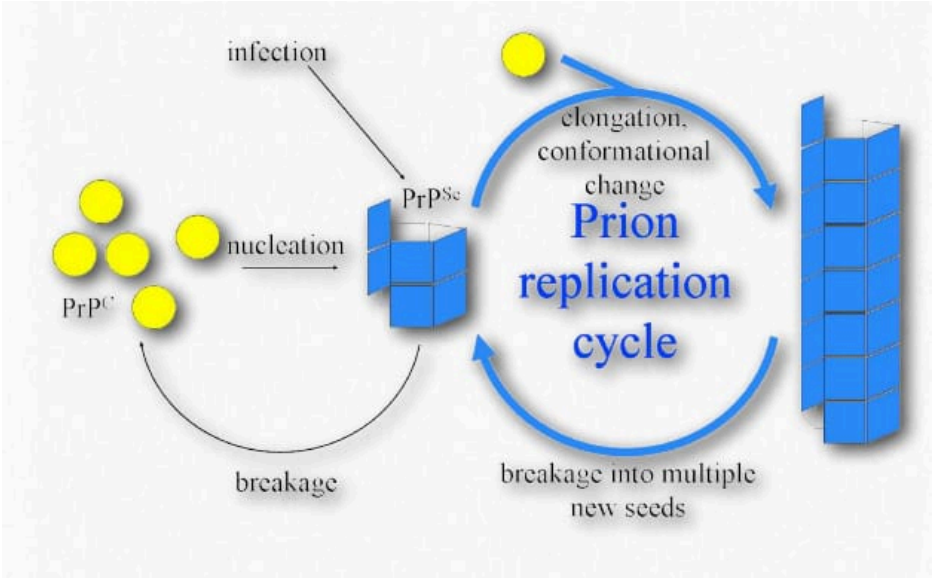


Figure 2.48 - One model of prion propagation Wikipedia

The function of PrPc is unknown. Mice lacking the PrP gene do not have major abnormalities. They do appear to exhibit problems with long term memory, suggesting a function for PrPc . Stanley Prusiner, who discovered prions and coined the term, received the Nobel Prize in Medicine in 1997 for his work. I think that if I chanced to be on A protein making up a prion I'd twist it and for goodness sakes Stop it from making fold mistakes

Amyloids

Amyloids are a collection of improperly folded protein aggregates that are found in the human body. As a consequence of their misfolding, they are insoluble and contribute to some twenty human diseases including important neurological ones involving prions. Diseases include (affected protein in parentheses) - Alzheimer's disease (Amyloid β), Parkinson's disease (α-synuclein), Huntington's disease (huntingtin), rheumatoid arthritis (serum amyloid A), fatal familial insomnia (PrPSc), and others.

Amino acid sequence plays a role in amyloidogenesis. Glutamine-rich polypeptides are common in yeast and human prions. Trinucleotide repeats are important in Huntington's disease. Where sequence is not a factor, hydrophobic association between β-sheets can play a role.

Amyloid β

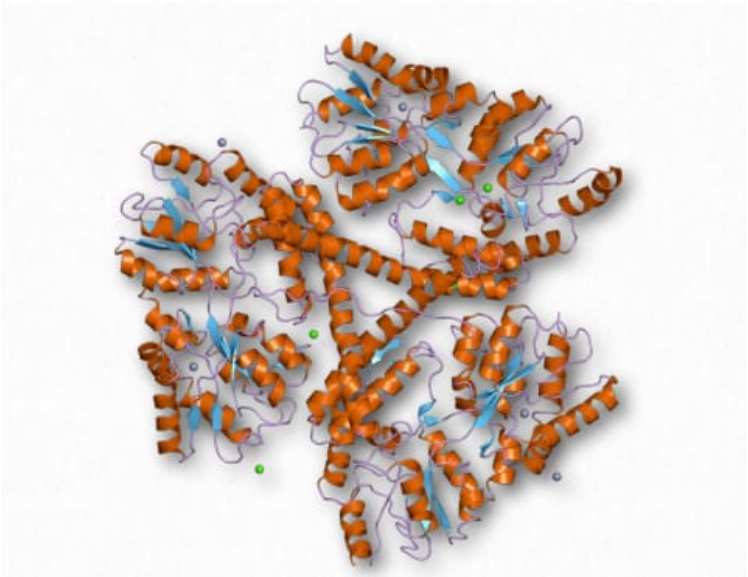


Figure 2.49 - Huntingtin

Amyloid β refers to collections of small proteins (36-43 amino acids) that appear to play a role in Alzheimer's disease. (Tau protein is the other factor.) They are, in fact, the main components of amyloid plaques found in the brains of patients suffering from the disease and arise from proteolytic cleavage of a larger amyloid precursor glycoprotein called Amyloid Precursor Protein, an integral membrane protein of nerve cells whose function is not known. Two proteases, β-secretase and γ- secretase perform this function. Amyloid β proteins are improperly folded and appear to induce other proteins to misfold and thus precipitate and form the amyloid characteristic of the disease. The plaques are toxic to nerve cells and give rise to the dementia characteristic of the disease.

It is thought that aggregation of amyloid β proteins during misfolding leads to generation of reactive oxygen species and that this is the means by which neurons are damaged. It is not known what the actual function of amyloid β is. Autosomal dominant mutations in the protein lead to early onset of the disease, but this occurs in no more than 10% of the cases. Strategies for treating the disease include inhibition of the secretases that generate the peptide fragments from the amyloid precursor protein.

Huntingtin

Huntingtin is the central gene in Huntington's disease. The protein made from it is glutamine rich, with 6-35 such residues in its wild-type form. In Huntington's disease, this gene is mutated, increasing the number of glutamines in the mutant protein to between 36 and 250. The size of the protein varies with the number of glutamines in the mutant protein, but the wild-type protein has over 3100 amino acids and a molecular weight of about 350,000 Da. Its precise function is not known, but huntingtin is found in nerve cells, with the highest level in the brain. It is thought to possibly play roles in transport, signaling, and protection against apoptosis. Huntingtin is also required for early embryonic development. Within the cell, huntingtin is found localized primarily with microtubules and vesicles.

Trinucleotide repeat

The huntingtin gene contains many copies of the sequence CAG (called trinucleotide repeats), which code for the many glutamines in the protein. Huntington's disease arises when extra copies of the CAG sequence are generated when the DNA of the gene is being copied. Expansion of repeated sequences can occur due to slipping of the polymerase relative to the DNA template during replication. As a result, multiple additional copies of the trinucleotide repeat may be made, resulting in proteins with variable numbers of glutamine residues. Up to 35 repeats can be tolerated without problem. The number of repeats can expand over the course of a person's lifetime, however, by the same mechanism. Individuals with 36-40 repeats begin to show signs of the disease and if there are over 40, the disease will be present.

Molecular chaperones

The importance of the proper folding of proteins is highlighted by the diseases associated with misfolded proteins, so it is no surprise, then, that cells expend energy to facilitate the proper folding of proteins. Cells use two classes of proteins known as molecular chaperones, to facilitate such folding in cells. Molecular chaperones are of two kinds, the chaperones, and the

chaperonins. An example of the first category is the Hsp70 class of proteins. Hsp stands for “heat shock protein”, based on the fact that these proteins were first observed in large amounts in cells that had been briefly subjected to high temperatures. Hsps function to assist cells in stresses arising from heat shock and exposure to oxidizing conditions or toxic heavy metals, such as cadmium and mercury. However, they also play an important role in normal conditions, where they assist in the proper folding of polypeptides by preventing aberrant interactions that could lead to misfolding or aggregation. The Hsp70 proteins are found in almost all cells and use ATP hydrolysis to stimulate structural changes in the shape of the chaperone to accommodate binding of substrate proteins. The binding domain of Hsp70s contains a β-barrel structure which wraps around the polypeptide chain of the substrate and has affinity for hydrophobic side chains of amino acids. As shown in Figure 2.50, Hsp70 binds to polypeptides as they emerge from ribosomes during protein synthesis. Binding of substrate stimulates ATP hydrolysis and this is facilitated by another heat shock protein known as Hsp40. The hydrolysis of ATP causes the Hsp70 to taken on a closed conformation that helps shield exposed hydrophobic residues and prevent aggregation or local misfolding.

After protein synthesis is complete, ADP is released and replaced by ATP and this results in release of the substrate protein, which then allows the full length polypeptide to fold correctly.

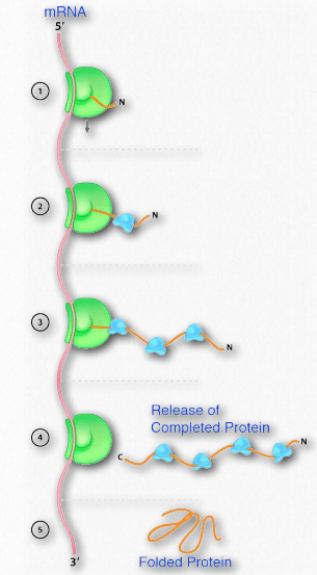


Figure 2.50 - Action of Hsp70 (blue) to facilitate proper folding of a protein (orange) Image by Aleia Kim

In heat shock

In times of heat shock or oxidative stress, Hsp70 proteins bind to unfolded hydrophobic regions of proteins to similarly prevent them from aggregating and allowing them to properly refold. When proteins are damaged, Hsp70 recruits enzymes that ubiquitinate the damaged protein to target them for destruction in proteasomes. Thus, the Hsp70 proteins play an important role in ensuring not only that proteins are properly folded, but that damaged or nonfunctional proteins are removed by degradation in the proteasome.

Chaperonins

A second class of proteins involved in assisting other proteins to fold properly are known as chaperonins. There are two primary categories of chaperonins - Class I (found in bacteria, chloroplasts, and mitochondria) and Class II (found in the cytosol of eukaryotes and archaebacteria). The best studied chaperonins are the GroEL/GroES complex proteins found in bacteria (Figure 2.51).

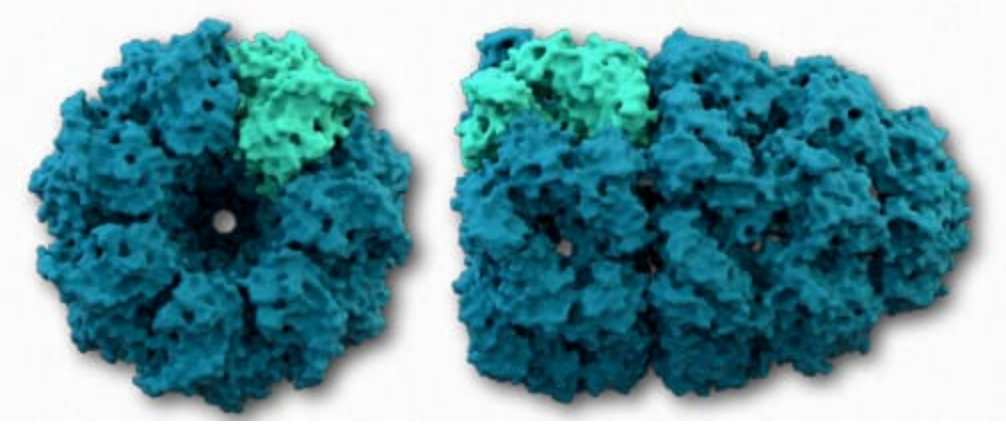


Figure 2.51 - View from bottom of GroEL (left) and GroEL/ GroES complex (right) Wikipedia

GroEL/GroES may not be able to undo aggregated proteins, but by facilitating proper folding, it provides competition for misfolding as a process and can reduce or eliminate problems arising from improper folding. GroEL is a double-ring 14mer with a hydrophobic region that can facilitate folding of substrates 15-60 kDa in size. GroES is a singlering heptamer that binds to GroEL in the presence of ATP and functions as a cover over GroEL. Hydrolysis of ATP by chaperonins induce large conformational changes that affect binding of substrate proteins and their folding. It is not known exactly how chaperonins fold proteins. Passive models postulate the chaperonin complex functioning inertly by preventing unfavorable intermolecular interactions or placing restrictions on spaces available for folding to occur. Active models propose that structural changes in the chaperonin complex induce structural changes in the substrate protein.

Protein breakdown

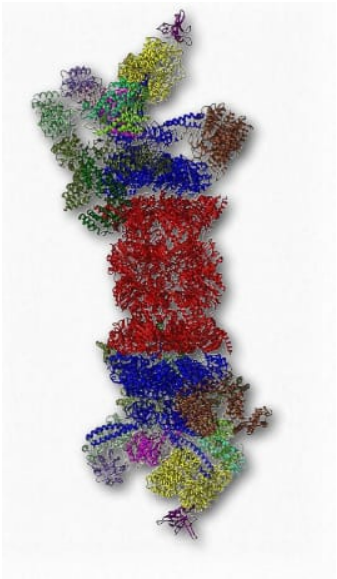


Figure 2.52 - 26S proteasome. Active site shown in red Wikipedia

Another protein complex that has an important function in the lifetime dynamics of proteins is the proteasome (Figure 2.52). Proteasomes, which are found in all eukaryotes and archaeans, as well as some bacteria, function to break down unneeded or damaged proteins by proteolytic degradation. Proteasomes help to regulate the concentration of some proteins and degrade ones that are misfolded. The proteasomal degradation pathway plays an important role in cellular processes that include progression through the cell cycle, modulation of gene expression, and response to oxidative stresses.

Degradation in the proteasome yields short peptides seven to eight amino acids in length. Threonine proteases play important roles. Breakdown of these peptides yields individual amino acids, thus facilitating their recycling in cells. Proteins are targeted for degradation in eukaryotic proteasomes by attachment to multiple copies of a small protein called ubiquitin (8.5 kDa - 76 amino acids). The enzyme catalyzing the reaction is known as ubiquitin ligase. The resulting polyubiquitin chain is bound by the proteasome and degradation begins. Ubiquitin was named due to it ubiquitously being found in eukaryotic cells.

Ubiquitin

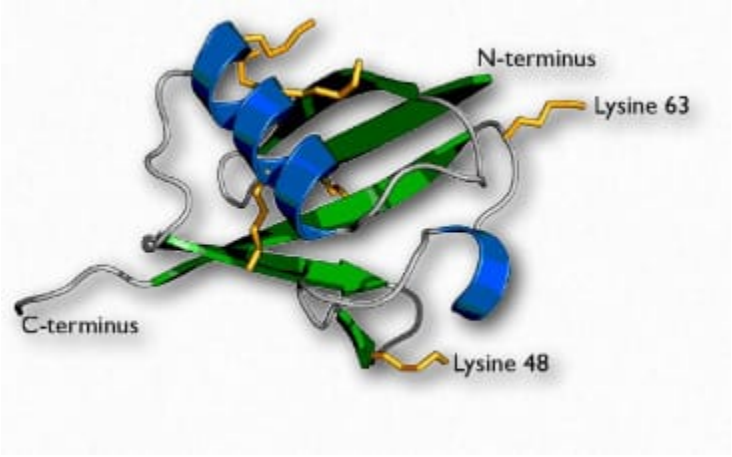


Figure 2.53 - Ubiquitin (lysine side chains shown in yellow) Wikipedia

Ubiquitin (Figure 2.53) is a small (8.5 kDa) multi-functional protein found in eukaryotic cells. It is commonly added to target proteins by action of ubiquitin ligase enzymes (E3 in Figure 2.54). One (ubiquitination) or many (polyubiquitination) ubiquitin molecules may be added. Attachment of the ubiquitin is through the side chain of one of seven different lysine residues in ubiquitin.

The addition of ubiquitin to proteins has many effects, the best known of which is targeting the protein for degradation in the proteasome. Proteasomal targeting is seen when polyubiquitination occurs at lysines #29 and 48. Polyubiquitination or monoubiquitination at other lysines can result in altered cellular location and changed protein-protein interactions. The latter may alter affect inflammation, endocytic trafficking, translation and DNA repair.

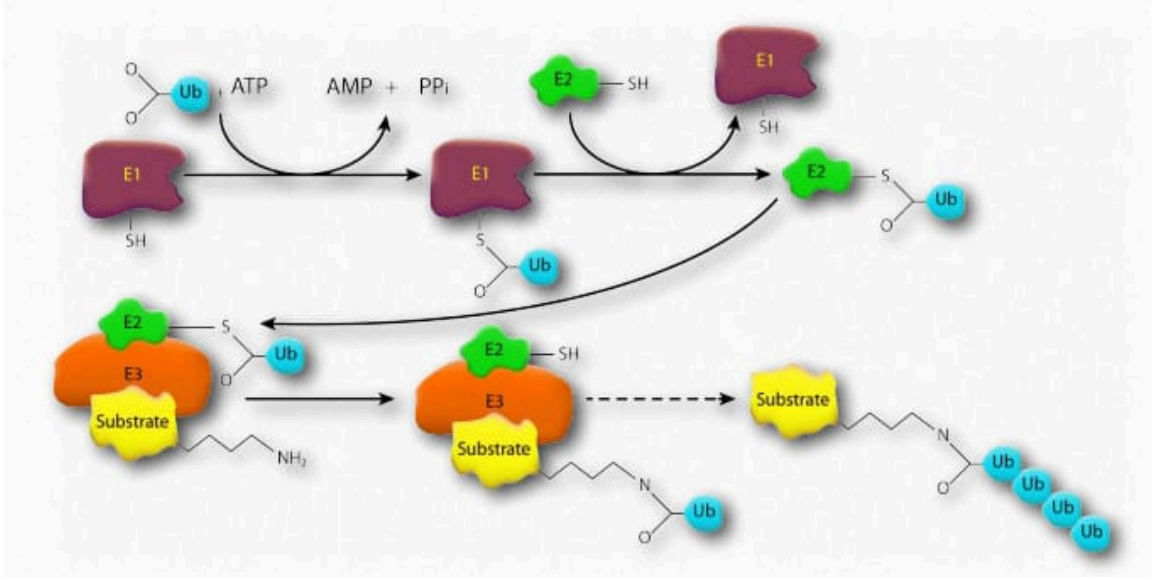


Figure 2.54 - Pathway for ubiquitination of a target substrate protein Image by Pehr Jacobson

Ubiquitin ligase malfunction

Parkin is a Parkinson's disease-related protein that, when mutated, is linked to an inherited form of the disease called autosomal recessive juvenile Parkinson's disease. The function of the protein is not known, but it is a component of the E3 ubiquitin ligase system responsible for transferring ubiquitin from the E2 protein to a lysine side chain on the target protein. It is thought that mutations in parkin lead to proteasomal dysfunction and a consequent inability to break down proteins harmful to dopaminergic neurons. This results in the death or malfunction of these neurons, resulting in Parkinson's disease.

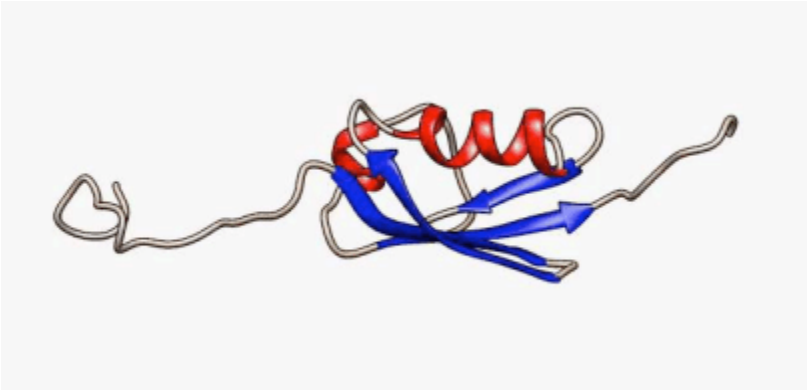
Intrinsically disordered proteins



Movie 2.1 - Dynamic movement of cytochrome C in solution Wikipedia

As is evident from the many examples described elsewhere in the book, the 3-D structure of proteins is important for their function. But, increasingly, it is becoming evident that not all proteins fold into a stable structure. Studies on the so-called intrinsically disordered proteins (IDPs) in the past cou- ple of decades has shown that many proteins are biologically active, even thought they fail to fold into stable structures. Yet other proteins exhibit regions that remain unfolded (IDP regions) even as the rest of the polypeptide folds into a structured form.

Intrinsically disordered proteins and disordered regions within proteins have, in fact, been known for many years, but were regarded as an anomaly. It is only recently, with the realization that IDPs and IDP regions are widespread among eukaryotic proteins, that it has been recognized that the observed disorder is a "feature, not a bug".



Movie 2.2 SUMO-1, a protein with intrinsically disordered sections Wikipedia

Comparison of IDPs shows that they share sequence characteristics that appear to favor their disordered state. That is, just as some amino acid sequences may favor the folding of a polypeptide into a particular structure, the amino acid sequences of IDPs favor their remaining unfolded. IDP regions are seen to be low in hydrophobic residues and unusually rich in polar residues and proline. The presence of a large number of charged amino acids in the IDPs can inhibit folding through charge repulsion, while the lack of hydrophobic residues makes it difficult to form a stable hydrophobic core, and proline discourages the formation of helical structures. The observed differences between amino acid sequences in IDPs and structured proteins have been used to design algorithms to predict whether a given amino acid sequence will be disordered.

What is the significance of intrinsically disordered proteins or regions? The fact that this property is encoded in their amino acid sequences suggests that their disorder may be linked to their function. The flexible, mobile nature of some IDP regions may play a crucial role in their function, permitting a transition to a folded structure upon binding a protein partner or undergoing post-translational modification. Studies on several wellknown proteins with IDP regions suggest some answers. IDP regions may enhance the ability of proteins like the lac repressor to translocate along the DNA to search for specific binding sites. The flexibility of IDPs can also be an asset in protein-protein interactions, especially for proteins that are known to interact with many different protein partners.

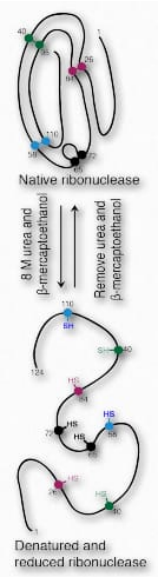


Figure 2.55 - Denaturation and renaturation of ribonuclease Wikipedia

For example, p53 has IDP regions that may allow the protein to interact with a variety of functional partners. Comparison of the known functions of proteins with predictions of disorder in these proteins suggests that IDPs and IDP regions may disproportionately function in signaling and regulation, while more structured proteins skew towards roles in catalysis and transport. Interestingly, many of the proteins found in both ribosomes and spliceosomes are predicted to have IDP regions that may play a part in correct assembly of these complexes. Even though IDPs have not been studied intensively for very long, what little is known of them suggests that they play an important and underestimated role in cells.

Metamorphic proteins

Another group of proteins that have recently changed our thinking about protein structure and function are the so-called metamorphic proteins. These proteins are capable of forming more than one stable, folded state starting with a single amino acid sequence. Although it is true that multiple folded conformations are not ruled out by the laws of physics and chemistry, metamorphic proteins are a relatively new discovery. It was known, of course, that prion proteins were capable of folding into alternative structures, but metamorphic proteins appear to be able to toggle back and forth between two stable structures. While in some cases, the metamorphic protein undergoes this switch in response to binding another molecule, some proteins that can accomplish this transition on their own. An interesting example is the signaling molecule, lymphotactin. Lymphotactin has two biological functions that are carried out by its two conformers- a monomeric form that binds the lymphotactin receptor and a dimeric form that binds heparin. It is possible that this sort of switching is more widespread than has been thought.

Refolding denatured proteins

All information for protein folding is contained in the amino acid sequence of the protein. It may seem curious then that most proteins do not fold into their proper, fully active form after they have been+++ denatured and the denaturant is removed. A few do, in fact. One good example is bovine ribonuclease (Figure 2.55). Its catalytic activity is very resistant to heat and urea and attempts to denature it don't work very well. However, if one treats the enzyme with β-mercaptoethanol (which breaks disulfide bonds) prior to urea treatment and/or heating, activity is lost, indicating that the covalent disulfide bonds help stabilize the overall enzyme structure and when they are broken, denaturation can readily occur. When the mixture cools back down to room temperature, over time some enzyme activity reappears, indicating that ribonuclease re-folded under the new conditions.

Interestingly, renaturation will occur maximally if a tiny amount of β-mercaptoethanol is left in the solution during the process. The reason for this is because β- mercaptoethanol permits reduction (and breaking) of accidental, incorrect disulfide bonds during the folding process. Without it, these disulfide bonds will prevent proper folds from forming.

Irreversible denaturation

Most enzymes, however, do not behave like bovine ribonuclease. Once denatured, their activity cannot be recovered to any significant There are not very many ways Inactivating RNase It's stable when it's hot or cold Because disulfides tightly hold If you desire to make it stall Use hot mercaptoethanol extent. This may seem to contradict the idea of folding information being

inherent to the sequence of amino acids in the protein. It does not.

Most enzymes don't refold properly after denaturation for two reasons. First, normal folding may occur as proteins are being made. Interactions among amino acids early in the synthesis are not "confused" by interactions with amino acids later in the synthesis because those amino acids aren't present as the process starts.

Chaperonins' role

In other cases, the folding process of some proteins in the cell relied upon action of chaperonin proteins (see HERE). In the absence of chaperonins, interactions that might result in misfolding occur, thus preventing proper folding. Thus, early folding and the assistance of chaperonins eliminate some potential "wrong-folding" interactions that can occur if the entire sequence was present when folding started.

Quaternary structure

A fourth level of protein structure is that of quaternary structure. It refers to structures that arise as a result of interactions between multiple polypeptides. The units can be identical multiple copies or can be different polypeptide chains. Adult hemoglobin is a good example of a protein with quaternary structure, being composed of two identical chains called α and two identical chains called β.

Though the α-chains are very similar to the β- chains, they are not identical. Both of the α- and the β-chains are also related to the single polypeptide chain in the related protein called myoglobin. Both myoglobin and hemoglobin have similarity in binding oxygen, but their behavior towards the molecule differ significantly. Notably, hemoglobin's multiple subunits (with quaternary structure) compared to myoglobin's single subunit (with no quaternary structure) give rise to these differences.

References

1. https://en.Wikipedia.org/wiki/Van_der_W aals_force 105

This page titled 2.3: Structure & Function- Proteins I is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

3.1: Basic Concepts in Membranes

Source: BiochemFFA_3_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Lipid bilayers

The protective membrane around cells contains many components, including cholesterol, proteins, glycolipids, glycerophospholipids, and sphingolipids. The last two of these will, when mixed vigorously with water, spontaneously form what is called a lipid bilayer (Figure 3.1), which serves as a protective boundary for the cell that is largely impermeable to the movement of most materials across it. With the notable exceptions of water, carbon dioxide, carbon monoxide, and oxygen, most polar/ionic require transport proteins to help them to efficiently navigate across the bilayer. The orderly movement of these compounds is critical for the cell to be able to 1) get food for energy; 2) export materials; 3) maintain osmotic balance; 4) create gradients for secondary transport; 5) provide electromotive force for nerve signaling; and 6) store energy in electrochemical gradients for ATP production (oxidative phosphorylation or photosynthesis). In some cases, energy is required to move the substances (active transport).

Facilitated Diffusion

In other cases, no external energy is required and they move by diffusion through specific cellular channels. This is referred to as facilitated diffusion. Before we discuss movement of materials across membranes, it is appropriate we discuss the composition of cellular membranes. Plasma membranes differ from cell walls both in the materials comprising them and in their flexibility. Cell walls will be covered near the end of this chapter.

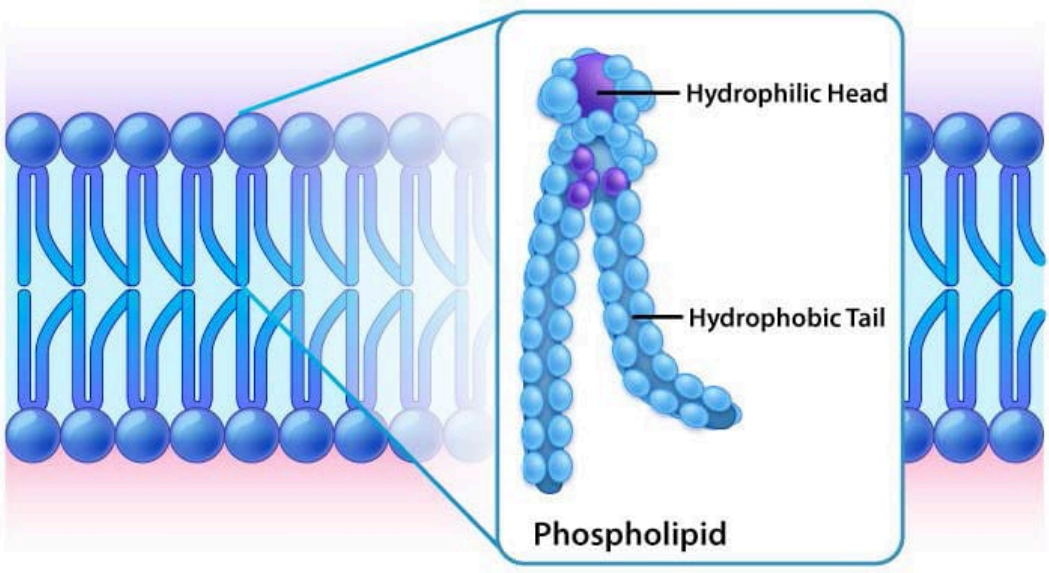


Figure 3.1: Lipid bilayer closeup Image by Aleia Kim

Though some cells do not have cell walls (animal cells) and others do (bacteria, fungi, and plants), there is commonality among cells in that they all possess plasma membranes. There is also commonality in the components of the membranes, though the relative amount of constituents varies. Figures 3.1 and 3.2 illustrate the structure and environments of plasma membranes. All plasma membranes contain a significant amount of amphiphilic substances linked to fatty acids. These include the glycerophospholipids and the sphingolipids. The fatty acid(s) are labeled as hydrophobic tails in the figures.

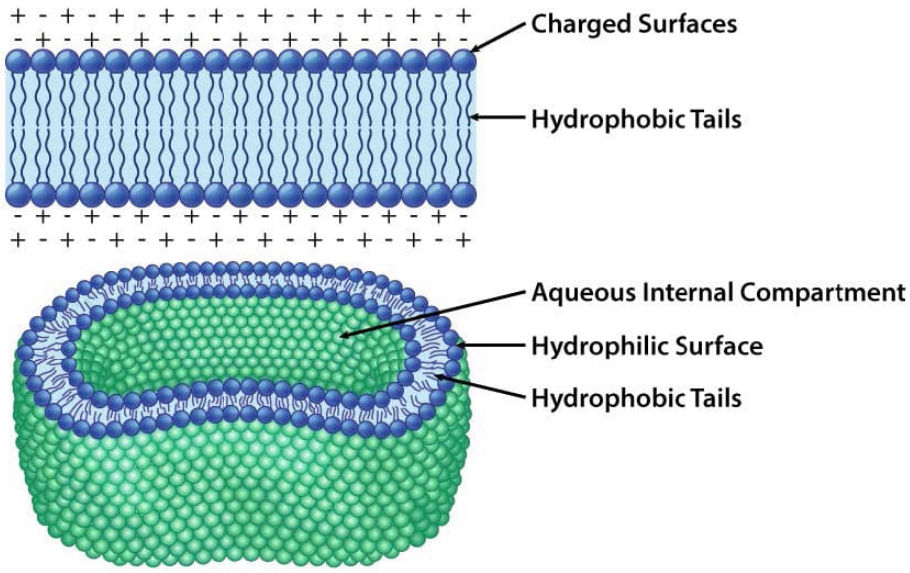


Figure 3.2: Organization of the lipid bilayer Image by Aleia Kim

Hydrophilic heads

The composition of the hydrophilic heads varies considerably. In glycerophospholipids, a phosphate is always present, of course, and it is often esterified to another substance to make a phosphatide (Figure 3.3). Common compounds linked to the phosphate (at the X position) include serine, ethanolamine, and choline. These vary in the their charges so in this way, the charge on the external or internal surface can be controlled. Cells tend to have more negative charges on the exterior half of the lipid bilayer (called the outer leaflet) and more positive charges on the interior half (inner leaflet).

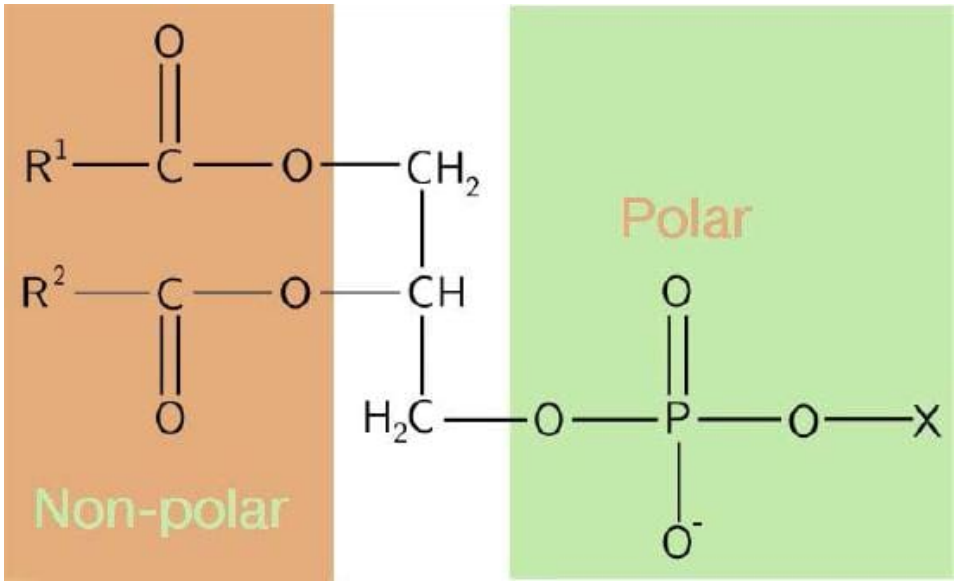


Figure 3.3 - Schematic diagram of a phosphatide

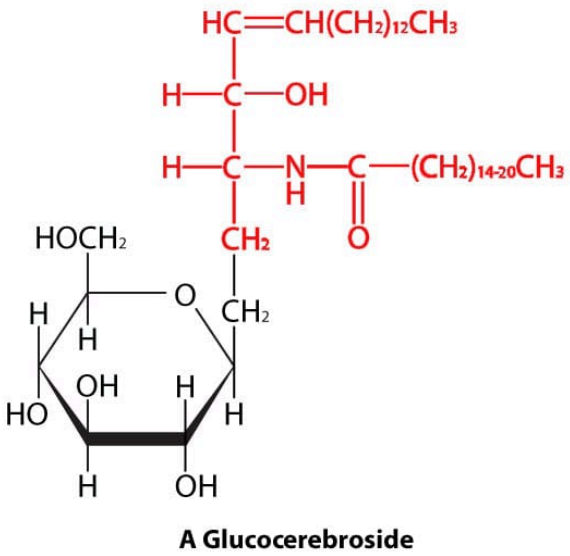


Figure 3.4: A sphingolipid. Polar head in black. Nonpolar tail in red Image by Aleia Kim

Sphingolipids

In sphingolipids (Figure 3.4), the hydrophilic head can contain a phosphate linked to ethanolamine or choline and this describes the structure of sphingomyelin, an important component of neural membranes. Most sphingolipids lack the phosphate and have instead a hydrophilic head of a single sugar (cerebrosides) or a complex oligosaccharide (gangliosides).

Water exclusion

In each case, the glycerophospholipid or sphingolipid has one end that is polar and one end that is non-polar. As we saw in the organization of amino acids with hydrophobic side chains occurring preferentially on the inside of a folded protein to exclude water, so too do the non-polar portions of these amphiphilic molecules arrange themselves so as to exclude water. Remember that the cytoplasm of a cell is mostly water and the exterior of the cell is usually bathed in an aqueous layer. It therefore makes perfect sense that the polar portions of the membrane molecules arrange themselves as they do - polar parts outside interacting with water and non-polar parts in the middle of the bilayer avoiding/excluding water.

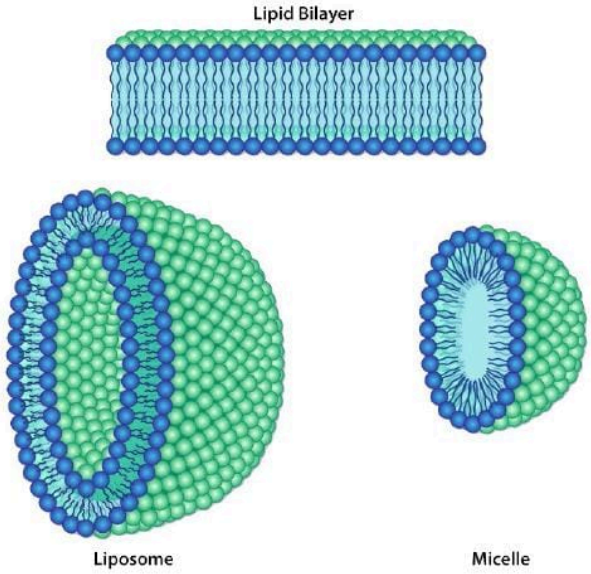


Figure 3.6 - Different lipid bilayer structures Image by Aleia Kim

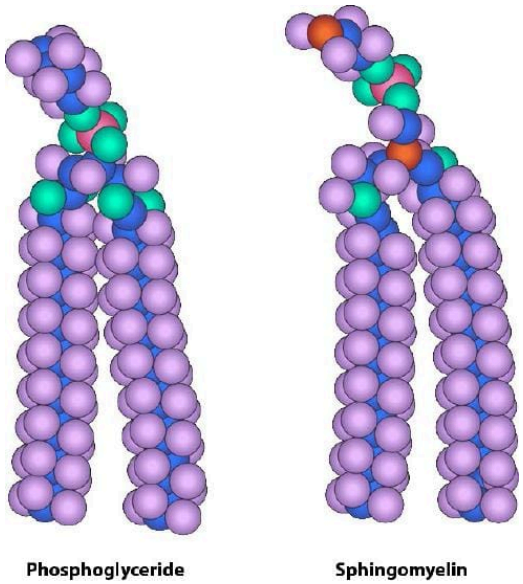


Figure 3.5 - Similarity of form between a phosphoglyceride and sphingomyelin Image by Aleia Kim

Composition Bias

The plasma membrane has distinct biases of composition relative to its inside and the outside (Figure 3.7). First, glycosylation (of lipids and proteins) has the sugar groups located almost exclusively on the outside of the cell, away from the cytoplasm (Figure 3.8). Among the membrane lipids, sphingolipids are much more commonly glycosylated than glycerophospholipids. In addition, some of the glyerophospholipids are found preferentially on one side or the other (Figure 3.7). Phosphatidylserine and phosphatidylethanolamine are found preferentially within the inner leaflet of the plasma membrane, whereas phosphatidylcholine tends to be located on the outer leaflet. In the process of apoptosis, the phosphatidylserines appear on the outer leaflet where they serve as a signal to macrophages to bind and destroy the cell. Sphingolipids are found preferentially in the plasma membrane and are almost completely absent from mitochondrial and endoplasmic reticulum membranes (Figure 3.9).

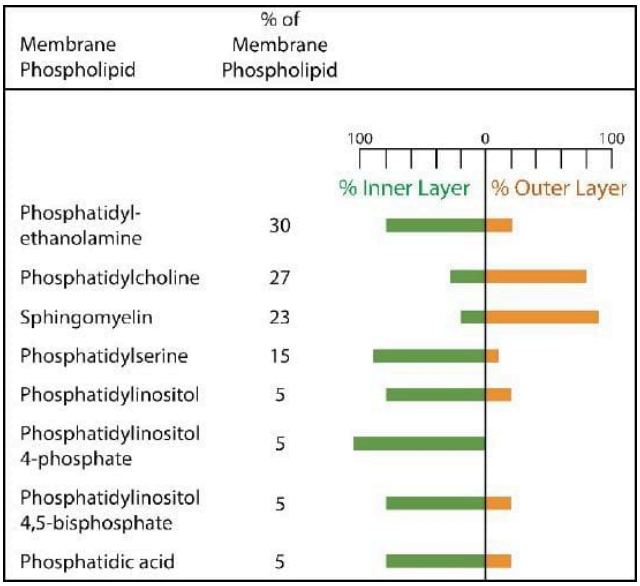


Figure 3.7: Differential distribution of membrane lipids by inner and outer leaflet Image by Pehr Jacobson

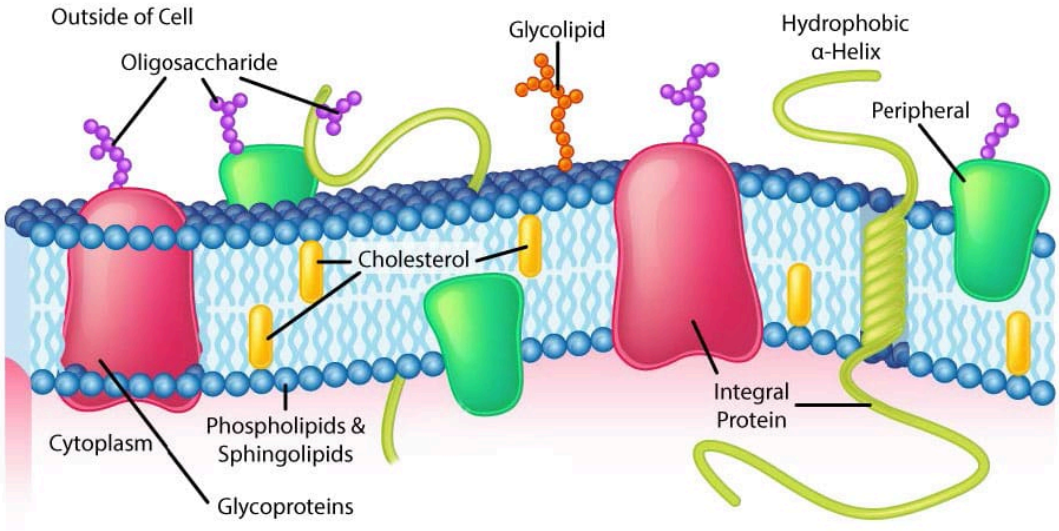


Figure 3.8 - Molecular organization of the lipid bilayer Image by Aleia Kim

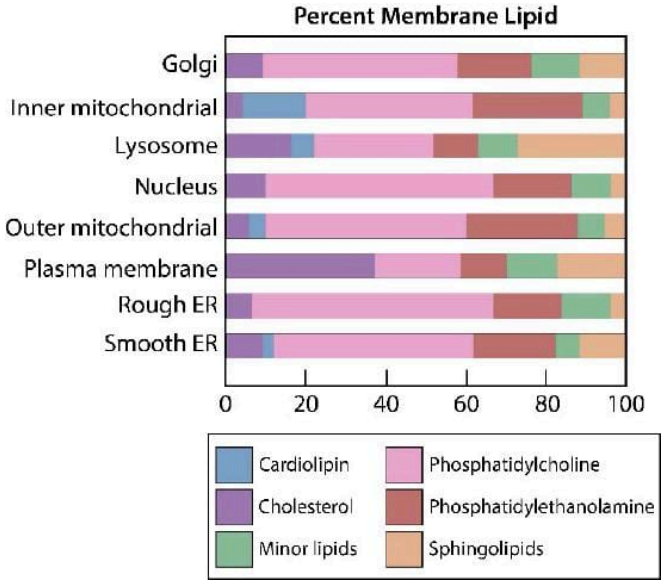


Figure 3.9 - Distribution of lipids in organelle membranes Image by Pehr Jacobson

Organelle membranes

Bias of lipid composition also exists with respect to organelle membranes. The unusual diphosphodiglycerolipid known as cardiolipin, for example, is almost only found in mitochondrial membranes (see HERE) and like phosphatidylserine, its movement is an important step in apoptosis. In signaling, phosphatidylinositols play important roles providing second messengers upon being cleaved (see HERE).

Lateral Diffusion

Movement of lipids within each leaflet of the lipid bilayer occurs readily and rapidly due to membrane fluidity. This type of movement is called lateral diffusion and can be measured by the technique called FRAP (Figure 3.10, see HERE also). In this method, a laser strikes and stains a section of the lipid bilayer of a cell, leaving a spot as shown in B. Over time, the stain diffuses out ultimately across the entire lipid bilayer, much like a drop of ink will diffuse throughout when added to a glass of water. A measurement of the rate of diffusion gives an indication of the fluidity of a membrane.

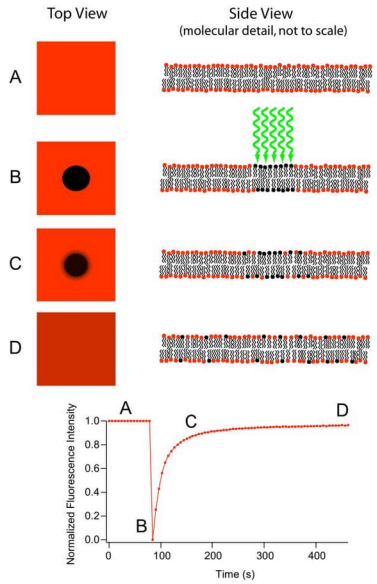


Figure 3.10 - Fluorescence recovery after photobleaching (FRAP)

Transverse Diffusion

While the movement in lateral diffusion occurs rapidly, movement of molecules from one leaflet over to the other leaflet occurs much more slowly. This type of molecular movement is called transverse diffusion and is almost nonexistent in the absence of enzyme action. Remember that there is a bias of distribution of molecules between leaflets of the membrane, which means that something must be moving them.

There are three enzymes that catalyze movement of compounds in transverse diffusion. Flippases move membrane glycerophospholipids/ sphingolipids from outer leaflet to inner leaflet (cytoplasmic side) of cell. Floppases move membrane lipids in the opposite direction. Scramblases move in either direction.

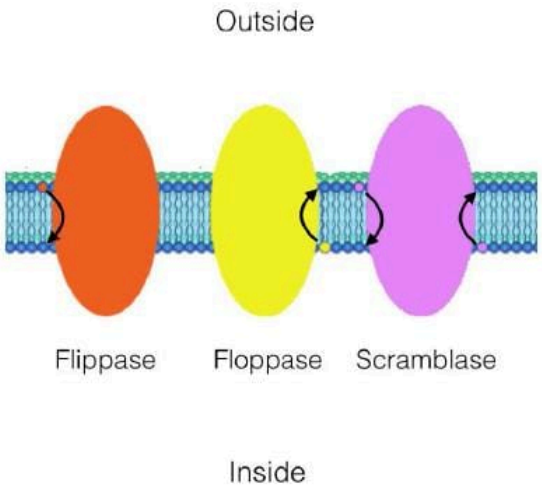


Figure 3.11 - Catalytic action of a flippase, a floppase, and a scramblase

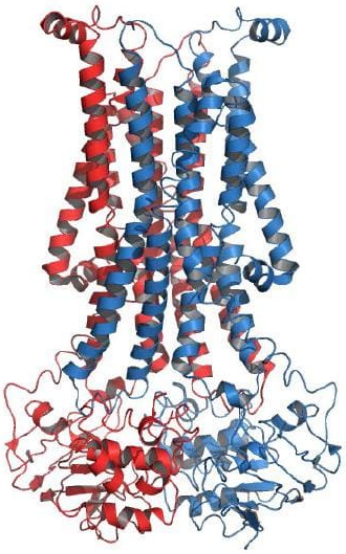


Figure 3.12 - Structure of a flippase Wikipedia

Other components of lipid bilayer

Besides glycerophospholipids and sphingolipids, there are other materials commonly found in lipid bilayers of cellular membranes. Two important prominent ones are cholesterol (Figure 3.13) and proteins. Besides serving as a metabolic precursor of steroid hormones and the bile acids, cholesterol's main role in cells is in the membranes. The flatness and hydrophobicity of the sterol rings allow cholesterol to interact with the nonpolar portions of the lipid bilayer while the hydroxyl group on the end can interact with the hydrophilic part.

Membrane fluidity

Cholesterol's function in the lipid bilayer is complex (Figure 3.13). It influences membrane fluidity. Figure 3.14 shows the phase transition for a membrane as it is heated, moving from a more "frozen" character to that of a more "fluid" one as the temperature rises. The mid-point of this transition, referred to as the Tm, is influenced by the fatty acid composition of the lipid bilayer compounds. Longer and more saturated fatty acids will favor higher Tm values, whereas unsaturation and short fatty acids will favor lower Tm values. It is for this reason that fish, which live in cool environments, have a higher level of unsaturated fatty acids in them - to use to make membrane lipids that will remain fluid at ocean temperatures. Interestingly, cholesterol does not change the Tm value, but instead widens the transition range between frozen and fluid forms of the membrane, allowing it to have a wider range of fluidity.

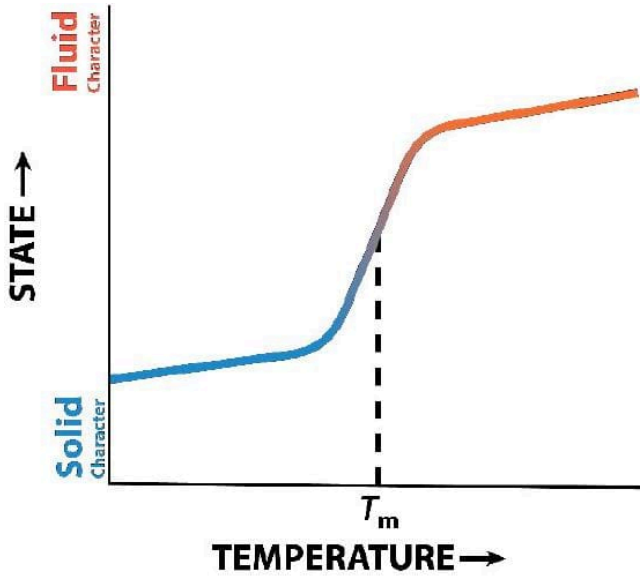


Figure 3.14 - Membrane transition temperature Image by Aleia Kim

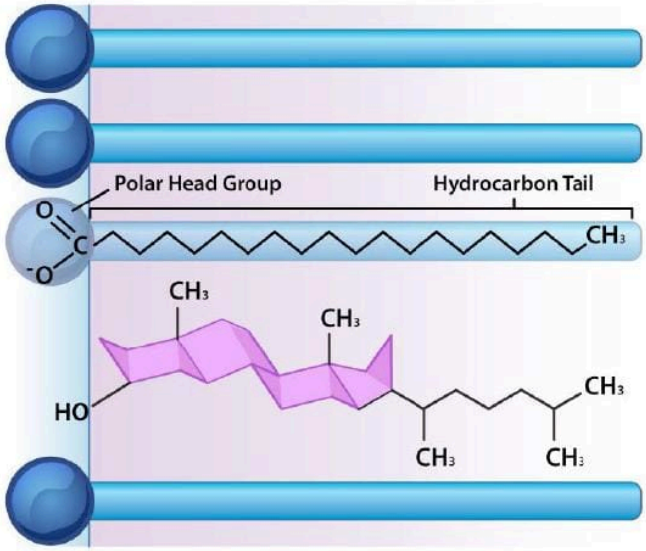


Figure 3.13 - Cholesterol in the lipid bilayer Image by Aleia Kim

Lipid Rafts

Cholesterol is also abundantly found in membrane structures called lipid rafts. Depicted in Figure 3.15, lipid rafts are organized structures within the membrane typically containing signaling molecules and other integral membrane proteins. Lipid rafts affect membrane fluidity, neurotransmission, and trafficking of receptors and membrane proteins.

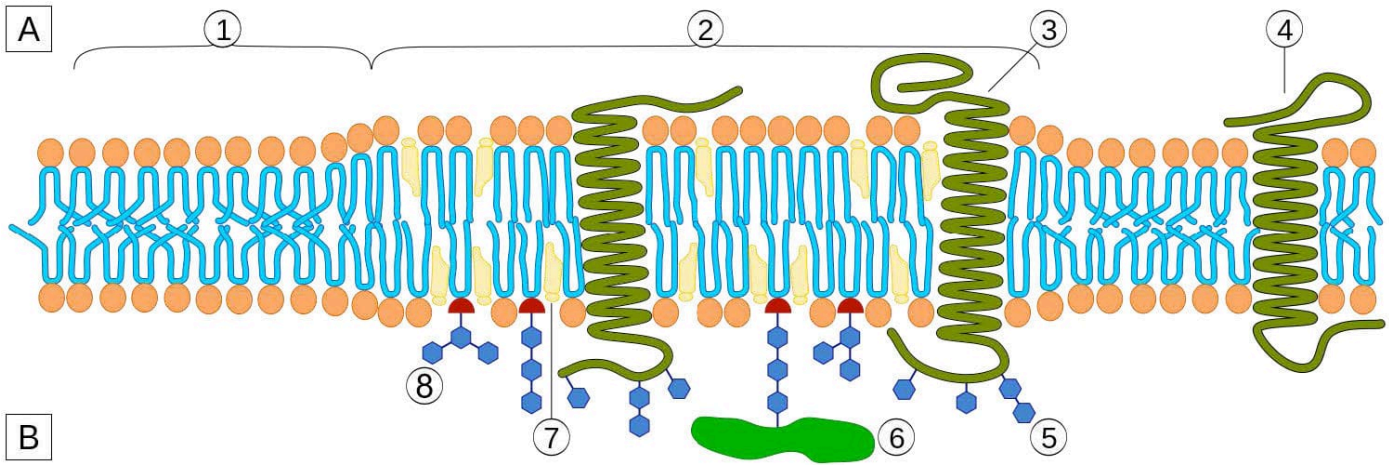


Figure 3.15 - A lipid raft - 1 = Non-raft membrane / 2 = Lipid raft / 3 = Lipid raft associated transmembrane protein / 4 = Non-raft membrane protein / 5 = Glycosylation modifications (on glycoproteins and glycolipids) / 6 = GPIanchored protein / 7 = Cholesterol / 8 = Glycolipid

Features

Distinguishing features of the rafts is that they are more ordered than the bilayers surrounding them, containing more saturated fatty acids (tighter packing and less disorganization, as a result) and up to 5 times as much cholesterol. They also are relatively rich in sphingolipids, with as much as 50% greater quantities of sphingomyelin than surrounding areas of the bilayer. The higher concentration of cholesterol in the rafts may be due to its greater ability to associate with sphingolipids (Figure 3.16). Some groups, such as prenylated proteins, like RAS, may be excluded from lipid rafts.

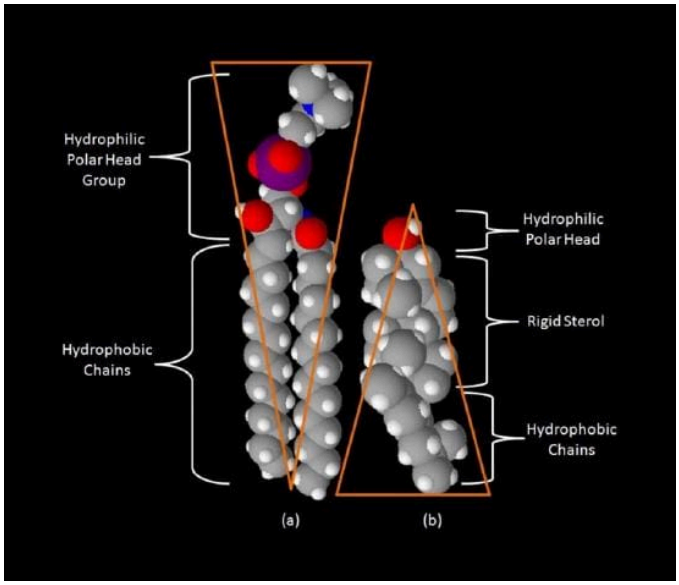


Figure 3.16 - A sphingolipid (left) associating with cholesterol (right)

Lipid rafts may provide concentrating platforms after individual protein receptors bind to ligands in signaling. After receptor activation takes place at a lipid raft, the signaling complex would provide protection from nonraft enzymes that could inactivate the signal. For example, a common feature of signaling systems is phosphorylation, so lipid rafts might provide protection against dephosphorylation by enzymes called phosphatases. Lipid rafts appear to be involved in many signal transduction processes, such as T cell antigen receptor signaling, B cell antigen receptor signaling, EGF receptor signaling, immunoglobulin E signaling, insulin receptor signaling and others. For more on signaling, see HERE.

Barrier

Transport of materials across membranes is essential for a cell to exist. The lipid bilayer is an effective barrier to the entry of most molecules and without a means of allowing food molecules to enter a cell, it would die. The primary molecules that move freely

across the lipid bilayer are small, uncharged ones, such as H2O, CO2, CO, and O2, so larger molecules, like glucose, that the cell needs for energy, would be effectively excluded if there were not proteins to facilitate its movement across the membrane.

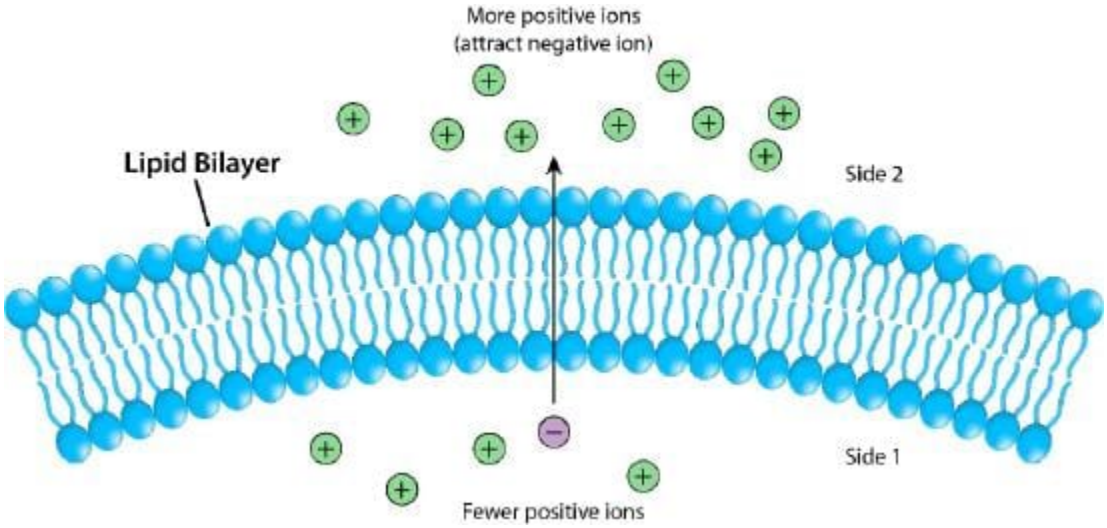


Figure 3.17 - The lipid bilayer as a barrier Image by Pehr Jacobson

Figure 3.17 depicts the barrier that the lipid bilayer provides to movement across it and the pressures (ionic attraction, in this case) that can affect movement. Potential energy from charge and concentration differences are harvested by cells for purposes that include synthesis of ATP, and moving materials against a concentration gradient in a process called active transport.

Membrane proteins

Proteins in a lipid bilayer can vary in quantity enormously, depending on the membrane. Protein content by weight of various membranes typically ranges between 30 and 75% by weight. Some mitochondrial membranes can have up to 90% protein. Proteins linked to and associated with membranes come in several types.

Transmembrane proteins

Transmembrane proteins are integral membrane proteins that completely span from one side of a biological membrane to the other and are firmly embedded in the membrane (Figure 3.18). Transmembrane proteins can function as docking sites for attachment (to the extracellular matrix, for example), as receptors in the cellular signaling system, or facilitate the specific transport of molecules into or out of the cell.

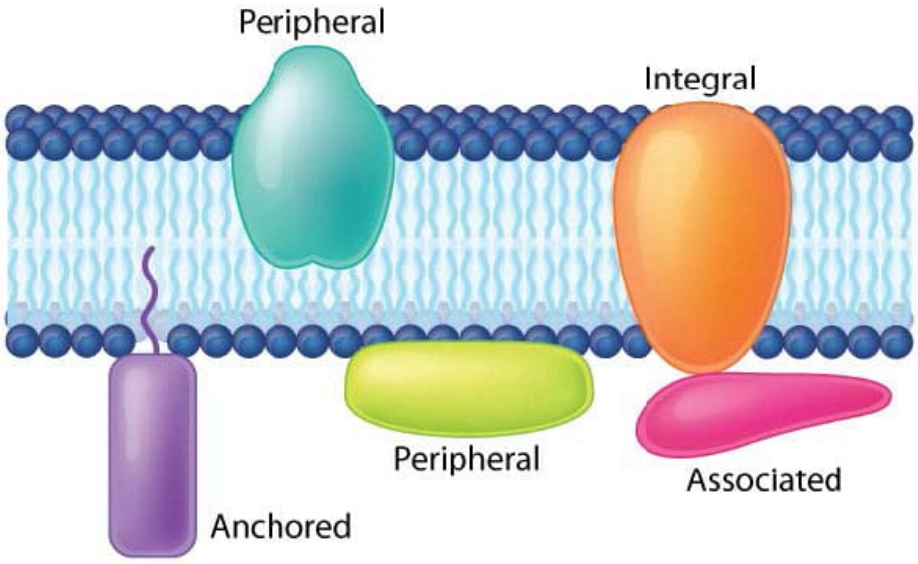


Figure 3.18 - Membrane protein types Image by Aleia Kim

Example of integrated/ transmembrane proteins include those involved in transport (e.g., Na+/K+ ATPase), ion channels (e.g., potassium channel of nerve cells) and signal transduction across the lipid bilayer (e.g., GProtein Coupled Receptors).

Peripheral membrane proteins interact with part of the bilayer (usually does not involve hydrophobic interactions), but do not project through it. A good example is phospholipase A2, which cleaves fatty acids from glycerophospholipids in membranes. Associated membrane proteins typically do not have external hydrophobic regions, so they cannot embed in a portion of the lipid bilayer, but are found near them. Such association may arise as a result of interaction with other proteins or molecules in the lipid bilayer. A good example is ribonuclease.

Anchored membrane proteins

Anchored membrane proteins are not themselves embedded in the lipid bilayer, but instead are attached to a molecule (typically a fatty acid) that is embedded in the membrane (Figure 3.19). The oncogene family of proteins known as ras are good examples. These proteins are anchored to the lipid bilayer by attachment to non-polar farnesyl groups catalyzed by the enzyme farnesyltransferase.

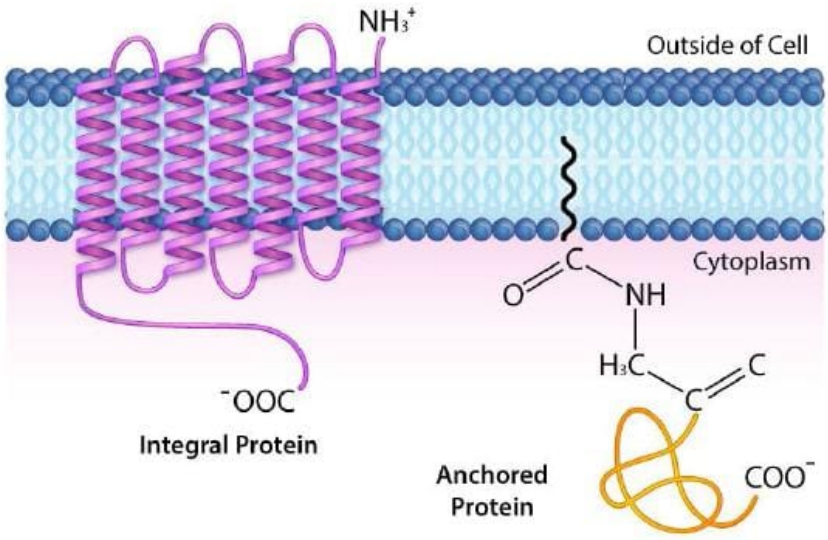


Figure 3.19 - Integral and anchored proteins Image by Aleia Kim

Finer classification

A more detailed classification scheme further categorizes the integral and anchored proteins into six different types (Figure 3.20). Type I and Type II have only one portion of the protein pass through the membrane. They differ in the orientation of the amine and carboxyl end with respect to inside/outside. Type I transmembrane proteins have the amino terminus on the outside and carboxy terminus on the inside, whereas Type II proteins have this reversed. Type III proteins are a single polypeptide chain that has multiple regions of it cross back and forth across the membrane, often to form a channel. Type IV is a multi-polypeptide protein which has multiple crossings of the membrane. Type V transmembrane proteins do not have a part of them that crosses the membrane, but they are anchored to the membrane by a lipid (such as a fatty acid) embedded in the lipid bilayer. Type VI transmembrane proteins both have a portion of them that crosses the membrane and they are attached to a lipid embedded in the lipid bilayer.

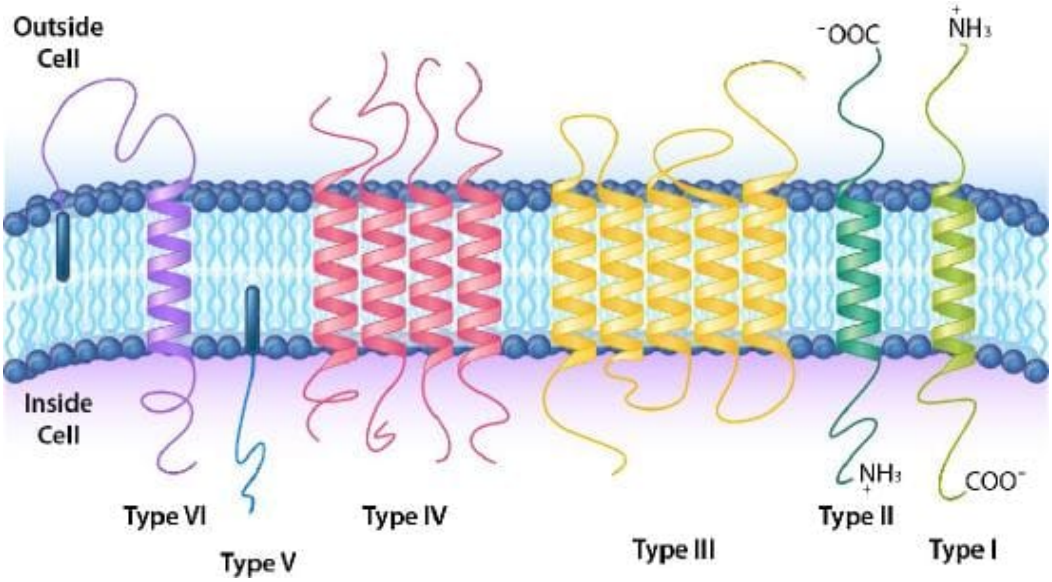


Figure 3.20 Types of integral and anchored membrane proteins Image by Aleia Kim

Blood Types

Cells have hundreds-thousands of membrane proteins and the protein composition of a membrane varies with its function and location. Glycoproteins embedded in membranes play important roles in cellular identification. Blood types, for example, differ from each other in the structure of the carbohydrate chains projecting out from the surface of the glycoprotein in their membranes (Figure 3.21).

	Group A	Group B	Group AB	Group O
Red blood cell type	A	B	AB	O
Antibodies in Plasma	Anti-B	Anti-A	None	Anti-A and Anti-B
Antigens in Red Blood Cell	A antigen	B antigen	A and B antigens	None

Figure 3.21

Osmotic Pressure

Membranes provide barriers/boundaries for most molecules, but the permeability of water across a lipid bilayer creates a variable that must be considered. The variable here is osmotic pressure. Osmotic pressure (loosely) refers to the tendency of a solution to take in water by the process of osmosis. In Figure 3.22, one can see a visual representation of the concept of the pressure.

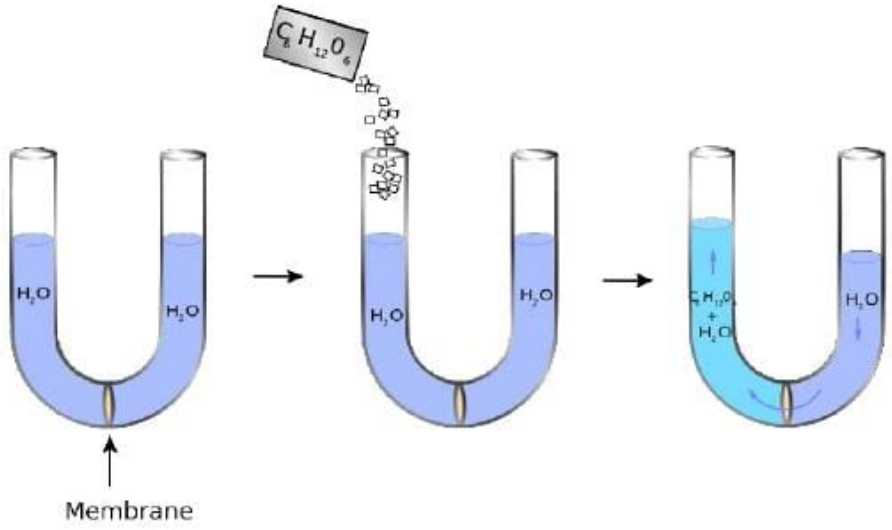


Figure 3.22

A U-shaped tube has at its bottom a semipermeable membrane. Water can pass through the membrane, but sugar molecules (C6H12O6) cannot. On the left side, sugar is added creating a concentration difference between the right and left chambers. Water diffuses across the membrane from right to left in an attempt to equalize the concentrations, causing the level of the right side to decrease and the left side to increase. The pressure resulting from the differences in height is felt at the membrane.

Equalizing concentrations

The liquid on the right does not completely move to the left, though, as might be expected if the only force involved is equalizing the concentration of sugar across the membrane (no sugar on right = no water). Instead, an equilibrium of sorts of water levels is reached even though the concentrations don't equal out. The pressure existing at the membrane then from the differences in level corresponds to the osmotic pressure of the mixture. The osmotic pressure of a solution is the pressure difference needed to halt the flow of solvent across a semipermeable membrane. Osmotic pressure can also be thought of as the pressure required to counter osmosis. The osmotic pres- Figure 3.21 - Blood types arise from cell surface glycoproteins Figure 3.22 - Osmotic pressure. Water diffuses leftwards to try to equalize the solute concentration. The pressure realized at the membrane in the right figure is the osmotic pressure sure of a dilute solution mathematically behaves like the ideal gas law

$$P_{osmotic} = \frac{nRT}{V} \tag{3.1.1}$$

where n is the number of moles, R is the gas constant, T is the temperature in Kelvin, and V is the volume.

It is more convenient in solutions to work with molarity, so

$$P_{osmotic} = MR^*T \tag{3.1.2}$$

where M is the molarity of the dissolved molecules, R* is the gas constant expressed in (L atm)/(K mol), and T is the temperature. The Greek letter Π is used to refer to the $P_{osmotic}$term, so

$$\Pi = MR^*T \tag{3.1.3}$$

Remember when calculating the molarity to include the molarity of each particle. For example, when one dissolves sucrose in solution, it does not split into smaller particles, so

$$Molarity_{Particles} = Molarity_{Sucrose} \tag{3.1.4}$$

However, for salts, like KOH, which forms two ions in solution (K+ and OH-),

$Molarity_{Particles}$= 2* $Molarity_{KOH}$.

Significant consideration

Osmotic pressure is a significant consideration for cells. Consider the fact that water can move freely across cellular membranes, but most of the contents of the cell, such as proteins, DNA, ions, sugars, etc., cannot. Second, the concentration of these compounds inside the cell is different than the concentration of them outside of the cell. Third, since water can move through the lipid bilayer, it will tend to move in the direction that will tend to equalize solute.

Hypotonic, hypertonic, isotonic

We consider three situations (Figure 3.23). First, if the concentration of solutes is greater inside the cell than outside, water will tend to move into the cell, causing the cell to swell. This circumstance is called hypotonic. Conversely, if the solute concentration is greater outside the cell than inside of it, water will exit the cell and the cell with shrink. This is a hypertonic situation. Last, if the concentrations of solutes into and outside of the cell is equal, this is called an isotonic solution. Here, no movement of water occurs across the cell membrane and the cell retains its size.

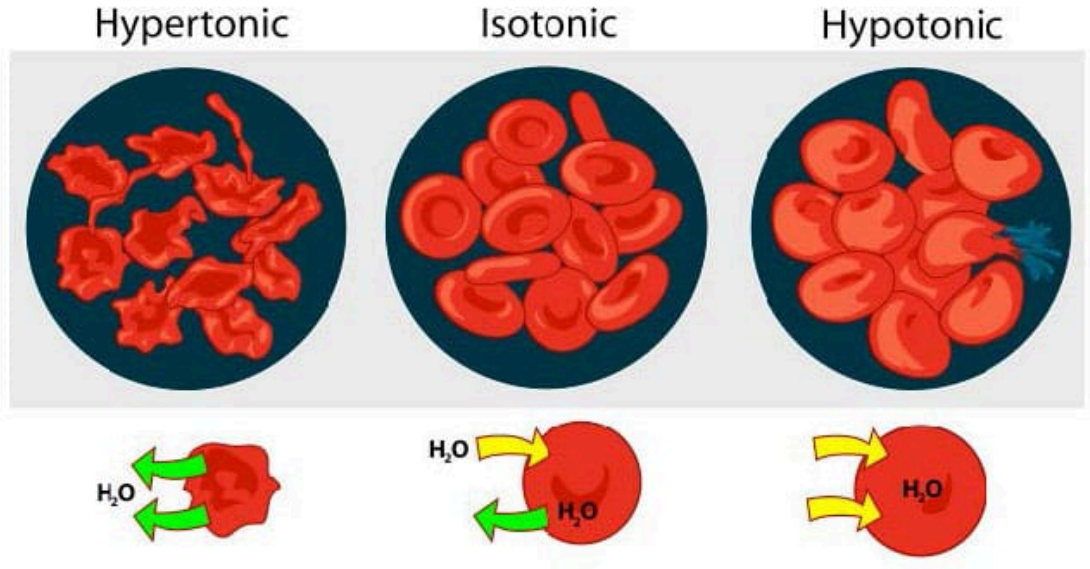
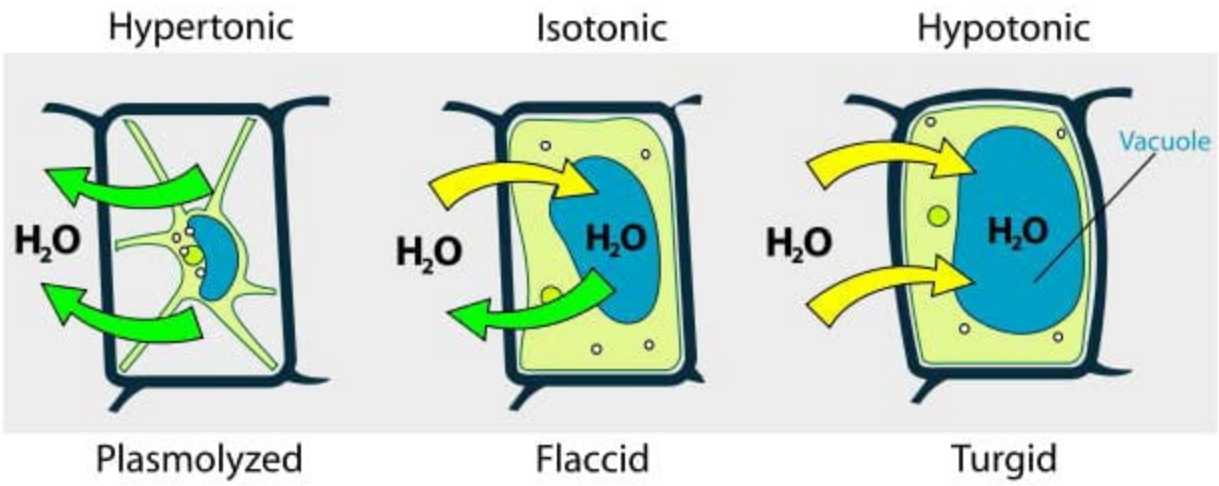


Figure 3.23 - The effect of three different osmotic conditions on red blood cells concentrations on either side of the membrane

If the osmotic pressure is greater than the forces holding together a cellular membrane, the cell will rupture. Because of this, some cells have built in defenses to prevent problems. Plant cells, for example have a fairly rigid cell wall that resists expansion in hypotonic solutions (Figure 3.24). Bacteria also have a cell wall that provides protection.



This page titled 3.1: Basic Concepts in Membranes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

3.2: Transport in Membranes

Source: BiochemFFA_3_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Movement of materials across membranes

As noted earlier, it is essential for cells to be able to uptake nutrients. This function along with movement of ions and other substances is provided by proteins/protein complexes that are highly specific for the compounds they move.

Selective movement of ions by membrane proteins and the ions' extremely low permeability across the lipid bilayer are important for helping to maintain the osmotic balance of the cell and also for providing for the most important mechanism for it to make ATP - the process of oxidative phosphorylation.

Terminology

A protein involved in moving only one molecule across a membrane is called a uniport (Figure 3.25). Proteins that move two molecules in the same direction across the membrane are called symports (also called synporters, synports, or symporters). If two molecules are moved in opposite directions across the bilayer, the protein is called an antiport. Proteins involved in moving ions are called ionophores.

If the action of a protein in moving ions across a membrane results in a net change in charge, the protein is described as electrogenic and if there is no change in charge the protein is described as electroneutral (Figure 3.26). When the driving force for movement through the membrane protein is simply diffusion, the process is called facilitated diffusion or passive transport and when the process requires other energy input, the process is called active transport.

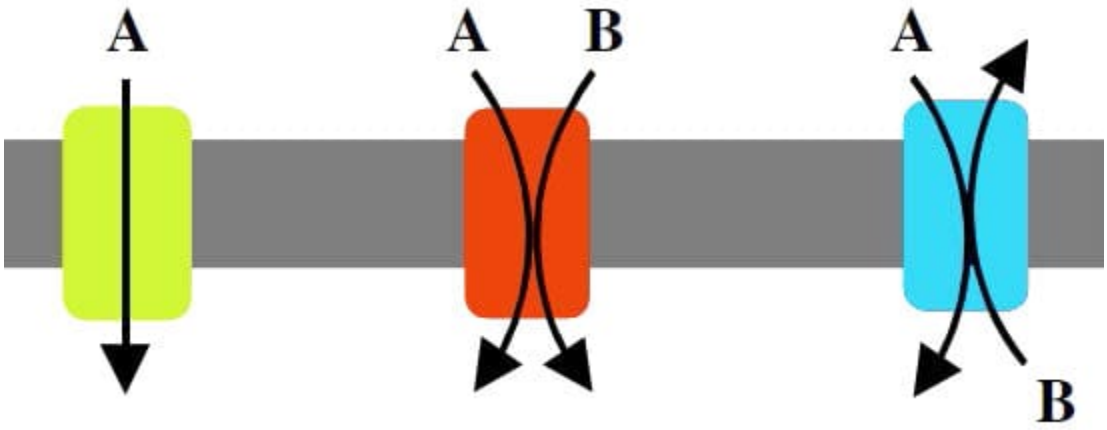


Figure 3.25 - A uniport, a symport, and an antiport

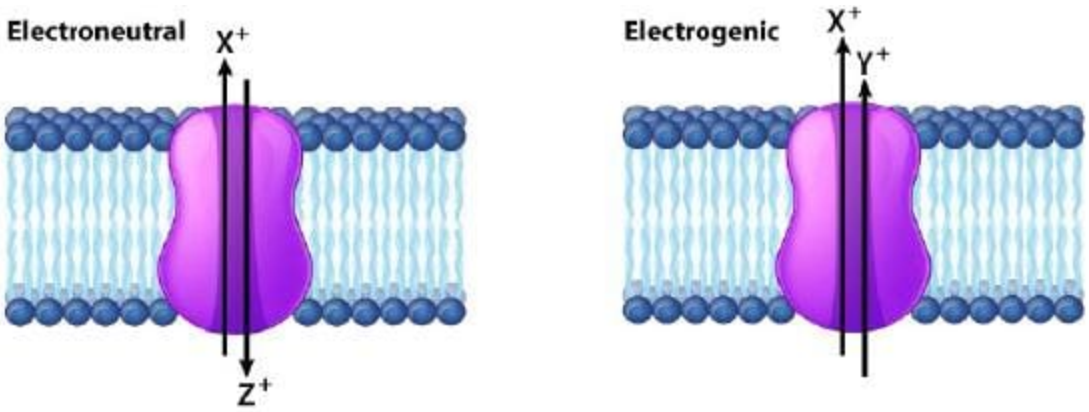


Figure 3.26 - Electroneutral and electrogenic transporters - Image by Aleia Kim

Channels and transporters

With respect to movement of materials through membrane proteins, there is a difference between channels (sometimes called pores) and transporters. Channels largely provide openings with some specificity and molecules pass through them at close to the rate of diffusion. They usually involve movement of water or ions. Examples would be the sodium or potassium channels of nerve cells. Transporters have high specificity and transfer rates that are orders of magnitude slower. Transport proteins include the sodium-potassium pump, the sodium-calcium exchanger, and lactose permease, amongst many others).

Facilitated diffusion

As noted, the driving force for facilitated diffusion is concentration, meaning that in facilitated diffusion, materials will only move from a higher concentration to a lower concentration and that at the end of the process, the concentration of materials on each side of a bilayer will be equal (Figure 3.28). This may work well in many cases.

For example, the blood concentration of glucose is sufficiently high that red blood cells can use facilitated diffusion as a means of acquiring glucose. Other cells, further removed from the blood supply where the glucose concentration is lower, must use active transport mechanisms because there is not a sufficient concentration of glucose to provide cells with the glucose they need.

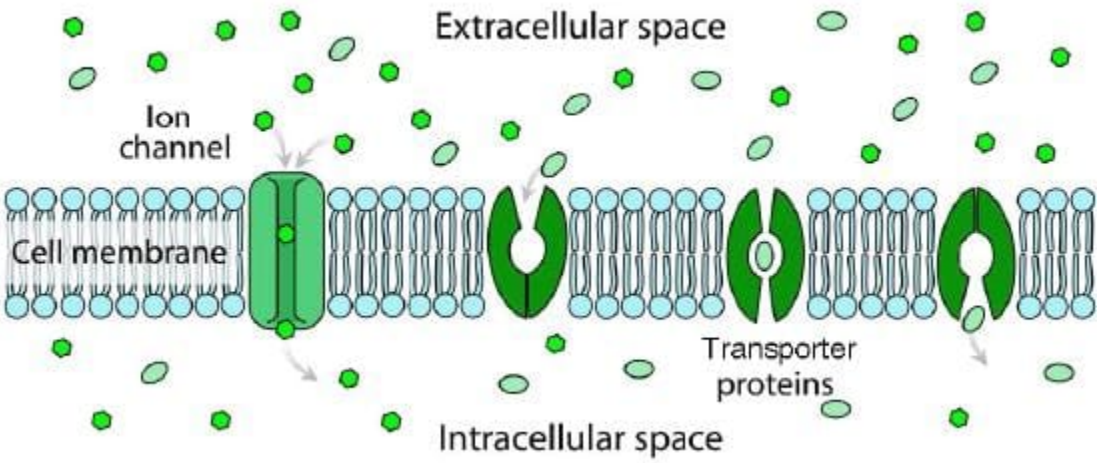


Figure 3.27 - Ion channel and transporter proteins

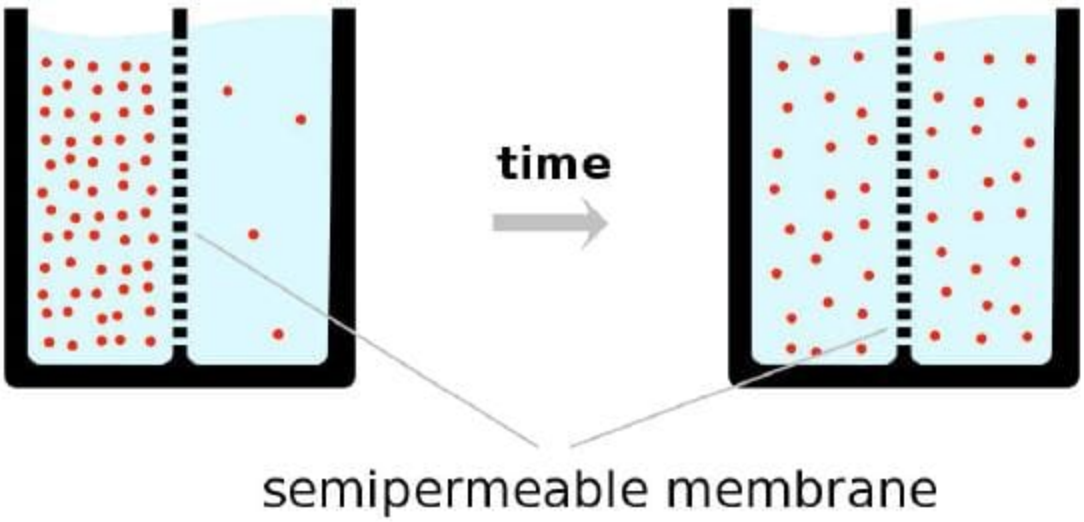


Figure 3.28 - In diffusion, solutes move from high concentration to lower concentration and eventually equalize

Ion channels

Ion channels are pore-forming membrane proteins in the membranes of all cells that regulate movement of selected ions across a membrane (Figures 3.29 & 3.30). They help to establish the resting membrane potential and to affect action potentials and other electrical signals. They are very important in the process of nerve transmission. Ion channels control the flow of ions across secretory and epithelial cells, and consequently help to regulate cell volume by affecting osmotic pressure.

Ion channels are essential features of almost all cells, functioning as selective “tunnels” that restrict movement through them to ions with specific characteristics (typically size). The size of the opening is very narrow (usually one or two atoms wide) and is able to select even against ions that are too small.

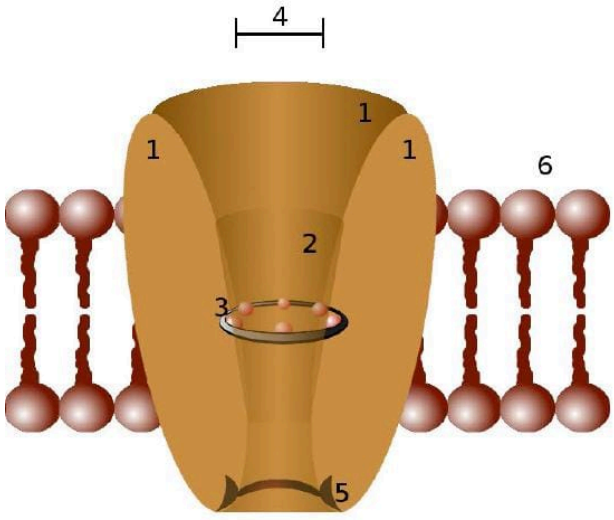


Figure 3.29 - Schematic structure of an ion channel protein - 1 = General structure / 2 = Entrance / 3 = Selectivity filter / 4 = Effective diameter / 5 = Controllable opening

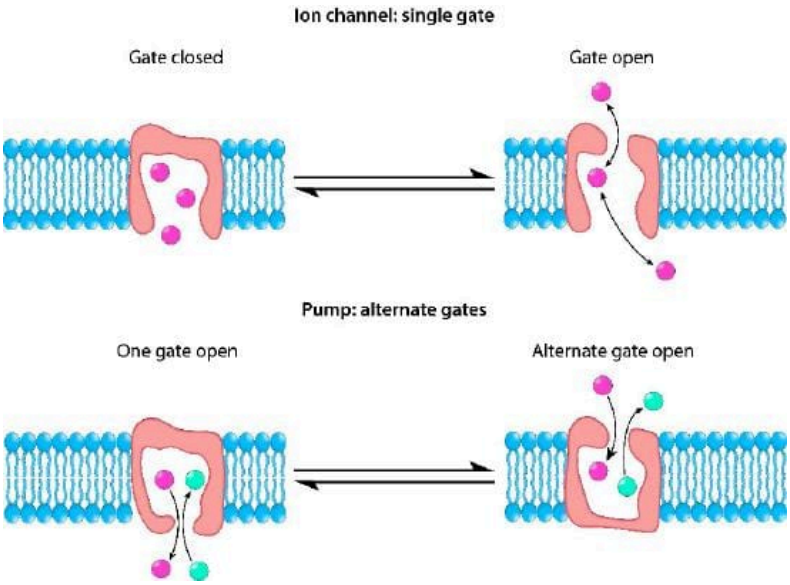


Figure 3.30 - Two types of ion channel - single and double gates - Image by Pehr Jacobson

Control mechanisms

Ion channels are controlled by mechanisms that include voltage, ligands, light, temperature, and mechanical deformation (stretch activated). Ligand-gated ion channels (LGICs) are transmembrane proteins which open to selectively allow ions such as Na+, K+, Ca++, or Cl– to pass through the membrane in response to the binding of a ligand messenger.

Sound waves cause mechanical deformation of hair cells in the ear. This results in the opening of ion channels and initiation of a nerve signal to the brain.

Sodium ion channels in the tongue for sugar receptors open in response to binding of sucrose, allowing sodium concentration in the nerve cell to increase and initiate a nerve signal to the brain. In this case, the default for the gate is to be closed and it opens in response to binding of a ligand (sucrose).

In light sensing cells of the eye, calcium gates are open by default, but stimulation by light causes them to close, triggering a series of events that result in a signal being sent the brain about the perception of light. Thus, in this case, the stimulus (light) causes an open channel to close.

Moving the other direction, nerve signals originating in the brain travel to muscle tissue and through a complicated set of exchanges, result in the opening of calcium gates of muscle cells, increasing the concentration of calcium and stimulating muscular contraction (see HERE).

Voltage gated channels are essential for transmission of nerve signals, a process discussed in more depth HERE.

Ion movement through channels

The ability of ion channels to select against ions too large is intuitive - the size of the opening in the ion channel simply isn't big enough for a larger ion to fit through the opening. Potassium, for example, passes through sodium channels rarely because the opening is too small.

Potassium channels that are selective for potassium ions must be big enough to allow potassium to enter, but if size were the only selection means, then sodium ions would also readily pass through potassium channels, since sodium ions (0.95 Å) are smaller than potassium ions (1.33 Å). In order for potassium channels to select against sodium ions and favor potassium ions, other considerations come into play.

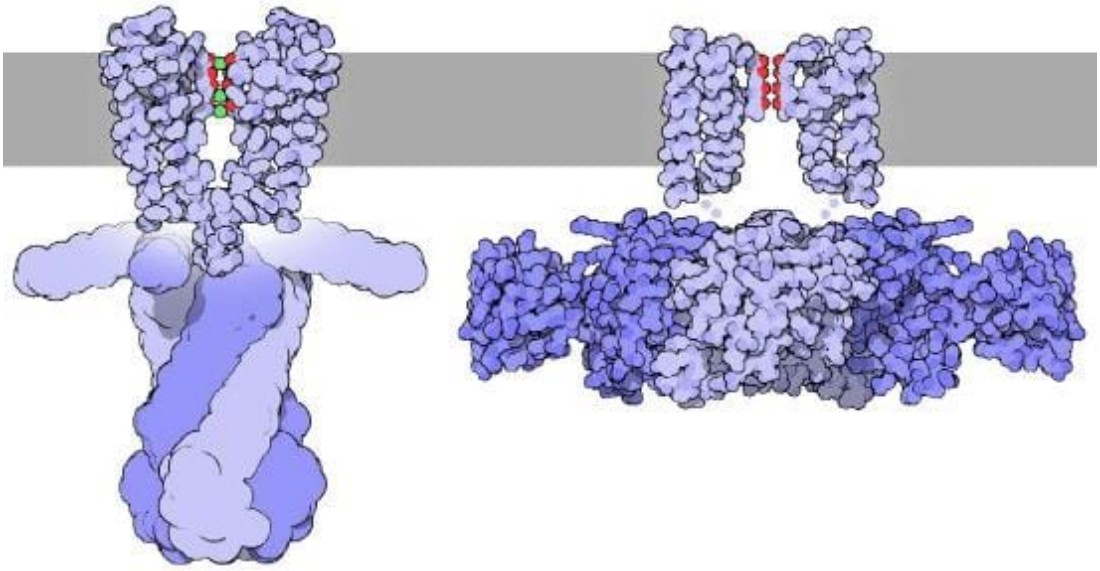


Figure 3.31 - A potassium channel closed (left) and open (right) - David Goodsell - "Molecule of the Month: Potassium Channels". (from Wikipedia)

Hydration shell

To understand this unique selectivity, it is important to understand how ions move through channels. Before an ion can pass through a channel, it must first be dissociated from (stripped of) the water molecules in its hydration shell - water molecules surrounding ions in aqueous solutions (Figure 3.32). This process requires an input of energy. The initial energy required to strip the water molecules from the hydration shell has been compared to the activation energy of an enzymatic reaction.

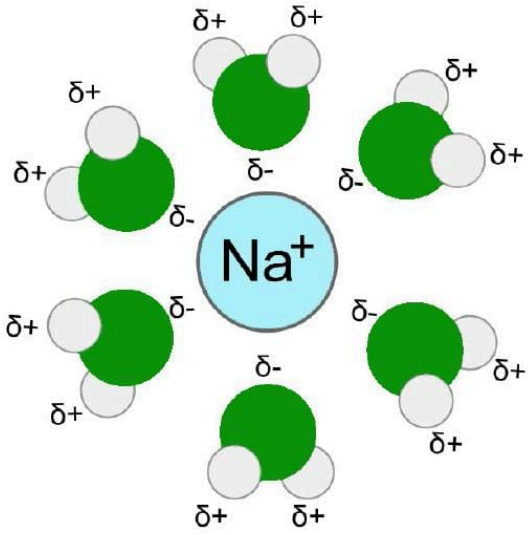


Figure 3.32 - Sodium surrounded by water molecules in a hydration shell

Comparable to enzymes

Just as enzymes lower the activation energy of enzymatic reactions and thus allow them to more readily occur, so too do channel proteins lower the energy requirements for a molecule to traverse a lipid bilayer. In the absence of the channel protein, the dehydration energy is mostly prohibitive for most polar molecules to occur, so very few make it across the lipid bilayer without the channel protein. This is why ion channel/transport proteins are so important to the cell.

After the water has been stripped, the ion can pass through the channel and when it arrives at the other side of the channel, the diffusing ion becomes rehydrated, thus regaining the energy that was required initially to strip away the water molecules from the

ion.

Selectivity of the potassium channel

The potassium channel (Figure 3.33) uses the dimensions of the potassium ion precisely to shepherd it through the channel. The sodium ion, which has different dimensions has a more difficult time making it through the channel despite its smaller size. The reason this is rooted in the energy required for dehydration.

For potassium ions, after the water has been stripped off, precisely positioned carbonyl groups along the channel help to stabilize the ion as it moves. The sodium ion, on the other hand is too small and does not make efficient connections with carbonyl groups and thus has a more difficult path. Because of this, the energy difference between dehydration and rehydration of a sodium ion in a potassium channel is energetically unfavorable (requires net input of energy) but the same process for a potassium ion is energetically favorable (results in a net gain of energy).

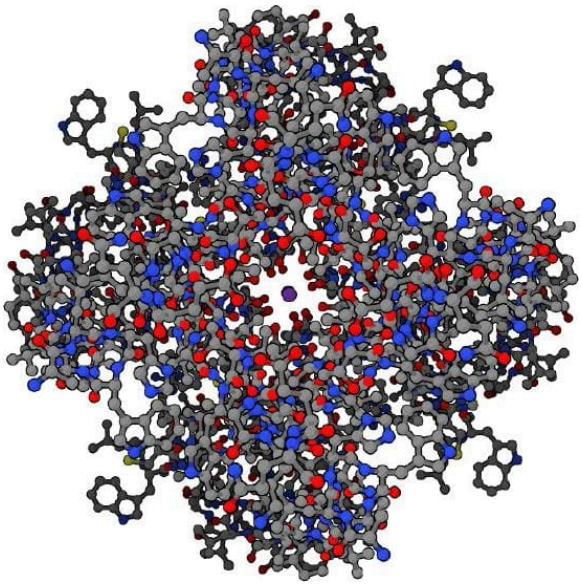
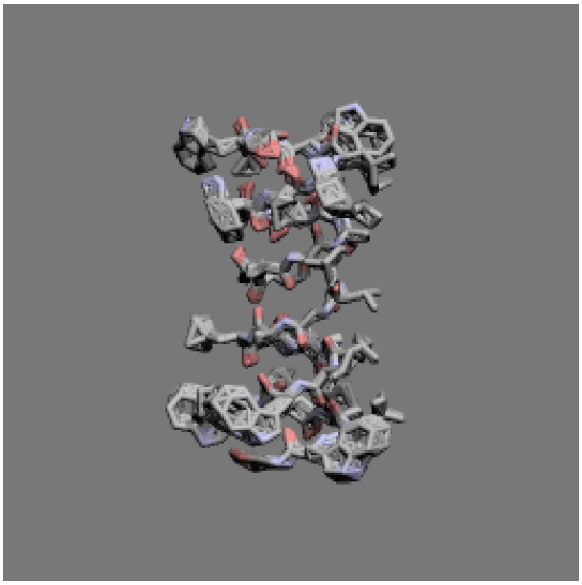


Figure 3.33 - A potassium ion (center) transiting the potassium channel - David Goodsell



Movie 3.1 - Gramicidin A Wikipedia (animated gif, download to view)

Energy factor

Thus the selection in favor of potassium and against sodium ions in a potassium channel is based on energy, not physical size, whereas in the selection of sodium ions over potassium ions in a sodium channel, size is the primary consideration.

Ion balance

The movement of ions across a lipid bilayer is tightly regulated, and with good reason. Maintaining a proper balance of ions inside and outside of cells is important for maintaining osmotic balance. It is also important inside and outside of organelles like the mitochondria and chloroplasts for energy generation. If the ionic balance of a cell is sufficiently disturbed by an uncontrolled ionophore, a cell may die.

Gramicidin

Gramicidins (Movie 3.1) are antibiotic polypeptides synthesized by the soil bacterium known as Bacillus brevis. These small pentadecapeptides (15 amino acids) are synthesized by the bacterium to kill other bacteria.

When released by the Bacillus brevis, the gramicidins insert themselves in the membranes of Gram positive bacteria and allow the movement of sodium ions into the target cells, ultimately killing them. Gramicidins can also cause hemolysis in humans so they cannot be used internally, but instead are used topically.

Aquaporins

Aquaporins are pore-containing integral membrane proteins that selectively permit passage of water molecules in and out of the cell, while preventing ions and other solutes from moving (Figures 3.34 & 3.35). Some aquaporins called aquaglyceroporins, also transport other small uncharged entities, such as glycerol, ammonia, urea, and CO2, across the membrane,. The water pores are completely impermeable to charged molecules, such as protons, which is important for the preserving the membrane's electrochemical potential difference.

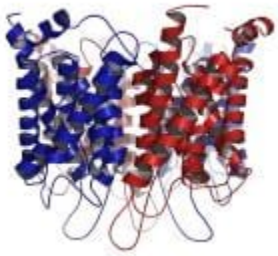


Figure 3.34 - View of aquaporin from the side - Wikipedia

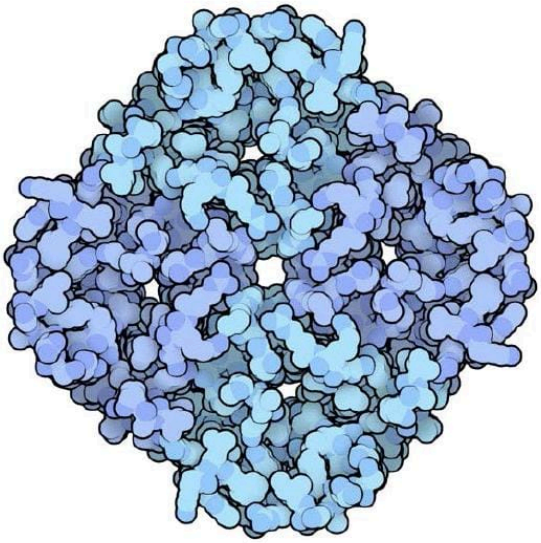


Figure 3.35 - View of aquaporin from the top - David Goodsell

Porins

Porins are proteins containing a β-barrel structure that crosses the cell membrane/wall and acts as a pore/channel through which specific molecules diffuse. Porins are found in the outer membrane of Gram-negative bacteria and some Gram-positive bacteria, mitochondria, and chloroplasts.

Porins typically transport only one group of molecules or, in some cases, one specific molecule. Antibiotics, such as β-lactam and fluoroquinolone pass through porins to reach the cytosol of Gram negative bacteria. Bacteria may develop resistance to these antibiotics when a mutation occurs to the porin involved that results in exclusion of the antibiotics that would otherwise pass through.

Transporter proteins

Not all facilitated transport occurs through ion channel proteins. Transporter proteins, as noted earlier (HERE and Figure 3.27) facilitate movement of materials across a lipid bilayer, but are slower than ion channels. Figure 3.36 illustrates a transporter protein in action. As can be seen, transporter proteins rely on a specific receptor site for proper recognition of the molecule to be moved.

Binding of the proper molecule causes a conformational change in the shape of the protein (an eversion) which results in a flipping of the open side of the protein to the other side of the lipid bilayer. In this way, the molecule is moved. Like ion channels, transporter proteins facilitate movement of materials in either direction, driven only by the concentration difference between one side and the other.

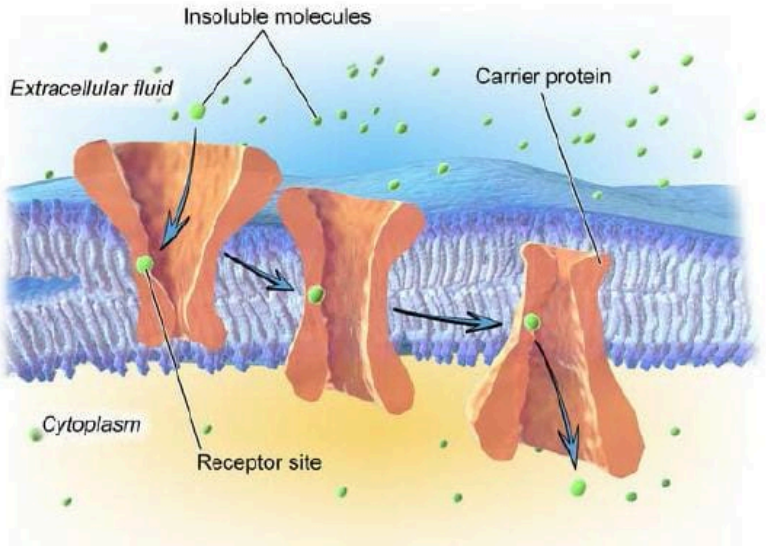


Figure 3.36 - Facilitated diffusion with a transporter protein - Wikipedia

Active transport

All of the transport mechanisms described so far are driven solely by a concentration gradient - moving from higher concentrations in the direction of lower concentrations. These movements can occur in either direction and, as noted, result in equal concentrations on either side of the bilayer, if allowed to go to completion. Many times, however, cells must move materials against a concentration gradient and when this occurs, another source of energy is required. This process is known as active transport.

A good definition of active transport is that in active transport, at least one molecule is being moved against a concentration gradient. A common, but not exclusive, energy source is ATP (see Na+/K+ ATPase), but other energy sources are also employed. For example, the sodium-glucose transporter uses a sodium gradient as a force for actively transporting glucose into a cell. Thus, it is important to know that not all active transport uses ATP energy.

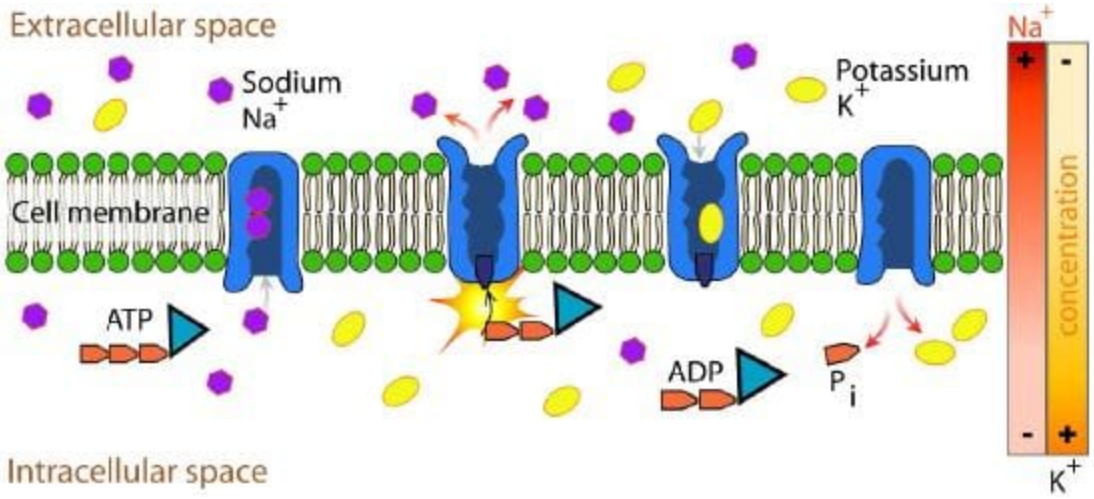


Figure 3.37 - An overview of active transport by the Na+K+ ATPase

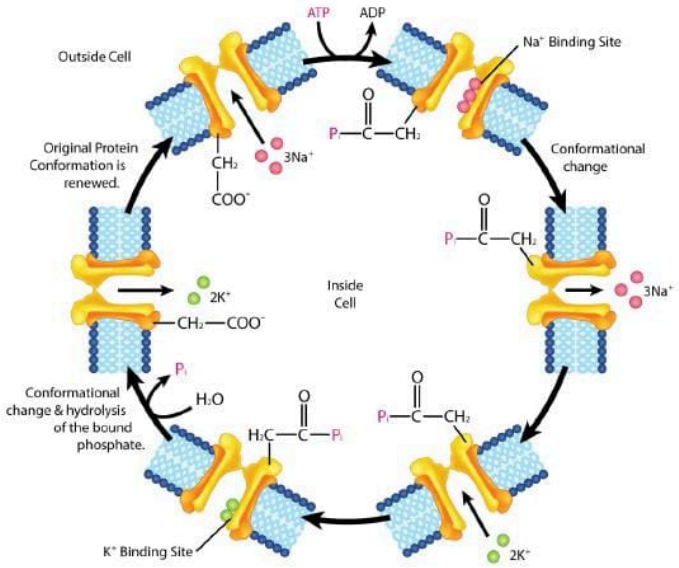


Figure 3.38 - Sequential steps in the active transport of ions by theNa+K+ ATPase - Wikipedia

Na+/K+ ATPase

An important integral membrane transport protein is the Na+/K+ ATPase antiport (Figures 3.37 and 3.38), which moves three sodium ions out of the cell and two potassium ions into the cell with each cycle of action. In each case, the movement of ions is against the concentration gradient. Since three positive charges are moved out for each two positive charges moved in, the system is electrogenic.

The protein uses the energy of ATP to create ion gradients that are important both in maintaining cellular osmotic pressure and (in nerve cells) for creating the sodium and potassium gradients necessary for signal transmission. Failure of the system to function results in swelling of the cell due to movement of water into the cell through osmotic pressure. The transporter expends about one fifth of the ATP energy of animal cells. The cycle of action occurs as follows:

1. Pump binds ATP followed by binding of 3 Na+ ions from cytoplasm of cell
2. ATP hydrolysis results in phosphorylation of aspartate residue of pump. ADP is released
3. Phosphorylated pump undergoes conformational change to expose Na+ ions to exterior of cell. Na+ ions are released.
4. Pump binds 2 extracellular K+ ions.
5. Pump dephosphorylates causing it to expose K+ ions to cytoplasm as pump returns to original shape.
6. Pump binds 3 Na+ ions, binds ATP and releases 2 K+ ions to restart process

The Na+/K+ ATPase is classified as a P-type ATPase. This category of pump is notable for having a phosphorylated aspartate intermediate and is present across the biological kingdoms - bacteria, archaeans, and eukaryotes.

ATPase types

ATPases have roles in either the synthesis or hydrolysis of ATP and come in several different forms.

- F-ATPases (F1FO-ATPases) are present in mitochondria, chloroplasts and bacterial plasma membranes and are the prime ATP synthesizers for these systems. Each uses a proton gradient as its energy source for ATP production. Complex V of the mitochondrion is an F-type ATPase.
- V-ATPases (V1VO-ATPases) are mostly found in vacuoles of eukaryotes . They utilize energy from ATP hydrolysis to transport solutes and protons into vacuoles and lysosomes, thus lowering their pH values.

The V-type and F-type ATPases are very similar in structure. The V-type (Figure 3.39) uses ATP to pump protons into vacuoles and lysosomes, whereas F-types use proton gradients of the mitochondria and chloroplasts to make ATP.

- A-ATPases (A1AO-ATPases) are found in archaeans and are similar to F-ATPases in function.
- P-ATPases (E1E2-ATPases) are in bacteria, fungi and in eukaryotic plasma membranes and organelles. They transport a diversity of ions across membranes. Each has a common mechanism of action which include autophosphorylation of a conserved aspartic acid side chain within it. Examples of P-type ATPases include the Na+/K+ ATPase and the calcium pump.
- E-ATPases are enzymes found on the cell surface. They hydrolyze a range of extracellular nucleoside triphosphates, including ATP.

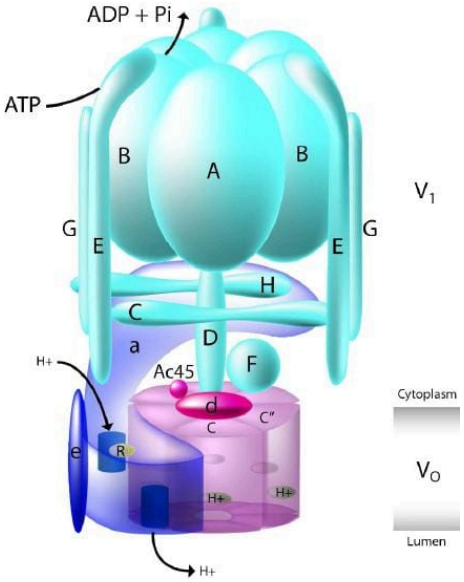


Figure 3.39 - Schematic structure of V-ATPases- Wikipedia

Nerve transmission

Now that you have seen how the Na+/K+ ATPase functions, it is appropriate to discuss how nerve cells use ion gradients created with it to generate and transmit nerve signals. Neurons are cells of the nervous system that use chemical and electrical signals to rapidly transmit information across the body (Figure 3.40). The sensory nerve system links receptors for vision, hearing, touch, taste, and smell to the brain for perception. Motor neurons run from the spinal cord to muscle cells. These neurons have a cell body and a very long, thin extension called an axon, that stretches from the cell body in the spinal cord all the way to the muscles they control. Nerve impulses travel down the axon to stimulate muscle contraction.

Signals travel through neurons, ultimately arriving at junctions with other nerve cells or target cells such as muscle cells. Note that neurons do not make physical contact with each other or with muscle cells. The tiny space between two neurons or between a neuron and a muscle cell is called the synaptic cleft. At the synaptic cleft, the neuron releases neurotransmitters that exit the nerve cell and travel across the junction to a recipient cell where a response is generated. That response may be creating another nerve signal, if the adjacent cell is a nerve cell or it may be a muscular contraction if the recipient is a muscle cell (Figure 3.41).

In considering information movement via nerve cells, then, we will discuss two steps - 1) creation and propagation of a signal in a nerve cell and 2) action of neurotransmitters exiting a nerve cell and transiting a synaptic junction.

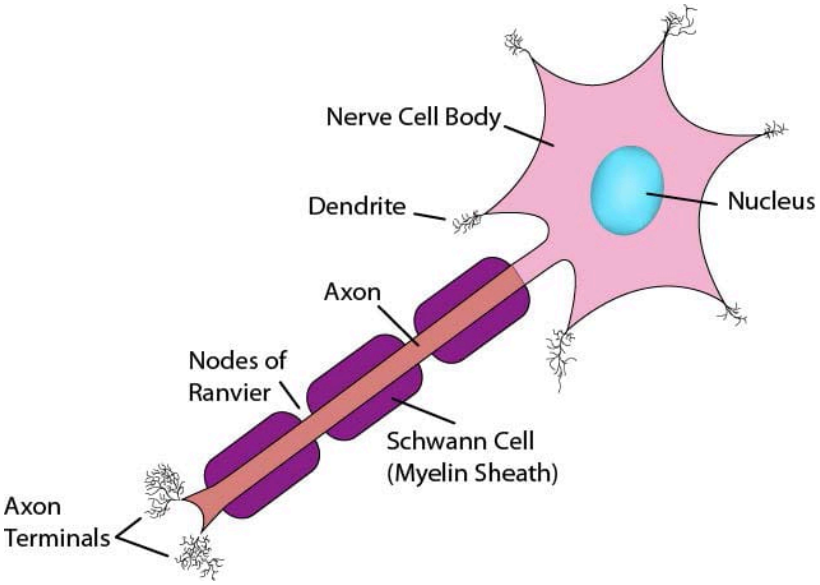


Figure 3.40 - Anatomy of a nerve cell - Image by Pehr Jacobson

Signal source

Creation of a nerve signal begins with a stimulus to the nerve cell. In the case of muscle contraction, the motor cortex of the brain sends signals to the appropriate motor neurons, stimulating them to generate a nerve impulse. How is such an impulse generated?

Resting potential

In the unstimulated state, all cells, including nerve cells, have a small voltage difference (called the resting potential) across the plasma membrane, arising from unequal pumping of ions across the membrane. The Na+/K+ ATPase, for example, pumps sodium ions out of the cell and potassium ions into cells. Since three sodium ions get pumped out for every two potassium ions pumped in, a charge and chemical gradient is created. It is the charge gradient that gives rise to the resting potential.

Altering the gradients of ions across membranes provide the driving force for nerve signals. This happens as a result of opening and closing of gated ion channels. Opening of gates to allow ions to pass through the membrane swiftly changes the ionic balance across the membrane resulting in a new voltage difference called the action potential. It is the action potential that is the impetus of nerve transmission.

Initiation of signal

The signal generated by a motor neuron begins with opening of sodium channels in the membrane of the nerve cell body causing a rapid influx of sodium ions into the nerve cell. This step, called depolarization (Figure 3.42), triggers an electrochemical signal - the action potential. Remember that the Na+/K+ ATPase has created a large sodium gradient, so sodium ions rush into the cell when sodium channels open. After the initial depolarization, potassium channel gates, responding to the depolarization, open, allowing potassium ions to rapidly diffuse out of the cell (remember K+ ions are more abundant inside of the cell). This phase is called the repolarization phase and during it, the sodium gates close.

The rapid exit of potassium ions causes the voltage difference to “overshoot” the resting potential and potassium gates close. This followed by the so-called refractory period, when the Na+/K+ ATPase begins its work to re-establish the original conditions by pumping sodium ions out and potassium ions into the nerve cell. Eventually, the system recovers and the resting potential is re-established. The initiating end of the nerve cell is then ready for another signal.

Propagation of action potential

What we have described here is only the initiation of the nerve signal in one part of the nerve cell. For the signal to be received,

the action potential must travel the entirety of the length of the nerve cell (the axon) and cause a chemical signal to be released into the synaptic cleft to get to its target. Propagating the nerve signal (action potential) in the original nerve cell is the function of all of the rest of the gated ion channels (Figure 3.43) positioned on the sides of the nerve cell. The sodium and potassium gates involved in propagation of the signal all act in response to voltage changes created by the electrochemical gradient moving down the nerve cell (Figure 3.44). Remember that opening of the initial gates at initiation of the signal created an influx of sodium ions and an efflux of potassium ions.

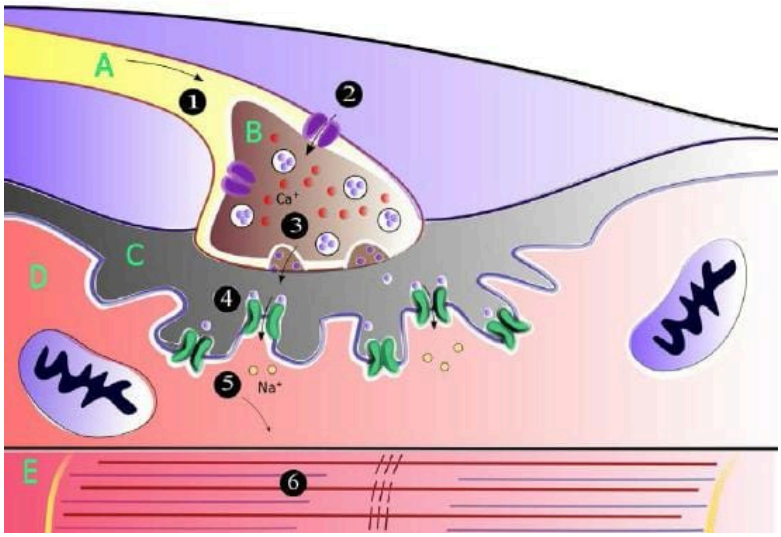


Figure 3.41 - (1) The action potential reaches the axon terminal; (2) voltage-dependent calcium gates open - calcium enters the axon terminal; (3) neurotransmitter vesicles fuse with the presynaptic membrane and release acetylcholine (ACh); (4) ACh binds to postsynaptic receptors on the sarcolemma; (5) ion channels open in response and allow sodium ions to flow into the muscle cell; (6) The flow of sodium ions into the muscle cell generates an action potential which travels to the myofibril and results in muscle contraction. - Wikipedia

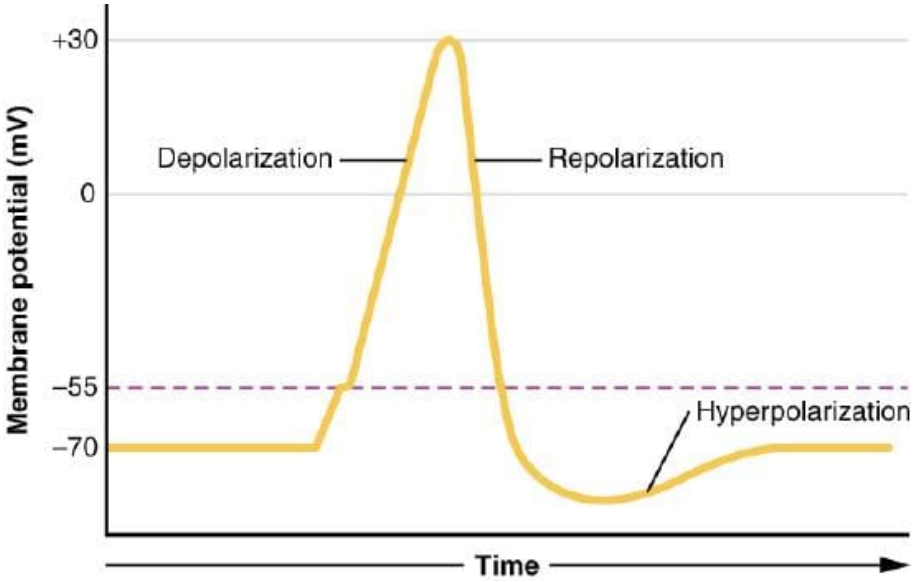


Figure 3.42 - Depolarization and repolarization of a nerve cell - Wikipedia

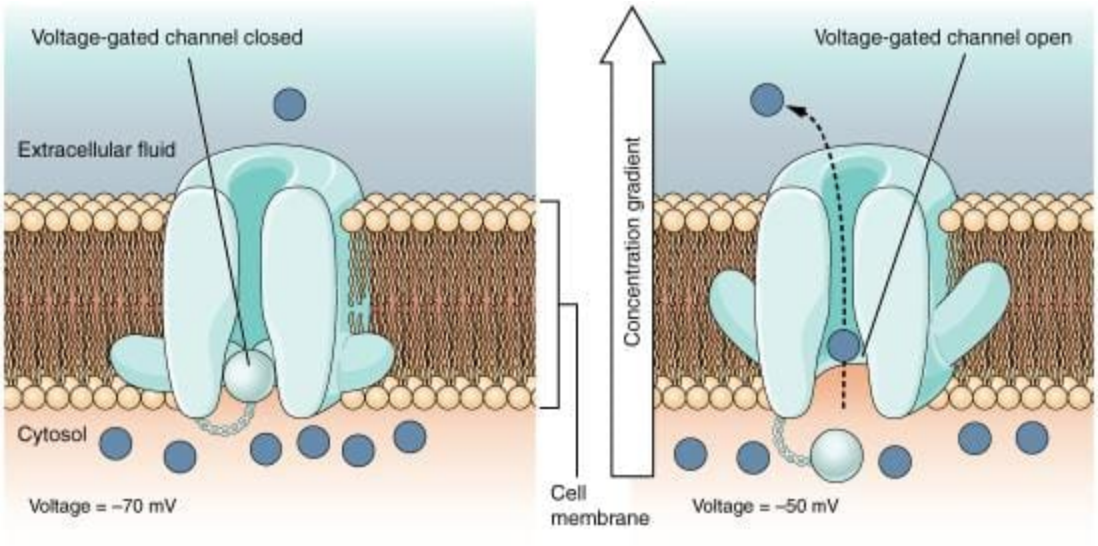


Figure 3.43 - Voltage gated ion channels - Wikipedia

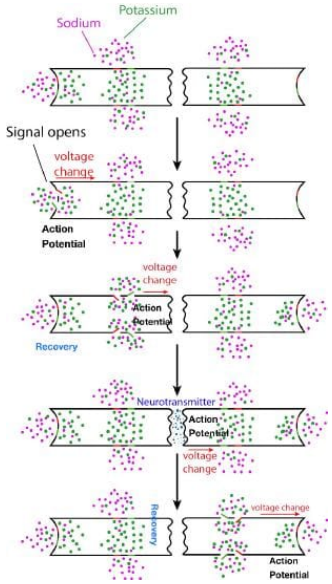


Figure 3.44 - Action potential moving across a synaptic junction of two nerve cells - Image by Pehr Jacobson

Moving signal

This chemical and electrical change that creates the action potential leaves the end of the nerve cell where it started and travels down the axon towards the other end of the nerve cell. Along the way, it encounters more sodium and potassium gated channels. In each case, these respond simply to the voltage change of the action potential and open and close, exactly in the same way the gates opened to start the signal. Thus, a rapid wave of increasing sodium ions and decreasing potassium ions moves along the nerve cell, propagated (and amplified) by gates opening and closing as the ions and charges move down the nerve cell. Eventually, the ionic tidal wave reaches the end of the nerve cell (axon terminal) facing the synaptic cleft.

Crossing the synaptic cleft

For the signal to be received by the intended target (postsynaptic cell) from the originating neuron (presynaptic neuron), it must cross the synaptic cleft and stimulate the neighboring cell (Figure 3.45). Communicating information across a synaptic cleft is the job of neurotransmitters. These are small molecules synthesized in nerve cells that are packaged in membrane vesicles called synaptic vesicles in the nerve cell. Neurotransmitters come in all shapes and chemical forms, from small chemicals like acetylcholine to peptides like neuropeptide Y. The most abundant neurotransmitter is glutamate, which acts at over 90% of the synapses in the human brain.

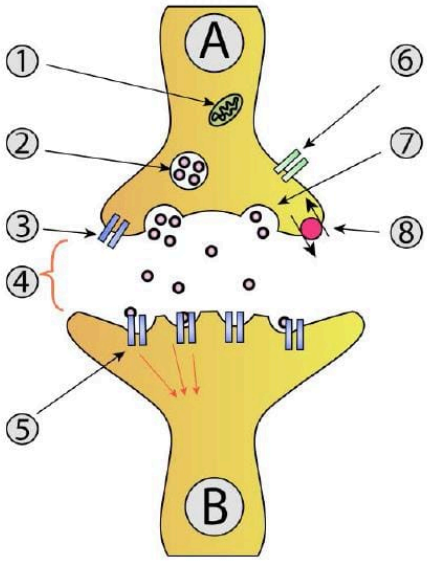


Figure 3.45 - A) Pre-synaptic neuron; B) post-synaptic neuron; 1) Mitochondria; 2) synaptic vesicle with neurotransmitters; 3) autoreceptor; 4) synapse with neurotransmitter released (serotonin); 5) postsynaptic receptors activated by neurotransmitter; 6) calcium channel; 7) exocytosis of a vesicle; 8) recaptured neurotransmitter. Wikipedia

Movie 3.2 - Movement of an action potential down a nerve cell - Wikipedia

Into the cleft

As the action potential in the presynaptic neuron approaches the axon terminus, synaptic vesicles begin to fuse with the membrane and their neurotransmitter contents spill into the synaptic cleft. Once in the cleft, the neurotransmitters diffuse, some of them reaching receptors on the postsynaptic cell. Binding of the neurotransmitter to the receptors on the membrane of the postsynaptic cell stimulates a response.

For motor neurons, the postsynaptic cell will be a muscle cell, and the response will be muscle contraction/relaxation. At this point, the originating nerve cell has done its job and communicated its information to its immediate target. If the postsynaptic cell is a nerve cell, the process repeats in that cell until it gets to its destination.

Na+/glucose transporter

Absorbing nutrients from the digestive system is necessary for animal life. The sodium/glucose transport protein is an electrogenic symporter that moves glucose into intestinal cells. It is found in the intestinal mucosa and the proximal tubule of the nephron of the kidney. The sodium/glucose transport system functions in the latter to promote reabsorption of glucose.

The pump works in conjunction with the Na+/K+ transport system. The gradient of sodium ions built up by the Na+/K+ pump is used as an energy source to drive movement of glucose into cells (see Figure 3.38). Use of an ion gradient established by a separate pump is known as secondary active transport. For intestinal mucosa, the pump transports glucose out of the gut and into gut cells. Later, the glucose is exported out the other side of the gut cells to the interstitial space for use in the body.

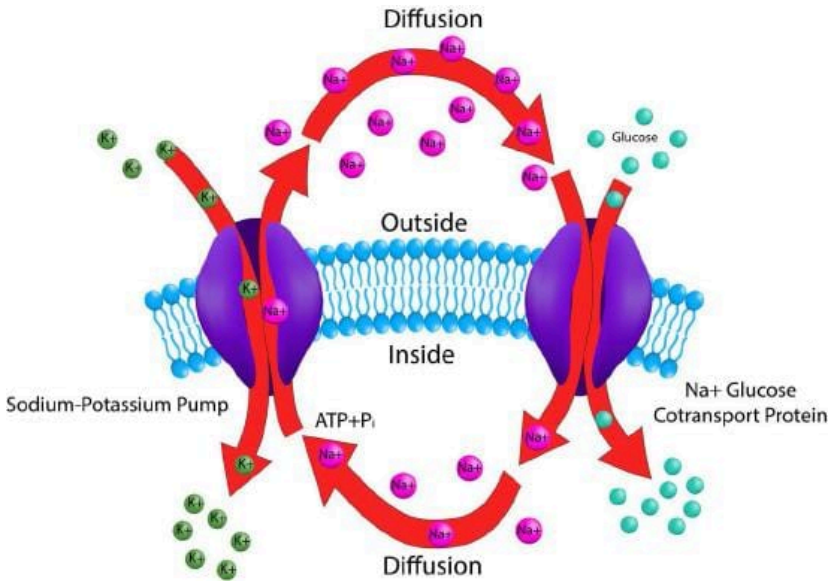


Figure 3.46 - The sodium/glucose pump (right) and the Na+/K+ ATPase (left). The sodium gradient generated by the Na+/K+ ATPase is used by the sodium/glucose pump to import glucose into the cell. Wikipedia

Calcium pumps

Calcium ions are necessary for muscular contraction and play important roles as signaling molecules within cells. In addition, when calcium concentrations rise too high, DNA in chromosomes can precipitate. Calcium concentration in cells is therefore managed carefully. It is kept very low in the cytoplasm as a result of action of pumps, both in the plasma membrane, which pump calcium outwards from the cytoplasm and in organelles, such as the endoplasmic reticulum (sarcoplasmic reticulum of muscle cells), which pump calcium out of the cytoplasm and into these organelles.

Opening of calcium channels, then, increases calcium concentration quickly in the cytoplasm resulting in a quick response, whether the intention is signaling or contraction of a muscle. After the response is generated, the calcium is pumped back out of the cytoplasm by the respective calcium pumps.

Some calcium pumps use ATP as an energy source to move calcium and others use ion gradients, such as sodium for the same purpose.

Na+/Ca++ transporter

One calcium pump of interest uses the sodium gradient as an energy source. It is the sodium/calcium pump. This electrogenic antiport system uses sodium's movement into the cell as a driving force to move calcium out of the cell, although its direction can reverse in some circumstances. The pump is a high capacity system to move a lot of calcium quickly, moving up to 5000 calcium ions per second and is found in many tissues with many functions.

Digitalis

One important function of the Na+/Ca++ pump occurs in heart cells. Ca++ is important for contraction of heart muscle. Calcium efflux from the cells is the normal operation of the pump, however, during the upstroke of the cycle, there is a large movement of sodium ions into the heart cell. When this occurs, the pump reverses and pumps Na+ out and Ca++ in briefly. Since calcium helps stimulate contraction of cardiac muscle, this can help make the heart beat stronger and is the focus of the use of digitalis to treat congestive heart failure.

Digitalis blocks the sodium-potassium ATPase and interferes with the sodium ion gradient. As noted above, when the Na+ gradient is oriented in the wrong direction, calcium is pumped in. Digitalis is therefore used to treat congestive heart failure because it increases the concentration of calcium in the heart cells, favoring more forceful beats.

ABC transporters

ABC transporters are another class of transmembrane proteins that use ATP energy to transport things against concentration gradients (Figures 3.47 & 3.48). This protein superfamily includes hundreds of proteins (48 in humans alone) and spans all extant phyla from prokaryotes to humans. These proteins function not only in membrane transport, but also in processes that include DNA repair and the process of translation.

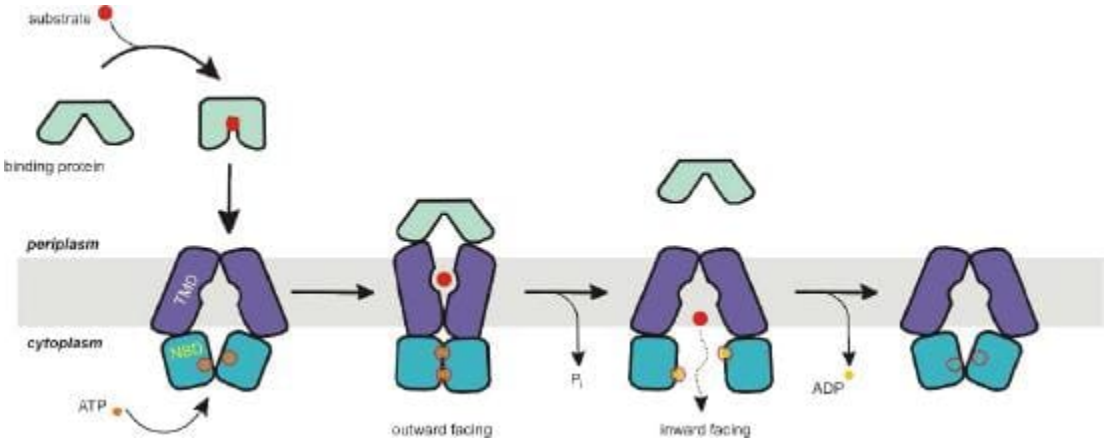


Figure 3.47 - Mechanism of an ABC importer

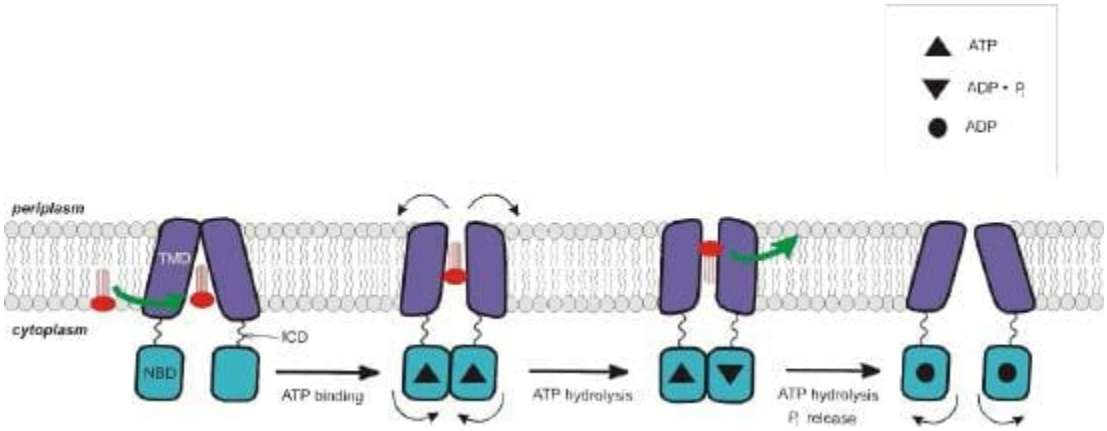


Figure 3.48 - Another view of an ABC exporter. NBD = Nucleotide Binding Domain

Transport

Substances that ABC transporters move across membranes include metabolic products, lipids, sterols, and drugs. ABC transporters function in multi-drug resistance of many cells, and provide resistance to antibiotics in bacteria as well as resistance to chemotherapy in higher cells by exporting drugs used to treat both of these types of cells.

ABC transporters are divided into three main groups - 1) importers (prokaryotes only); 2) exporters (prokaryotes and eukaryotes), and 3) non-transporters with roles in DNA repair and translation. All ABC transport proteins have four protein domains - two that are cytoplasmic and two that are membrane bound. They are alternately open to the cytoplasmic or extracellular (or periplasmic) regions and this is controlled by hydrolysis of ATP.

Disease

ABC transporters have roles in cystic fibrosis and other inherited human diseases. They are very involved in development of resistance to multiple drugs by a diverse group of cells. ABC transporters provide multi-drug resistance by expelling drug(s) from cells. ABCB1 protein, for example, pumps tumor suppression drugs out of the cell. Another ABC transporter known as Pgp transports organic cationic or neutral compounds.

Cystic fibrosis

Cystic fibrosis (CF) is a an autosomal recessive genetic disorder arising from mutations in both copies of the gene for the cystic fibrosis transmembrane conductance regulator (CFTR) protein. This ABC transporter system, which moves chloride and thiocyanate ions across epithelial tissue membranes exerts its effect mostly in the lungs but the pancreas, liver, kidneys, and intestine are all also affected by it.

Function

CFTR has roles in the production of sweat, mucus, and digestive fluids. Manifestations of the disease include breathing difficulty and overproduction of mucus in the lungs. When CFTR is functional, these fluids are normally thin, but when the gene is non-functional, they become much thicker and are points of infection.

CFTR contains two ATP-hydrolyzing domains and two cell membrane-crossing domains with 6 α-helices each. It can be activated by phosphorylation by a cAMP-dependent protein kinase. The carboxyl end of CFTR is linked to the cytoskeleton by a PDZ domain.

Lactose permease

Another integral membrane protein performing active transport is lactose permease. It facilitates the movement of the sugar lactose across the lipid bilayer of the cell membrane (Figures 3.49- 3.51). The transport mechanism is classified as a secondary active transport since it exploits the inwardly directed H+ electrochemical gradient as an energy source. When lactose is transported into cells, it is broken down into its substituent monosaccharide sugars - glucose and galactose - for energy creation.

The enzyme catalyzing this reaction is known as lactase and deficiency of it in humans leads to lactose intolerance (see HERE).

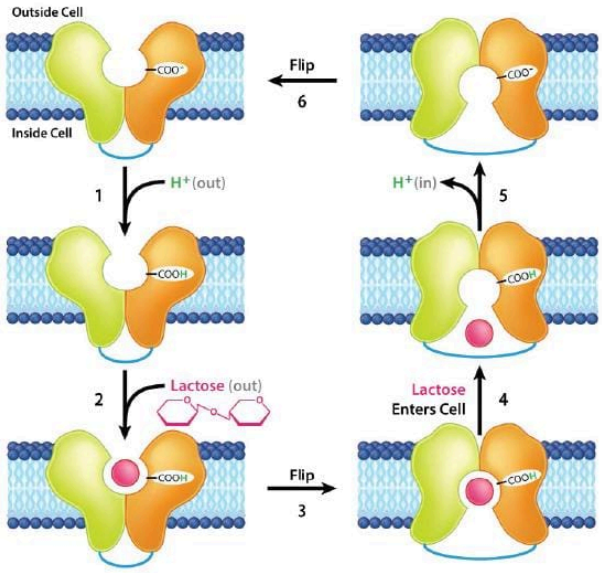


Figure 3.49 - Mechanism of action of lactose permease - Note that the proton acts by altering the charge of a carboxyl group - Image by Aleia Kim

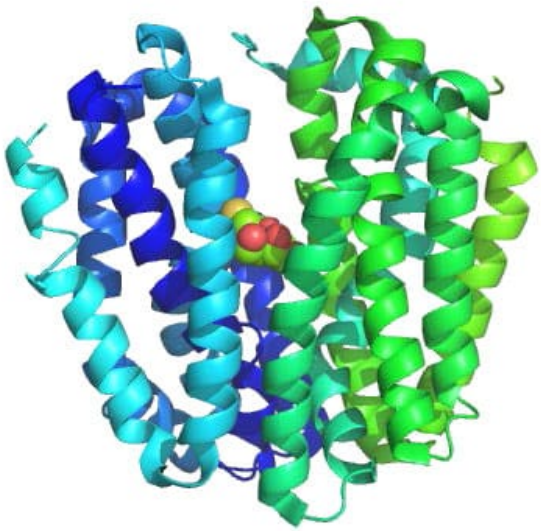


Figure 3.50 - Structure of lactose permease - Wikipedia

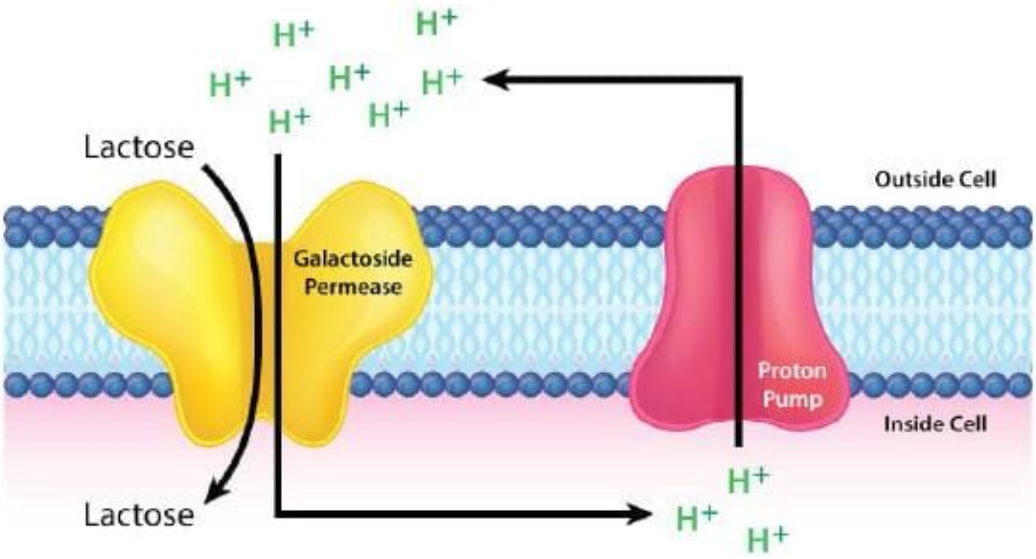


Figure 3.51 - Lactose (galactoside) permease is a secondary transporter, using a proton gradient as an energy source to pump lactose into the cell - Image by Aleia Kim

GLUTs

GLUTs (GLUcose Transport proteins) are uniport, type III integral membrane proteins that participate in the transport of glucose across membranes into cells. GLUTs are found in all phyla and are abundant in humans, with 12 GLUT genes. GLUT1, in erythrocytes is well-studied. Through GLUT 1, glucose enters and passes through it via facilitated diffusion at a rate that is 50,000 higher than in its absence. GLUTs of various types are found in different cells of the body. The one in red blood cells is known as GLUT 1 and has 12 membrane-spanning hydrophobic helices.

Though the structure of GLUT 1 is not known, it is speculated that the 12 helices form a chamber able to form hydrophilic bonds with glucose to facilitate its passage.

GLUT 1 levels in erythrocytes go up as glucose levels decrease and decrease when glucose levels go down. GLUT 1 can also transport ascorbate (vitamin C) in addition to glucose in mammals (such as humans) that do not produce their own vitamin C.

Glut 4

GLUT 4 is regulated by insulin and is found primarily in adipose and striated muscle tissue. Insulin alters intracellular trafficking pathways in response to increases in blood sugar to favor movement of various GLUT proteins (including GLUT 4) from intracellular vesicles to the cell membrane, thus stimulating uptake of the glucose. GLUT 4 is also found in the hippocampus where, if trafficking is disrupted, the result can be depressive behavior and cognitive dysfunction.

For all of the GLUT proteins, a key to keeping the glucose in the cell is phosphorylation of it by the glycolysis enzyme, hexokinase, in the cytoplasm. Phosphorylated molecules cannot enter GLUTs and don't have an easy means of exiting the cell.

Glucose Transporters

Name	Tissue	K_M (μM)	Comments
GLUT1	All mammalian tissues	1	Basal glucose uptake
GLUT2	Liver and pancreatic β cells	15-20	Participates in insulin regulation in the pancreas; removes excess glucose from the blood in the liver
GLUT3	All mammalian tissues	1	Basal glucose uptake
GLUT4	Muscle and fat cells	5	Concentration in plasma membrane of muscles rises with endurance training
GLUT5	Small intestine	---	Fructose transport

Figure 3.52 - GLUTs found in cells - Image by Aleia Kim

Distance Ed

To the tune of "Mister Ed"

Metabolic Melodies Website HERE

A course is a source,

of course, of course

Of all of the knowledge that we endorse

A major force for better/worse is the campus Distance Ed

It's true to outsource a college course

There are a few standards to be enforced

The long and short's we reinforce the campus Distance Ed

Bridge

A classroom class meets every week the same time every day

But Distance Ed is most unique - its flexible schedule's okay

E-course is a source, of course, of course

Of online assistance for lab reports

You're not enrolled in an online course?

Then sign up for this!

"You'll love Distance Ed"

Recording by David Simmons

Lyrics by Kevin Ahern
Recording by David Simmons Lyrics by Kevin Ahern

313

It's one o'clock and
Ahern's talkin'
Henderson and
Hasselbalch and
pKa's and
Buffers I should know
This song's for BB three five oh

I hope that maybe
He'll think the way we
Wrote our answers
Wasn't crazy
I really need the
Partial credit - so
This song's for BB three five oh

It's really groovy
That it improves me
Watching lectures
In Quicktime movies
I really need to
Go and download those
Podcasts for BB three five oh

This Song's For BB 3-5-0
To the tune of "This Land is My Land"

Metabolic Melodies Website HERE

I'm feeling manic
I'm in a panic
I'd better study
My old organic
It has reactions
That I need to know
This song's for BB three five oh

I know he said it
That's why I dread it
'cause I skipped Friday's
Extra credit
'twil pro'bly haunt me
That lowly ze-ro
Grade in BB three five oh

It could be steric
Or esoteric
That carbons get so
Anomeric
I'm too hysteric
Better let it go
This song's for BB three five oh

Recording by Tim Karplus

Lyrics by Kevin Ahern
Recording by Tim Karplus Lyrics by Kevin Ahern

This page titled 3.2: Transport in Membranes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

3.3: Other Considerations in Membranes

Source: BiochemFFA_3_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

There are many functions and factors relating to cell membranes that don't fit into broad categories. Those items will be the focus of this section.

Endocytosis

Besides transporter proteins and ion channels, another common way for materials to get into cells is by the process of endocytosis. Endocytosis is an alternate form of active transport for getting materials into cells. Some of these processes, such as phagocytosis, are able to import much larger particles than would be possible via a transporter protein. Like transporter proteins, endocytosis uses energy for the purpose (though it is not as visible as with protein transporters), but unlike protein transporters, the process is not nearly as specific for individual molecules.

As a result, the process usually involves the importation of many different molecules each time it occurs. The list of compounds entering cells in this way includes LDLs and their lipid contents, but it also include things like iron (packaged in transferrin), vitamins, hormones, proteins, and even some viruses sneak in by this means. There are three types of endocytosis we will consider (Figure 3.53).

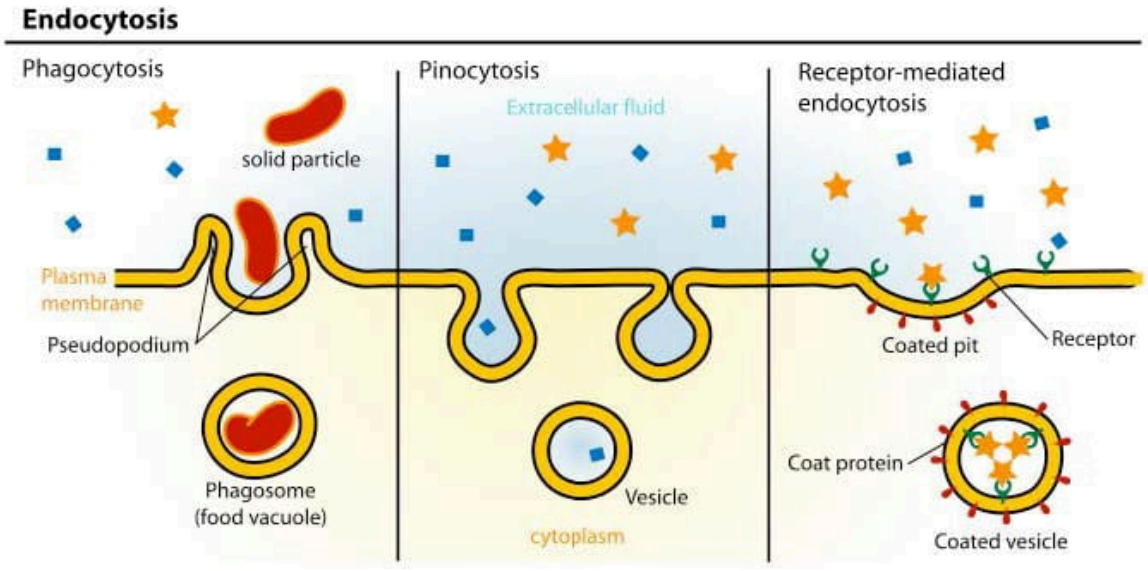


Figure 3.53 - Three types of endocytosis

Receptor mediated endocytosis

The process of receptor mediated endocytosis is a relatively specific means of bringing molecules into cells because it requires the incoming material to be somehow associated with a specific cell surface receptor. In the example of Figure 3.53, the receptor is the cellular LDL receptor. Clathrin-coated invaginations, as shown in the figure are known as "coated pits." The mechanism proceeds with an inward budding of the plasma membrane receptor (coated vesicles). Binding of the ligand (ApoB-100 of the LDL, for example, in Figure 3.54) to the LDL receptor leads to formation of a membrane invagination. The absorbed LDL particle fuses to form an early endosome (Figure 3.55) and contents are subsequently sorted and processed for use by the cell.

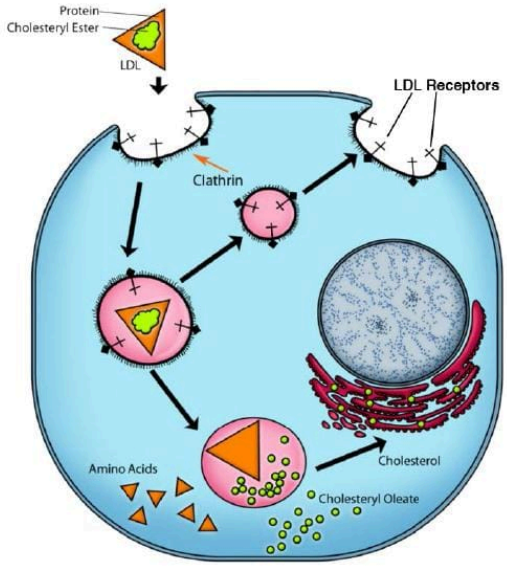


Figure 3.54 - Overview of clathrin-based receptor mediated endocytosis - Image by Aleia Kim

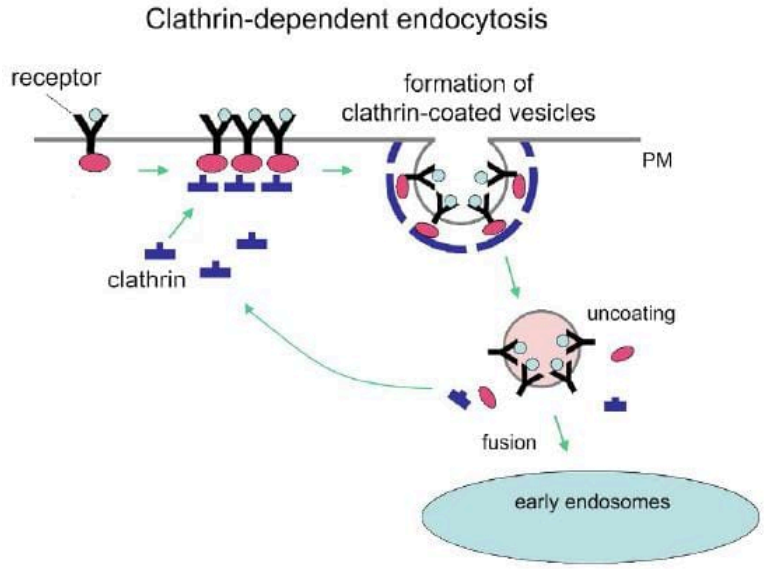


Figure 3.55 - Clathrin-mediated endocytosis of receptors - Wikipedia

The components from the coated vesicle are recycled to the plasma membrane for additional actions. Receptor mediated endocytosis can also play a role in internalization of cellular receptors that function in the process of signaling. Here, a receptor bound to a ligand is brought into the cell and may ultimately generate a response in the nucleus.

While receptor mediated endocytosis of receptors can have the effect of communicating a signal inwards to the cell, it can also reduce the total amount of signaling occurring, since the number of receptors on the cell surface is decreased by the process.

Non-clathrin endocytosis

There are three types of endocytosis occurring in cells that do not involve clathrin. They are 1) caveolae-based endocytosis, 2) macropinocytosis, and 3) phagocytosis. Caveolae-based endocytosis is based on a receptor molecule known as caveolin. Caveolins are a class of integral membrane proteins that compartmentalize and concentrate signaling molecules in the process of endocytosis. They are named for the cave-like caveolae structures of the plasma membrane where they are found.

Caveolins

Caveolins have affinity for cholesterol and associate with it in the membrane of cells, causing the formation of membrane invaginations of about 50 nm. The caveolin proteins can oligomerize and this is important for the coating and formation of the cave-like structures.

There are three caveolin genes found in vertebrate cells, CAV1, CAV2, and CAV3. Down-regulation of caveolin-1 results in less efficient cellular migration in vitro. Caveolins are implicated in both formation and suppression of tumors. High expression of them inhibits cancer-related growth factor signaling pathways, but some caveolin-expressing cancer cells are more aggressive and metastatic, possible due to an enhanced capacity for anchorage-independent growth.

Macropinocytosis

A phenomenon known as “cell drinking,” macropinocytosis literally involves a cell “taking a gulp” of the extracellular fluid. It does this, as shown in Figure 3.56, by a simple invagination of ruffled surface features of the plasma membrane. A pocket results, which pinches off internally to create a vesicle containing extracellular fluid and dissolved molecules. Within the cytosol, this internalized vesicle will fuse with endosomes and lysosomes. The process is non-specific for materials internalized.

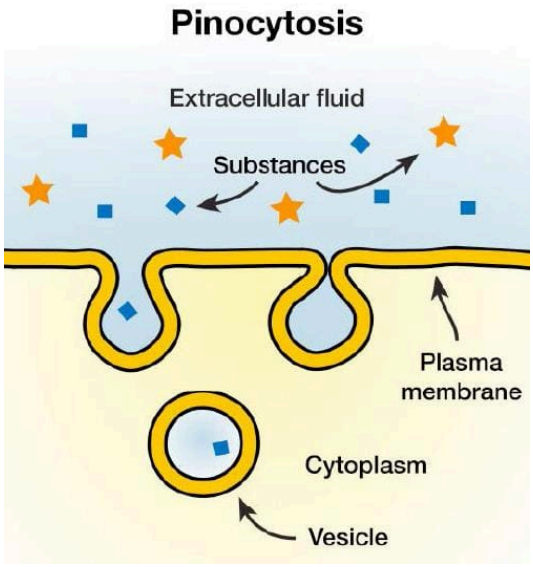


Figure 3.56 - Macropinocytosis

Phagocytosis

Phagocytosis is a process whereby relatively large particles (0.75 µm in diameter) are intenalized. Cells of the immune system, such as neutrophils, macrophages, and others, use phagocytosis to internalize cell debris, apoptotic cells, and microorganisms.

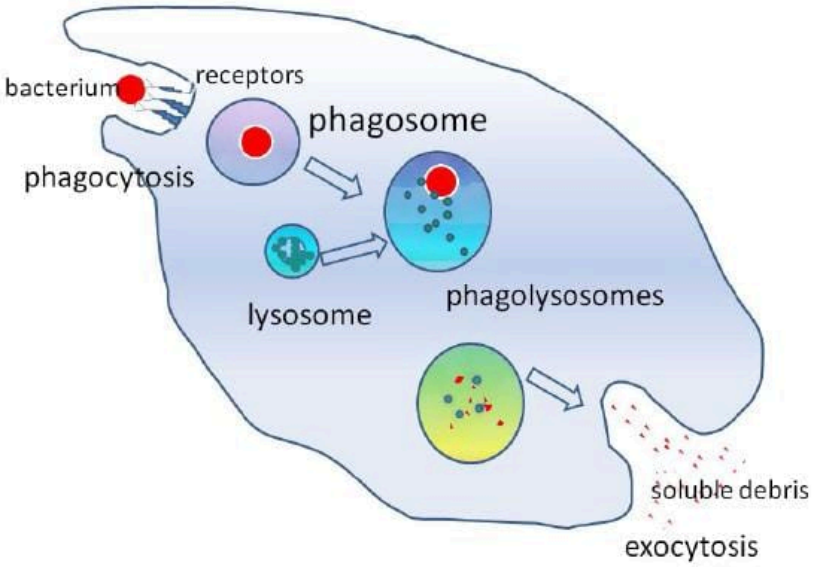


Figure 3.57 - Generalized scheme for phagocytosis of a bacterium - Wikipedia

The process operates through specific receptors on the surface of the cell and phagocytosing cell engulfs its target by growing around it. The internalized structure is known as a phagosome, which quickly merges with a lysosome to create a phagolysosome (Figure 3.58), which subjects the engulfed particle to toxic conditions to kill it, if it is a cell, and/or to digest it into smaller pieces. In some cases, as shown in the figure, soluble debris may be released by the phagocytosing cell.

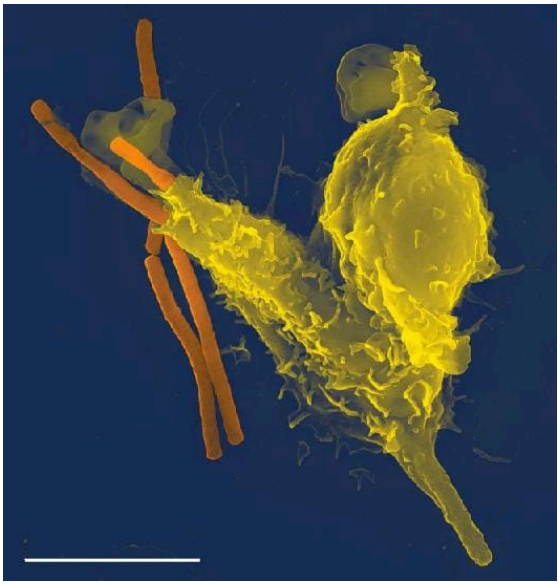


Figure 3.58 - Phagocytosis by a neutrophil (yellow) of an Anthrax bacillus (orange) - Wikipedia

Endosomes

Internalized material from endocytosis that doesn't involve phagocytosis passes through an internalized structure called an endosome. Endosomes are membrane bounded structures inside of eukaryotic cells that play a role in endocytosis (Figure 3.59). They have a sorting function for material internalized into the cell, providing for retrieval of materials not destined for destruction in the lysosomes. LDLs, for example, are targeted after endocytosis to the endosomes for processing before part of them is delivered to the lysosome. The endosomes can also receive molecules from the trans-Golgi network. These can be delivered to the lysosomes, as well, or redirected back to the Golgi. Endosomes come in three forms - 1) early, 2) late, and 3) recycling.

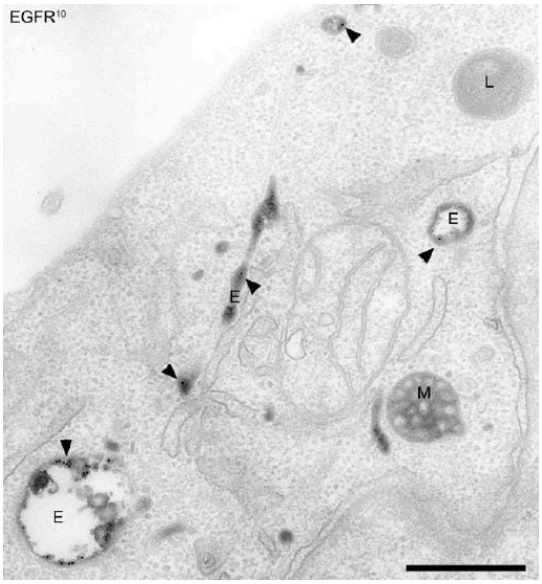


Figure 3.59 - Internalization of the epidermal growth factor receptor (EGFR) into endosomes. Early (E) and late (M) endosomes and lysosomes (L) are labeled. - Wikipedia

Exocytosis

The process of exocytosis is used by cells to export molecules out of cells that would not otherwise pass easily through the plasma membrane. In the process, secretory vesicles fuse with the plasma membrane and release their contents extracellularly. Materials, such as proteins and lipids embedded in the membranes of the vesicles become a part of the plasma membrane when fusion between it and the vesicles occurs.

Membrane fusion

Fusion is a membrane process where two distinct lipid bilayers merge their hydrophobic cores, producing one interconnected structure. Membrane fusion involving vesicles is the mechanism by which the processes of endocytosis and exocytosis occur.

When the fusion proceeds through both leaflets of both bilayers, an aqueous bridge results and the contents of the two structures mix. Common processes involving membrane fusion (Figure 3.60) include fertilization of an egg by a sperm, separation of membranes in cell division, transport of waste products, and neurotransmitter release (Figure 3.61). Artificial membranes such as liposomes can also fuse with cellular membranes. Fusion is also important for transporting lipids from the point of synthesis inside the cell to the membrane where they are used. Entry of pathogens can also be governed by fusion, as many bilayer-coated viruses use fusion proteins in entering host cells.

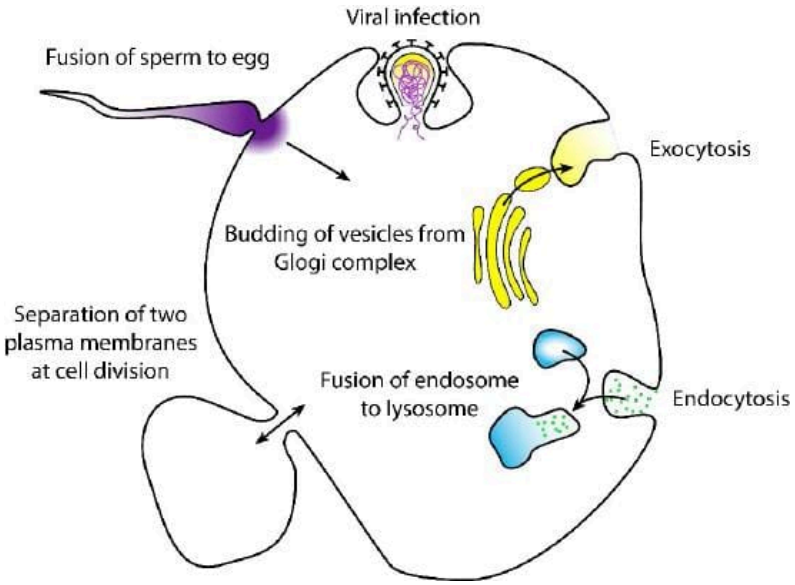


Figure 3.60 - Cell membrane fusions - Image by Pehr Jacobson

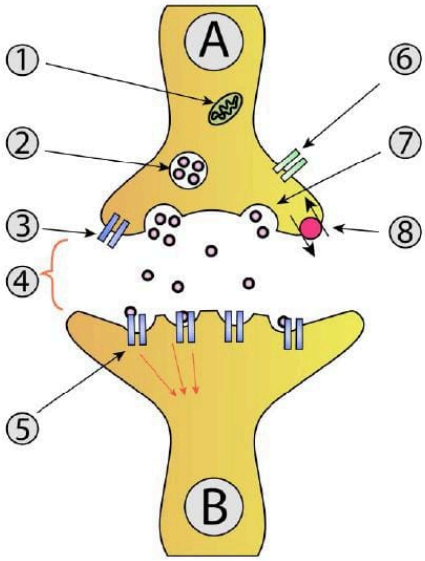


Figure 3.61 - Release of neurotransmitters (small circles) from presynaptic neuron A to postsynaptic neuron B. 1 = Mitochondrion / 2 = Synaptic vesicle with neurotransmitter / 3 = Autoreceptor / 4 = Synaptic cleft / 5 = Neurotransmitter receptor / 6 = Calcium channel / 7 = Fused vesicle releasing neurotransmitter / 8 = Neurotransmitter re-uptake pump - Wikipedia

SNARE proteins

Mediation of fusion of vesicles in exocytosis is carried out by proteins known as SNAREs (Soluble NSF Attachment Protein REceptor). This large superfamily of proteins spans a wide biological range, from yeast to mammals. Common vesicle fusions occur when synaptic vesicles dock with neurons (Figure 3.61) and release neurotransmitters. These are well-studied. The SNAREs involved in this process can be proteolytically cleaved by bacterial neurotoxins that give rise to the conditions of botulism and tetanus.

SNAREs are found in two locations. v-SNAREs are found in the membranes of transport vesicles during the budding process, whereas t-SNARES can be found in the membranes of targeted compartments.

The act of vesicle fusion coincides with increases of intracellular calcium. Fusion of synaptic vesicles in neurotransmission results in activation of voltage-dependent calcium channels in the targeted cell. Influx of calcium helps to stimulate vesicle fusion. In the endocrine system, binding of hormones to G protein coupled receptors activate the IP3/DAG system to increase levels of calcium.

In the process of membrane fusion (Figure 3.62), the v-SNAREs of a secretory vesicle (upper left) interact with the t-SNAREs of a target membrane (bottom). The v- and t-SNAREs “zipper” themselves together to bring the membrane vesicle and the target membrane closer together.

Zippering also causes flattening and lateral tension of the curved membrane surfaces, favoring hemifusion of the outer layers of each membrane. Continued tension results in subsequent fusion of the inner membranes as well, yielding opening of the contents of the vesicle chamber to its target (usually outside the cell).

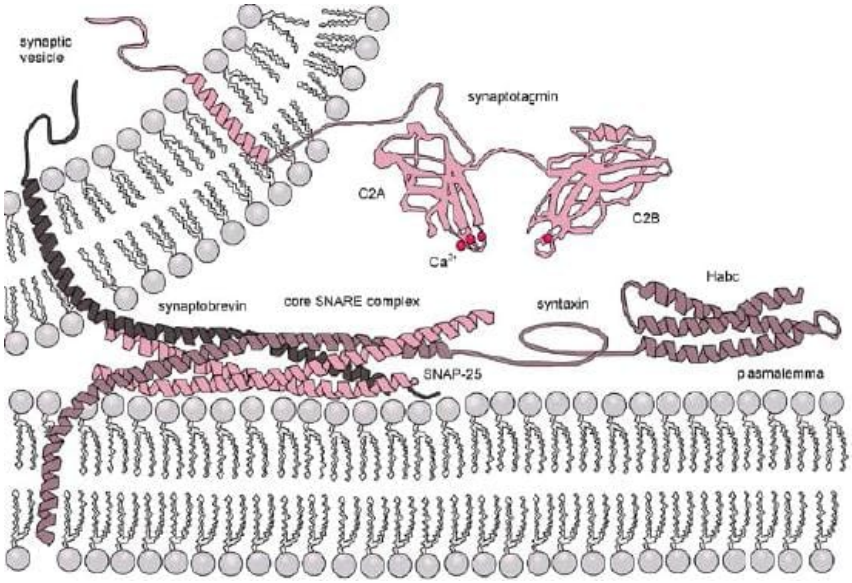


Figure 3.62 - Proteins involved in vesicle fusion in neurotransmission. A SNARE complex between α-helices of synaptobrevin, syntaxin and SNAP-25 intertwine and “zip” membranes together. Synaptotagmin is a calcium sensor regulating the process of zipping - Wikipedia

Shuttles

Another way to transport items across a membrane for which there is no specific transport system available is the use of shuttles. Shuttles are important when there is no transport mechanism for moving material across a membrane for which no transport system exists.

A great example is NADH. NADH is an important electron carrier that is produced in the cytoplasm and mitochondria of the cell. NADH produced in the mitochondrion goes directly to the electron transport system and delivers electrons to Complex I. NADH produced in the cytoplasm (such as from glycolysis) does not have this option, since the inner membrane of the mitochondrion is impermeable to the molecule and no transporter exists to move it across. The important part of the NADH is its electron cargo, so cells have evolved two ways to move the electrons into the mitochondrial matrix apart from NADH.

Both methods involve shuttles. In each case, an acceptor molecule receives electrons from NADH and the reduced form of the acceptor molecule is transported. It gets transported into the matrix where it is oxidized (electrons are lost) and then donated to the electron transport system.

Glycerol phosphate shuttle

The first of these methods is the least efficient, but it is rapid. It found commonly in muscles which have needs for rapid energy and brain tissue. This shuttle is referred to as the glycerol phosphate shuttle and is shown in Figure 3.63. It operates in the intermembrane space between the inner and outer mitochondrial membranes. The outer mitochondrial membrane is very porous, allowing many materials to pass freely through it. In the intermembrane space, the cytoplasmic enzyme, glyceraldehyde-3-phosphate dehydrogenase (cGPD) catalyzes transfer of electrons from NADH to dihdydroxyacetone phosphate (#2 in the figure), yielding NAD+ and glyceraldehyde-3-phosphate (#1 in the figure). The glyceraldehyde-3-phosphate then binds to a glyceraldehyde-3-phosphate dehydrogenase (mGPD) embedded in the outer portion of the inner mitochondrial membrane. mGPD catalyzes the transfer of electrons from glyceraldehyde-3-phosphate to FAD, producing dihycroxyacetone phosphate and FADH2. FADH2 then transfers its electrons to the electron transport system through CoQ (Q above), forming CoQH2 (QH2 above). As will be discussed in the section on electron transport, this is not an efficient shuttle system because it does not result in

production of as much ATP as occurs when electrons are transferred to NAD+ instead of FAD.

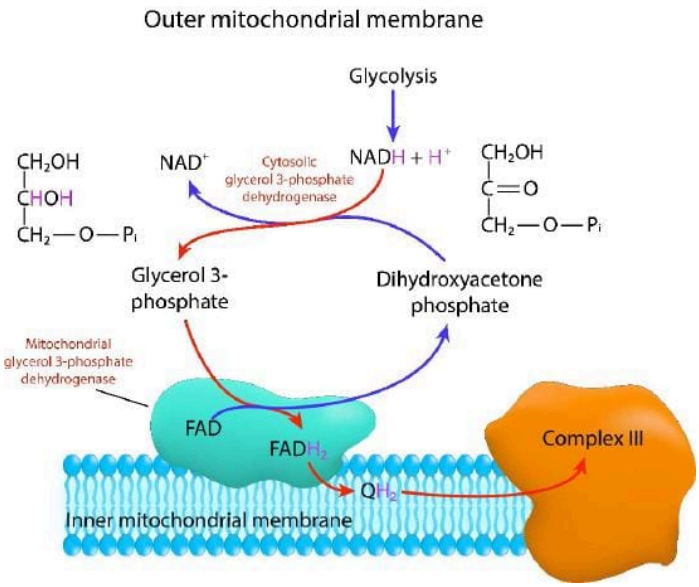


Figure 3.63 - Glycerol phosphate shuttle system in the intermembrane space of a mitochondrion - Image by Pehr Jacobson

Malate-aspartate shuttle

A more efficient system of transferring electrons is the malate-aspartate shuttle and it is shown in Figure 3.64. As is apparent in the figure, this shuttle involves more steps than the glycerol phosphate shuttle, but the advantage of the malate-aspartate shuttle is that it is more efficient. NADH outside of the mitochondrion transfers its electrons to the shuttle and then NADH is re-made on the inside of the shuttle. No energy is expended in the process.

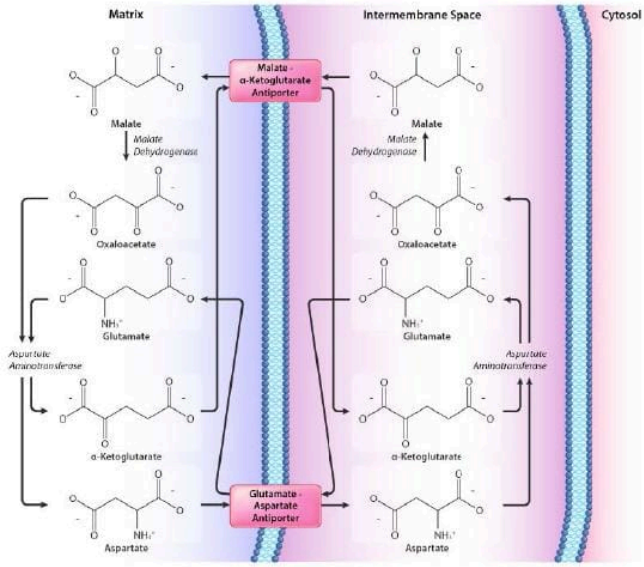


Figure 3.64 - The malate aspartate shuttle - Image by Aleia Kim

When NADH accumulates in the cytoplasm, it moves to the intermembrane space where the enzyme malate dehydrogenase catalyzes the transfer of electrons to oxaloacetate to yield NAD+ and malate. A transport system for malate moves malate into the mitochondrial matrix in exchange for α-ketoglutarate.

Inside the mitochondrion, malate is reoxidized to oxaloacetate and electrons are given to NAD+ to recreate NADH. NADH then donates electrons to Complex I of the electron transport system. That's really all there is to the shuttle. The remaining steps are simply to balance the equation of the process. Oxaloacetate accepts an amine group from glutamic acid to yield aspartic acid and α-ketoglutarate. Aspartate then moves out of the mitochondrion through an antiport transport protein that swaps it for glutamate.

A series of reactions in the intermembrane space balance the equation.

It is easy to get lost in the mess of balancing equations. The most important thing to understand here is that oxaloacetate accepts electrons on the outside to become malate which is the carrier of electrons across the membrane. Once inside the matrix, malate is converted back to oxaloacetate and its electrons are given to NAD+, forming NADH. Everything else is simple equation balancing.

Acetyl-CoA shuttle

Another kind of shuttle also involves the mitochondrion and in this case, the item being moved is a molecule, not a pair of electrons. The molecule of interest here is acetyl-CoA, for which no transport system operates, but which is needed in the cytoplasm for fatty acid synthesis when the cell has abundant energy.

Acetyl-CoA is mostly in the mitochondrion and so long as the citric acid cycle is operating efficiently, its concentration is relatively stable. However, when the citric acid cycle slows, acetyl-CoA and the citrate made from it in the cycle begin to accumulate.

A membrane transport system for citrate exists, so it gets moved out to the cytoplasm. In the cytoplasm, an enzyme known as citrate lyase cleaves citrate to acetyl-CoA and oxaloacetate. Oxaloacetate can be reduced to malate and moved back into the mitochondrion.

As for acetyl-CoA, the more of that cells have in the cytoplasm, the more likely they will begin making fatty acids and fat, since acetyl-CoA is the starting material for fatty acid synthesis, which occurs in the cytoplasm. When does this process occur? As noted above, it occurs when the citric acid cycle stops and this occurs when levels of NADH and FADH2 increase. These, of course, increase when one is not burning off as many calories as one is consuming as a byproduct of respiratory control. Lack of exercise leads to higher levels of reduced electron carriers.

Cell junctions

Cells in multicellular organisms are in close contact with each other and links between them are called junctions. In vertebrate organisms, there are three main types of cell junctions and one of them (gap junctions) is important for movement of materials between cells. The three types are

1. Gap junctions
2. Adherens junctions, (Anchoring Junctions, desmosomes and hemidesmosomes)
3. Tight junctions

Cell junctions in multicellular plants are structured differently from those in vertebrates and are called plasmodesmata. They too function in exchange of materials between individual cells.

Gap junctions

Gap junctions are specialized structures made up of two sets of structures called connexons (one from each cell - see Figure 3.65) directly link the cytoplasms of the connected cells. Gap junctions are regulated to control the flow of molecules, ions, and electrical impulses between cells. In plants, similar structures known as plasmodesmata traverse the cell wall (Figure 3.66) and perform similar functions.

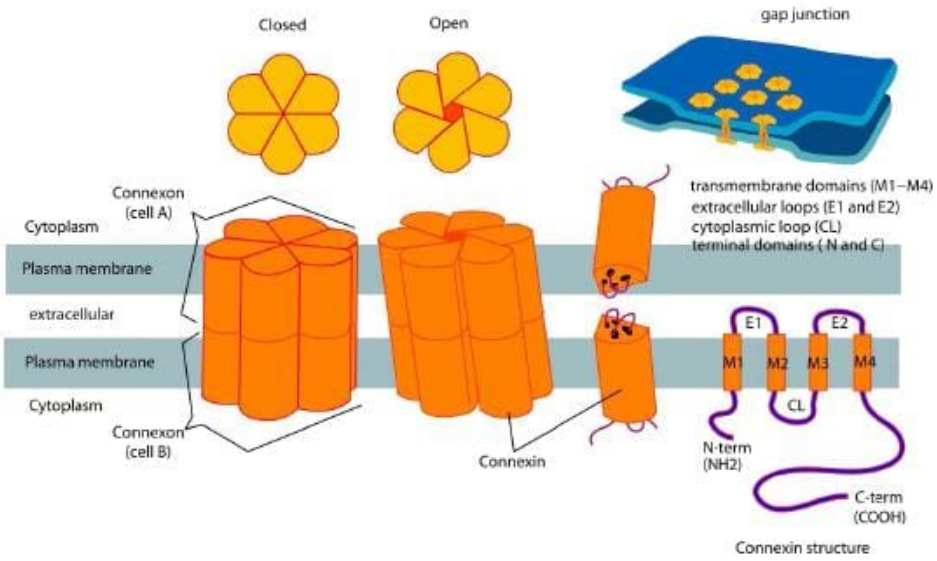


Figure 3.65 - Structure of connexons joining two cells. Bundles of six copies of connexin proteins in the plasma membrane of each cell comprise the connexon structures

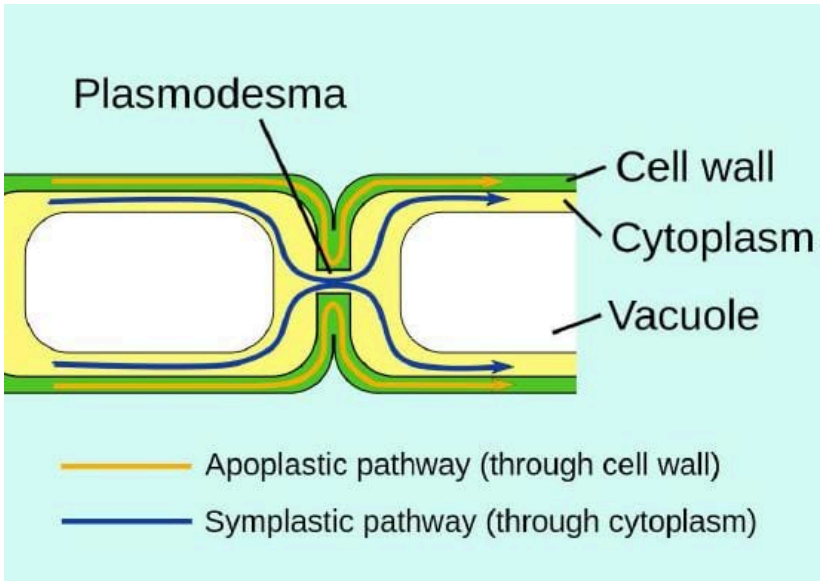


Figure 3.66 - Two means of intercellular communication in plant cells - apoplastic pathway (through cell wall) and symplastic pathway (through the plasodesma)

Adherens junctions

Adherens junctions (Figure 3.67) are protein complexes on the cytoplasmic side of the cell membranes of epithelial and endothelial tissues that link cells to each other or to the extracellular matrix. They correspond to the fascia adherens found in non-epithelial/non-endothelial cells.

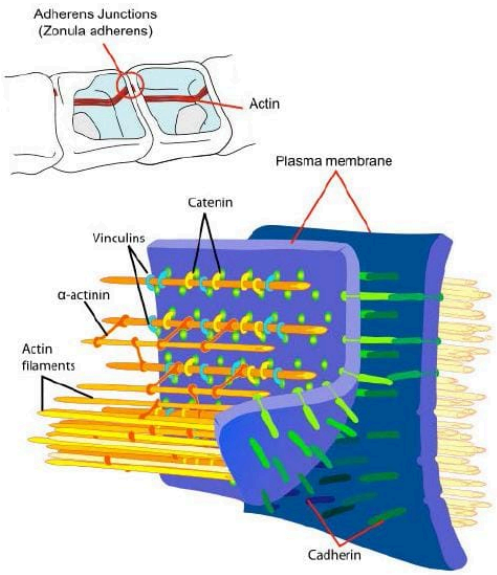


Figure 3.67 - Adherens junction

Adherens junctions contain the following proteins - 1) α-catenin (binds cadherin through β-catenin); 2) β-catenin (attachment for α-catenin to cadherin; 3) γ-catenin (binds to cadherin); 4) cadherins (group of transmembrane proteins that dimerize with cadherins on adjacent cells; 5) p120 (also called Δ-catenin - binds to cadherin); 6) plakoglobin (catenin family protein homologous to and acting like β-catenin); 7) actin; 8) actinin; and 9) vinculin. Adherens junctions may help to maintain the actin contractile ring which forms in the process of cytokinesis.

Tight junctions

Tight junctions (Figure 3.68) are a network of protein strands that seal cells together and restrict the flow of ions in the spaces between them. The effect of their structure is to restrict the movement of materials through tissues by requiring them to pass through cells instead of around them. Tight junctions join together the cytoskeletons of cells and through their structure maintain their apical and basolateral polarity.

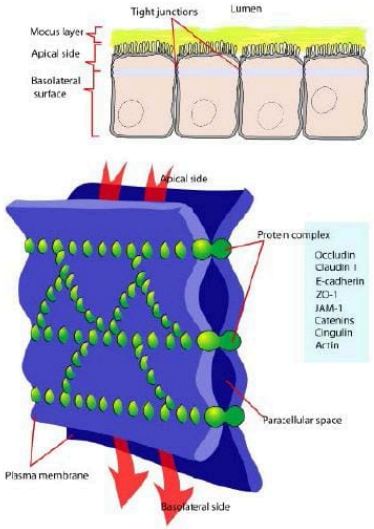


Figure 3.68 - Tight junctions

GPI anchors

Membrane proteins attached to glycosylphosphatidylinositol (also known as a GPI anchor) are referred to as being glypiated. The proteins, which play important roles in many biochemical processes, are attached to the GPI anchor at their carboxyl terminus.

Phospholipases, such as phospholipase C can cut the bond between the protein and the GPI, freeing the protein from the outer cell membrane. Proteins destined to be glypiated have two signal sequences. They are 1) An N-terminal signal sequence and 2) A C-terminal signal sequence that is recognized by a GPI transamidase (GPIT). The N-terminal signal sequences is responsible for directing co-translational transport into the endoplasmic reticulum. The C-terminal sequence is recognized by GPI transamidase, which links the carboxy terminus of a protein to the GPI anchor.

Liposomes

The spontaneous ability of phosphoglycerolipid and sphingolipid compounds to form lipid bilayers is exploited in the formation of artificial membranous structures called liposomes (Figure 3.69). Liposomes are useful for delivering their contents into cells via membrane fusion. In the figure, items targeted for delivery to cells would be encased in the middle circular region of the liposome and when the liposome fuses with the cell membrane, it will deliver the contents directly into the cytoplasm.

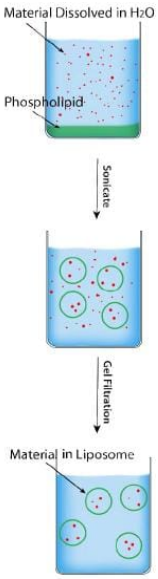


Figure 3.70 - Formation of liposomes from phospholipids in water - Image by Pehr Jacobson

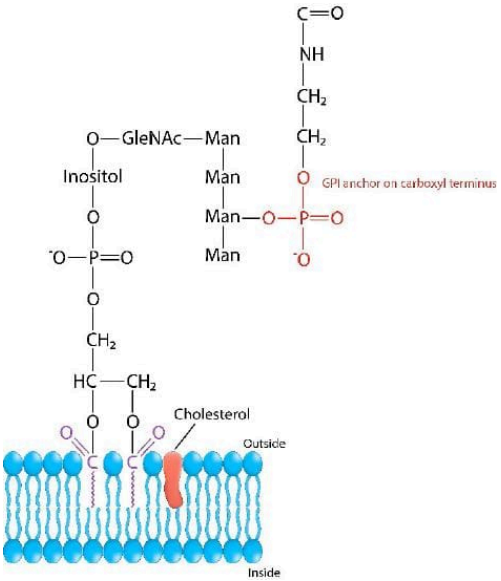


Figure 3.69 - A glypiated protein linked to a membrane-embedded inositol-based molecule. The protein portion is above the red phosphate ion - Image by Pehr Jacobson

Hydropathy index

The interior portion of the lipid bilayer is very hydrophobic, which makes it very restrictive to movement of ions and polar substances across it. This property also places limitations on the types of amino acids that will interact with it as well. Because of this, transmembrane protein domains found in integral membrane proteins are biased in the amino acids that interact with either the lipid bilayer or the aqueous material on either side of it.

Amino Acid	Hydropathy index
Alanine	1.8
Arginine	−4.5
Asparagine	−3.5
Aspartic acid	−3.5
Cysteine	2.5
Glutamic acid	−3.5
Glutamine	−3.5
Glycine	−0.4
Histidine	−3.2
Isoleucine	4.5
Leucine	3.8
Lysine	−3.9
Methionine	1.9
Phenylalanine	2.8
Proline	−1.6
Serine	−0.8
Threonine	−0.7
Tryptophan	−0.9
Tyrosine	−1.3
Valine	4.2

Figure 3.71 - Hydropathy index for amino acids. More positive values indicate higher hydrophobicity. Wikipedia

Hydrophobic amino acids are found within the bilayer, whereas hydrophilic amino acids are found predominantly on the surfaces. An additional clue to identifying membrane crossing regions of a protein is that tryptophan or tyrosine is commonly positioned at non-polar/polar interface(s) of the lipid bilayer for integral proteins. Such an organization of amino acids can be recognized by computer analysis of amino acid sequences using what is called a hydropathy index/score (Figure 3.71). Though the names and the scorings vary, the idea is to assign a number (usually positive) to amino acid side chains with higher hydrophobicity and negative to those that are ionic. With these scores, a computer program can easily find the average scores of short amino acid segments (say 3 amino acids long) and then plot those values on a graph of hydrophobicity index versus position in polypeptide chain. Doing that for a transmembrane protein such as glycophorin results in the plot shown in Figure 3.72. It is apparent in the analysis that this is a transmembrane protein that has seven domains crossing the lipid bilayer, as labeled.

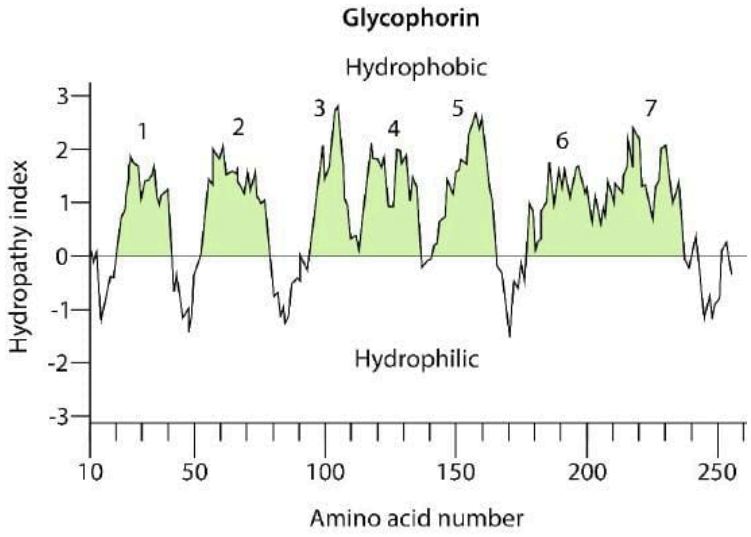


Figure 3.72 - Hydropathy index plot for glycophorin. Each lipid bilayer-crossing domain noted with a number - Image by Pehr Jacobson

Cell walls

Cells walls are found in many cells, including plants, fungi, and bacteria, but are not found in animal cells. They are designed to provide strength and integrity and at least some protection against bursting from osmotic pressure (Figures 3.73-3.75).

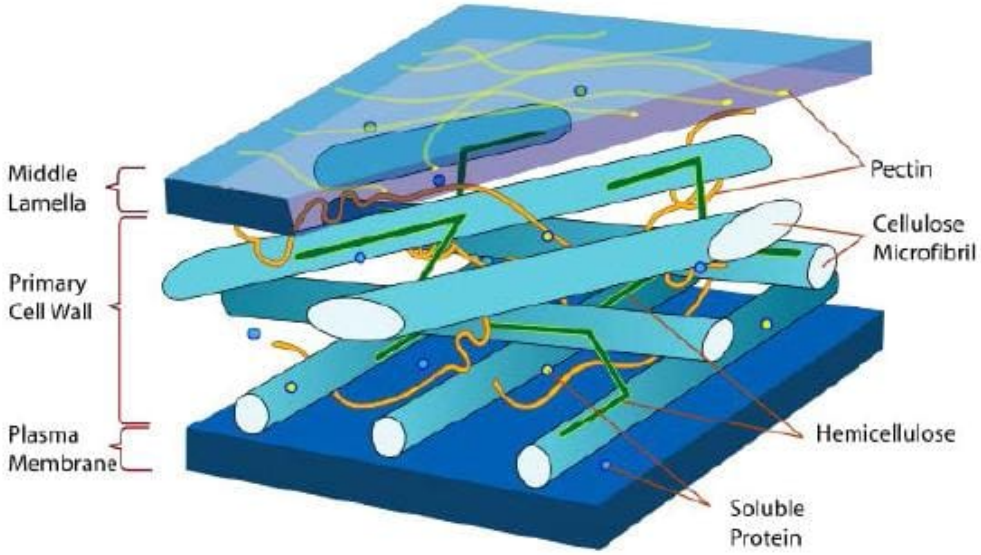


Figure 3.73 - Plant cell wall. Direction of the cytoplasm is down

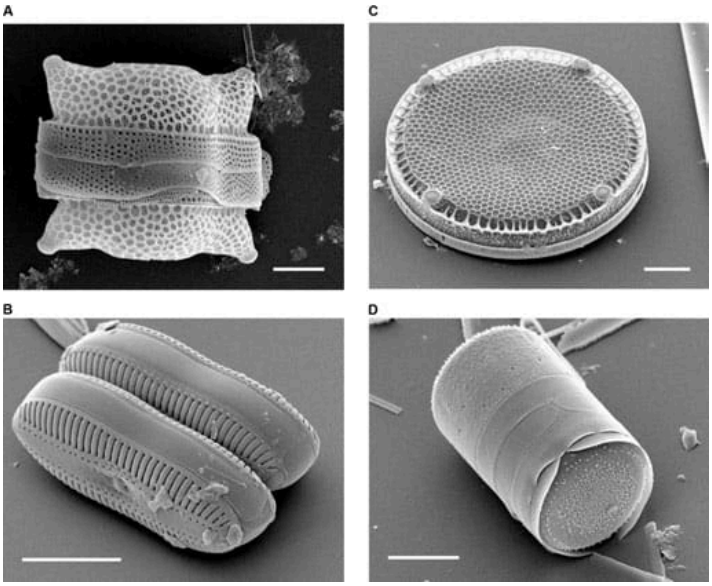


Figure 3.74 - Cell walls of diatoms - Wikipedia

Gram positive bacteria (Figure 3.75) have the simplest cell wall design. Moving from outside the cell towards the cytoplasm there is an outer peptidoglycan layer for the cell wall followed by a periplasmic space, a plasma membrane, and then the cytoplasm. Gram negative bacteria add a second protective layer external to all of this, so for them, starting at the outside and moving inwards, one encounters an outer lipopolysaccharide layer, a periplasmic space, the peptidoglycan cell wall, a second periplasmic space, a plasma membrane and then the cytoplasm.

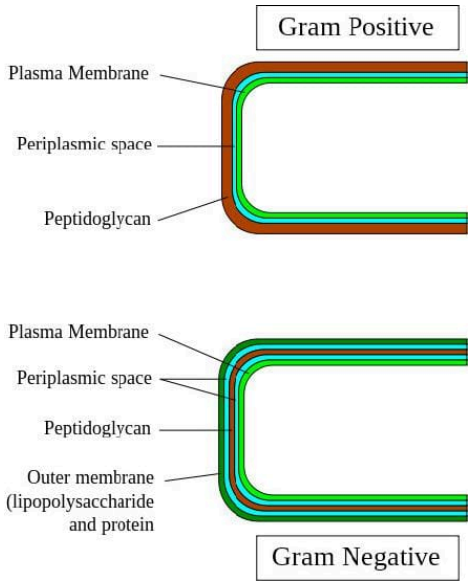


Figure 3.75 - Gram positive versus Gram negative bacteria cell coverings - Wikipedia

Herbaceous plants have a rigid outer cell wall (primarily composed of cellulose, hemicellulose, and pectin) and an inner plasma membrane. Woody plants add a second level of wall with lignin between the cellulosic wall and the plasma membrane of herbaceous plants.

BB Wonderland

To the tune of “Winter Wonderland”

Metabolic Melodies Website HERE

Milam Hall - It’s 12:30

And Ahern’s gettin’ wordy

He walks to and fro’

While not talkin’ slow

Givin’ it to B-B-4-5-0

I was happy when the term got started

Lecture notes and videos galore

MP3s got added to my iPod

But recitations sometimes were a bore

And exams bit me roughly

When the curve turned out ugly

I don’t think it’s so

My scores are too low

Slidin’ by in B-B-4-5-0

Final-LY there’s an examination

On December 9th at 6:00 pm

I'll have my card packed with information

So I don't have to memorize it then

And I'll feel like a smarty

With my jam-packed note-cardy

Just one more to go

And then ho-ho-ho

I'll be done with B-B-4-5-0

Recording by David Simmons

Lyrics by Kevin Ahern
Recording by David Simmons Lyrics by Kevin Ahern

334

Thank God There's a Video

To the tune of "Thank God I'm a Country Boy"

Metabolic Melodies Website HERE

There's a bundle of things a student oughta know

And Ahern's talk isn't really very slow

Learnin' ain't easy / the lectures kinda blow

Thank God there's a video

Well we've gone through the cycles and their enzymes too
Studying the regulation everything is new
I gotta admit that I haven't got a clue
What am I gonna do?

So I got me a note card and bought me a Stryer
Got the enzymes down and the names he requires
I hope that I can muster up a little more desire
Thank God there's a video

Just got up to speed about the N-A-D
Protons moving through Complex Vee
Electrons dance in the cytochrome C
Gotta hear the MP3

Fatty acid oxidation makes acetyl-CoA
Inside the inner matrix of the mitochondri-ay
It's very complicated, I guess I gotta say
Thank God there's a video

So I got me a note card and bought me a Stryer
Got the enzymes down and the names he requires
I hope that I can muster up a little more desire

Thank God there's a video

Replication's kind of easy in a simple kind of way
Copyin' the bases in the plasmid DNAs
Gs goes with Cs and Ts go with As
Thanks to polymerase

And the DNA's a template for the RNA
Helices unwinding at T-A-T-A
Termination happens, then the enzyme goes away
Don't forget the poly-A

So I got me a note card and bought me a Stryer
Got the enzymes down and the names he requires
I think that I can muster up a little more desire
Thank God there's a video

Recording by David Simmons

Lyrics by Kevin Ahern
Recording by David Simmons Lyrics by Kevin Ahern

This page titled 3.3: Other Considerations in Membranes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

4.1: Basic Principles of Catalysis

A printable version of this section is here: BiochemFFA_4_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

If there is a magical component to life, an argument can surely be made for it being catalysis. Thanks to catalysis, reactions that can take hundreds of years to complete in the uncatalyzed "real world," occur in seconds in the presence of a catalyst. Chemical catalysts, such as platinum, can speed reactions, but enzymes (which are simply super-catalysts with a "twist," as we shall see) put chemical catalysts to shame (Figure 4.1). To understand enzymatic catalysis, it is necessary first to understand energy. Chemical reactions follow the universal trend of moving towards lower energy, but they often have a barrier in place that must be overcome. The secret to catalytic action is reducing the magnitude of that barrier.

Enzyme	Nonenzymatic Half-Life	Uncatalyzed Rate (k_{un} s^{-1})	Catalyzed Rate (k_{cat} s^{-1})	Rate Enhancement (k_{cat} s^{-1}/k_{un} s^{-1})
OMP decarboxylase	78,000,000 years	28×10^{-16}	39	1.4×10^{17}
Staphylococcal nuclease	130,000 years	1.7×10^{-13}	95	5.6×10^{14}
Carboxypeptidase A	7.3 years	3.0×10^{-9}	578	1.9×10^{11}
Ketosteroid isomerase	7 weeks	1.7×10^{-7}	66,000	3.9×10^{11}
Triose phosphate isomerase	1.9 days	4.3×10^{-6}	4,300	1.0×10^{9}
Chorismate mutase	7.4 hours	2.6×10^{-5}	50	1.9×10^{6}
Carbonic anhydrase	5 seconds	1.3×10^{-1}	1×10^{6}	7.7×10^{6}

Abbreviations:
OMP - Orotine monophosphate
AMP - Adenosine monophosphate

Figure 4.1 - Rate enhancement for several enzymes Image by Aleia Kim

Before discussing enzymes, it is appropriate to pause and discuss an important concept relating to chemical/biochemical reactions. That concept is *equilibrium* and it is very often misunderstood. The "equi" part of the word relates to equal, as one might expect, but it does not relate to absolute concentrations. What happens when a biochemical reaction is at equilibrium is that the concentrations of reactants and products do not change over time. This does not mean that the reactions have stopped. Remember that reactions are reversible, so there is a forward reaction and a reverse reaction: if you had 8 molecules of A, and 4 of B at the beginning, and 2 molecules of A were converted to B, while 2 molecules of B were simultaneously converted back to A, the number of molecules of A and B remain unchanged, i.e., the reaction is at equilibrium. However, you will notice that this does not mean that there are equal numbers of A and B molecules.

Concentration Matters

So, contrary to the perceptions of many students, the concentrations of products and reactants are not equal at equilibrium, unless the ΔG°' for a reaction is zero, because when this is the case,

$$\Delta G = \ln\left(\frac{[\text{Products}]}{[\text{Reactants}]}\right) \tag{4.1.1}$$

since the ΔG°' is zero. Because ΔG itself is zero at equilibrium, then

$$[Products] = [Reactants]. \tag{4.1.2}$$

This is the only circumstance where

$$[Products] = [Reactants] \tag{4.1.3}$$

at equilibrium. Reiterating, at equilibrium, the concentrations of reactant and product do not change over time. That is, for a reaction $A \rightleftharpoons B[A]$ at time zero when equilibrium is reached, $[A]_{T_0}$, will be the same 5 minutes later (assuming A and B are

chemically stable). Thus,

$$[A]_{T_0} = [A]_{T+5} \tag{4.1.4}$$

Similarly,

$$[B]_{T_0} = [B]_{T+5} \tag{4.1.5}$$

For that matter, at any amount of time X after equilibrium has been reached,

$$[A]T0 = [A]T + 5 = [A]TX \tag{4.1.6}$$

and

$$[B]T0 = [B]T + 5 = [B]TX \tag{4.1.7}$$

However, unless ΔG°' = 0, it is wrong to say [A]T0 = [B]T0 As we study biochemical reactions and reaction rates, it is important to remember that 1) reactions do not generally start at equilibrium; 2) all reactions move in the direction of equilibrium; and 3) reactions in cells behave just like those in test tubes - they do not begin at equilibrium, but they move towards it. Dynamic reactions The reactions occurring in cells, though, are very dynamic and complex. In a test tube, they can be studied one at a time. In cells, the product of one reaction is often the substrate for another one. Reactions in cells are interconnected in this way, giving rise to what are called metabolic pathways. There are, in fact, thousands of different interconnected reactions going on continuously in cells.

Attempts to study a single reaction in the chaos of a cell is daunting to say the least. For this reason, biochemists isolate enzymes from cells and study reactions individually. It is with this in mind that we begin our consideration of the phenomenon of catalysis by describing, first, the way in which enzymes work.

Activation energy

Figure 4.2 schematically depicts the energy changes that occur during the progression of a simple reaction. In order for the reaction to proceed, an activation energy must be overcome in order for the reaction to occur.

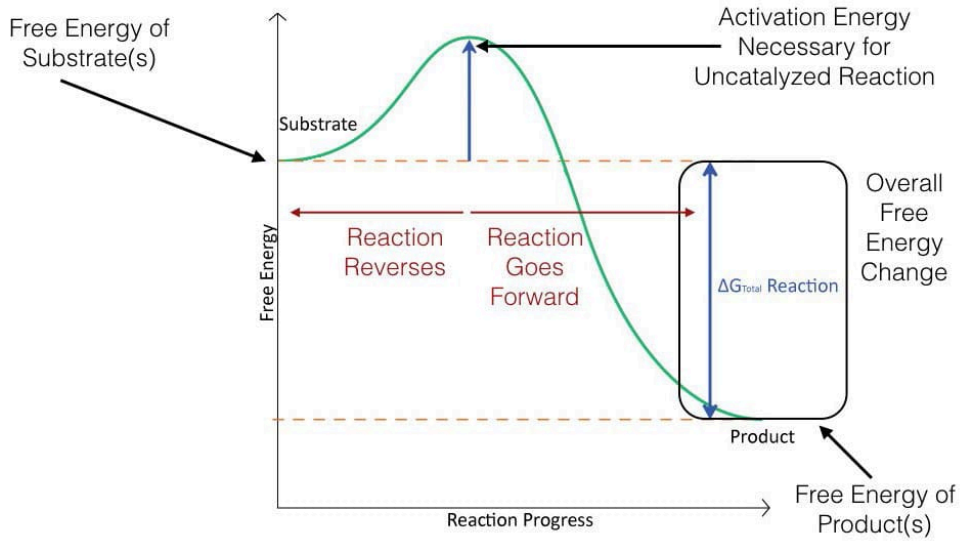


Figure 4.2 - Energy changes during the course of an uncatalyzed reaction. Image by Aleia Kim

In Figure 4.3, the activation energy for a catalyzed reaction is overlaid. As you can see, the reactants start at the same energy level for both catalyzed and uncatalyzed reactions and that the products end at the same energy for both as well. The catalyzed reaction, however, has a lower energy of activation (dotted line) than the uncatalyzed reaction. This is the secret to catalysis - overall ΔG for a reaction does NOT change with catalysis, but the activation energy is lowered.

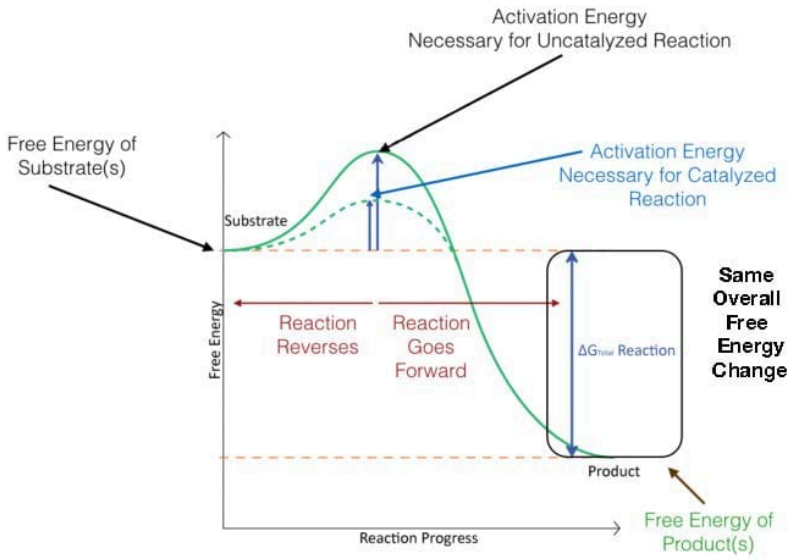


Figure 4.3 - Energy changes during the course of an uncatalyzed reaction (solid green line) and a catalyzed reaction (dotted green line). Image by Aleia Kim

Reversibility

The extent to which reactions will proceed forward is a function of the size of the energy difference between the product and reactant states. The lower the energy of the products compared to the reactants, the larger the percentage of molecules that will be present as products at equilibrium. It is worth noting that since an enzyme lowers the activation energy for a reaction that it can speed the reversal of a reaction just as it speeds a reaction in the forward direction. At equilibrium, of course, no change in concentration of reactants and products occurs. Thus, enzymes speed the time required to reach equilibrium, but do not affect the balance of products and reactants at equilibrium.

Exceptions

The reversibility of enzymatic reactions is an important consideration for equilibrium, the measurement of enzyme kinetics, for Gibbs free energy, for metabolic pathways, and for physiology. There are some minor exceptions to the reversibility of reactions, though. They are related to the disappearance of a substrate or product of a reaction. Consider the first reaction below which is catalyzed by the enzyme carbonic anhydrase:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^- + H^+ \tag{4.1.8}$$

In the forward direction, carbonic acid is produced from water and carbon dioxide. It can either remain intact in the solution or ionize to produce bicarbonate ion and a proton. In the reverse direction, water and carbon dioxide are produced. Carbon dioxide, of course, is a gas and can leave the solution and escape.

When reaction molecules are removed, as they would be if carbon dioxide escaped, the reaction is pulled in the direction of the molecule being lost and reversal cannot occur unless the missing molecule is replaced. In the second reaction occurring on the right, carbonic acid (H2CO3) is “removed” by ionization. This too would limit the reaction going back to carbon dioxide in water. This last type of “removal” is what occurs in metabolic pathways. In this case, the product of one reaction (carbonic acid) is the substrate for the next (formation of bicarbonate and a proton).

In the metabolic pathway of glycolysis, ten reactions are connected in this manner and reversing the process is much more complicated than if just one reaction was being considered.

General mechanisms of action

As noted above, enzymatically catalyzed reactions are orders of magnitude faster than uncatalyzed and chemical-catalyzed

reactions. The secret of their success lies in a fundamental difference in their mechanisms of action.

Every chemistry student has been taught that a catalyst speeds a reaction without being consumed by it. In other words, the catalyst ends up after a reaction just the way it started so it can catalyze other reactions, as well. Enzymes share this property, but in the middle, during the catalytic action, an enzyme is transiently changed. In fact, it is the ability of an enzyme to change that leads to its incredible efficiency as a catalyst.

Changes

These changes may be subtle electronic ones, more significant covalent modifications, or structural changes arising from the flexibility inherent in enzymes, but not present in chemical catalysts. Flexibility allows movement and movement facilitates alteration of electronic environments necessary for catalysis. Enzymes are, thus, much more efficient than rigid chemical catalysts as a result of their abilities to facilitate the changes necessary to optimize the catalytic process.

Substrate binding

Another important difference between the mechanism of action of an enzyme and a chemical catalyst is that an enzyme has binding sites that not only 'grab' the substrate (molecule involved in the reaction being catalyzed), but also place it in a position to be electronically induced to react, either within itself or with another substrate.

The enzyme itself may play a role in the electronic induction or the induction may occur as a result of substrates being placed in very close proximity to each other. Chemical catalysts have no such ability to bind substrates and are dependent upon them colliding in the right orientation at or near their surfaces.

Active site

Reactions in an enzyme are catalyzed at a specific location within it known as the 'active site'. Substrates bind at the active site and are oriented to provide access for the relevant portion of the molecule to the electronic environment of the enzyme where catalysis occurs.

Enzyme flexibility

As mentioned earlier, a difference between an enzyme and a chemical catalyst is that an enzyme is flexible. Its slight changes in shape (often arising from the binding of the substrate itself) help to optimally position substrates for reaction after they bind.

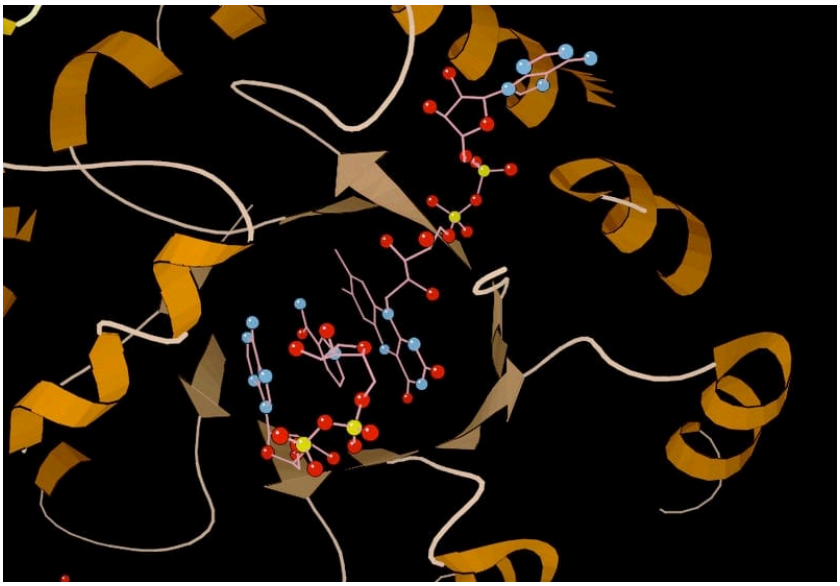


Figure 4.4 - Substrate binding by methylenetetrahydrofolate reductase

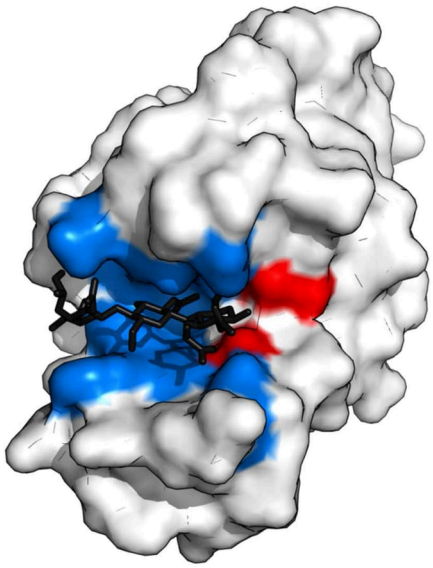


Figure 4.5 - Lysozyme with substrate binding site (blue), active site (red) and bound substrate (black). Wikipedia

Induced fit

These changes in shape are explained, in part, by Koshland's Induced Fit Model of Catalysis (Figure 4.6), which illustrates that not only do enzymes change substrates, but that substrates also transiently change enzyme structure. At the end of the catalysis, the enzyme is returned to its original state. Koshland's model is in contrast to the Fischer Lock and Key model, which says simply that an enzyme has a fixed shape that is perfectly matched for binding its substrate(s). Enzyme flexibility also is important for control of enzyme activity. Enzymes alternate between the T (tight) state, which is a lower activity state and the R (relaxed) state, which has greater activity.

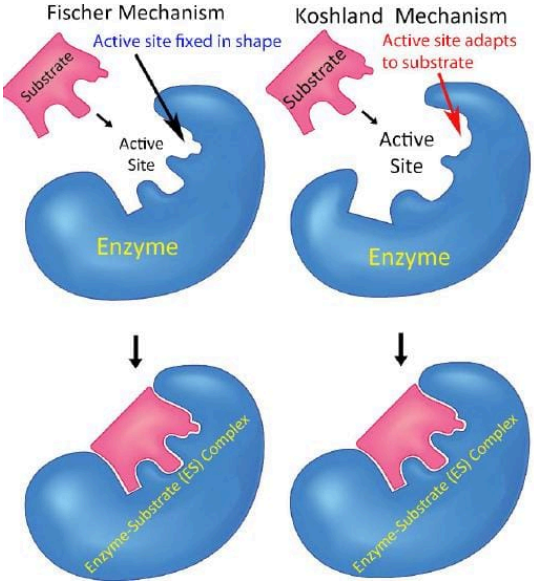


Figure 4.6 - Fischer's lock and key model (left) Vs. Koshland's induced fit model (right). Image by Aleia Kim

Induced Fit

The Koshland Induced Fit model of catalysis postulates that enzymes are flexible and change shape on binding substrate. Changes in shape help to 1) aid binding of additional substrates in reactions involving more than one substrate and/or 2) facilitate formation of an electronic environment in the enzyme that favors catalysis. This model is in contrast to the Fischer Lock and Key Model of catalysis which considers enzymes as having pre-formed substrate binding sites.

Ordered binding

The Koshland model is consistent with multi-substrate binding enzymes that exhibit ordered binding of substrates. For these systems, binding of the first substrate induces structural changes in the enzyme necessary for binding the second substrate.

There is considerable experimental evidence supporting the Koshland model. Hexokinase, for example, is one of many enzymes known to undergo significant structural alteration after binding of substrate. In this case, the two substrates are brought into very close proximity by the induced fit and catalysis is made possible as a result.

Reaction types

Enzymes that catalyze reactions involving more than one substrate, such as

$$A + B \rightleftharpoons C + D \tag{4.1.9}$$

can act in two different ways. Enzymatic reactions can be of several types, as shown in Figure 4.7. In one mechanism, called sequential reactions, at some point in the reaction, both substrates will be bound to the enzyme. There are, in turn, two different ways in which this can occur - random and ordered.

Figure 4.7 - Categories of enzymatic reactions

Types of Reactions	
Single Substrate - Single Product	A ⇄ B
Single Substrate - Multiple Products	A ⇄ B + C
Multiple Substrates - Single Products	A + B ⇄ C
Multiple Substrates - Multiple Products	A + B ⇄ C + D

Consider lactate dehydrogenase, which catalyzes the reaction below:

$$NADH + Pyruvate \rightarrow Lactate + NAD^{+} \tag{4.1.10}$$

This enzyme requires that NADH must bind prior to the binding of pyruvate. As noted earlier, this is consistent with an induced fit model of catalysis. In this case, binding of the NADH changes the enzyme shape/environment so that pyruvate can bind and without binding of NADH, the substrate cannot access the pyruvate binding site. This type of multiple substrate reaction is called **sequential ordered binding**, because the binding of substrates must occur in the right order for the reaction to proceed.

Random binding

A second mechanism of binding/catalysis is exhibited by creatine kinase which catalyzes the following reaction:

$$Creatine + ATP \rightarrow \text{Creatine phosphate} + ADP \tag{4.1.11}$$

For this enzyme, substrates can bind to it in any order. Creatine kinase displays sequential random binding. It is worth mentioning that random binding is not inconsistent with Koshland's induced fit model. Rather, random binding simply means that the enzyme's induced fit doesn't affect substrate binding sites and involves other parts of the enzyme. In summary, sequential binding can occur as ordered binding or as random binding.

Double displacement reactions

Not all enzymes that catalyze multi-substrate reactions, though, bind A and B by the sequential mechanisms above. This other category of enzyme includes those that exhibit what are called "ping-pong" (or double displacement) mechanisms. In these enzymes, the enzyme functions as both a catalyst and a carrier of a group between individually bound substrates. Examples of this type of enzyme include the class of enzymes known as transaminases. A general form of the reactions catalyzed by these enzymes is shown in Figure 4.8.

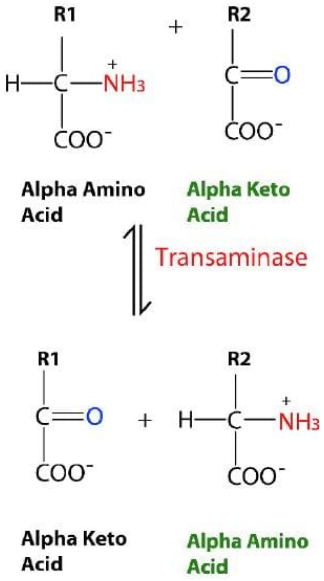


Figure 4.8 - Double displacement reaction mechanism of a transaminase

In reversible transaminase reactions, an oxygen and an amine are swapping between the molecules. It can be represented as follows (where N is the amine and O is the oxygen).

$$A{=}O + C{=}N \rightleftarrows B{-}N + D{=}O$$

This reaction occurs not as one transfer reaction swapping the N and the O, but rather as a set of two half-reactions. In this case, the enzyme is a both donor and a carrier of the group being swapped. The first half-reaction goes as follows

$$A{=}O + Enz{-}N \rightleftarrows B{-}N + Enz{=}O$$

Next a second half-reaction goes as

$$C{-}N + Enz{=}O \rightleftarrows D{=}O + Enz{=}N$$

The sum of these half-reactions then is

$$A{=}O + C{=}N \rightleftarrows B{-}N + D{=}O$$

Note that at no time did the enzyme bind both A and C simultaneously. It is also important to recognize that the enzyme existed in two states - Enz=O and Enz-N. The shuffling of the enzyme between these two states is what gives rise to the ping-pong name of this mechanism - it literally goes back and forth like a ping-pong ball in a table tennis match.

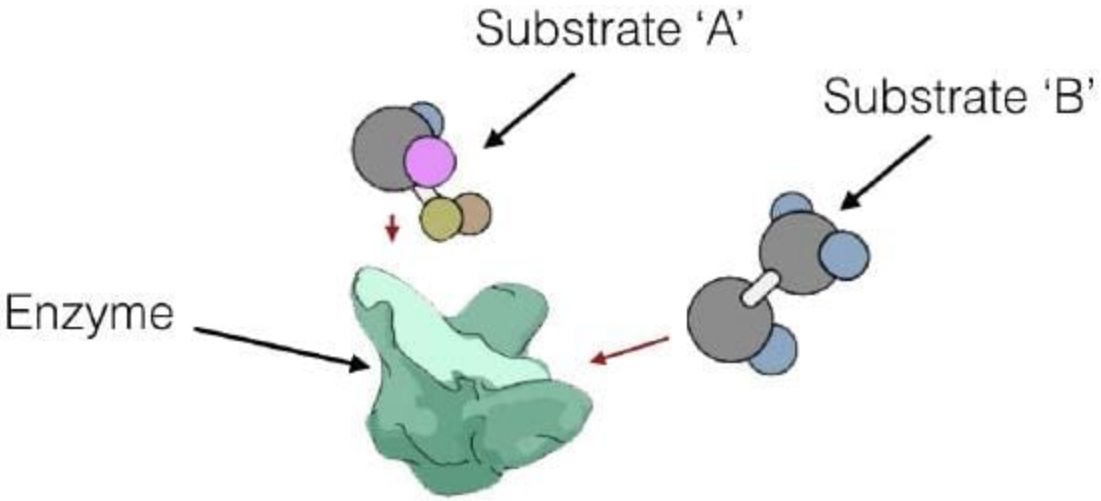


Figure 4.9 - Binding of substrates (A+B) to free enzyme

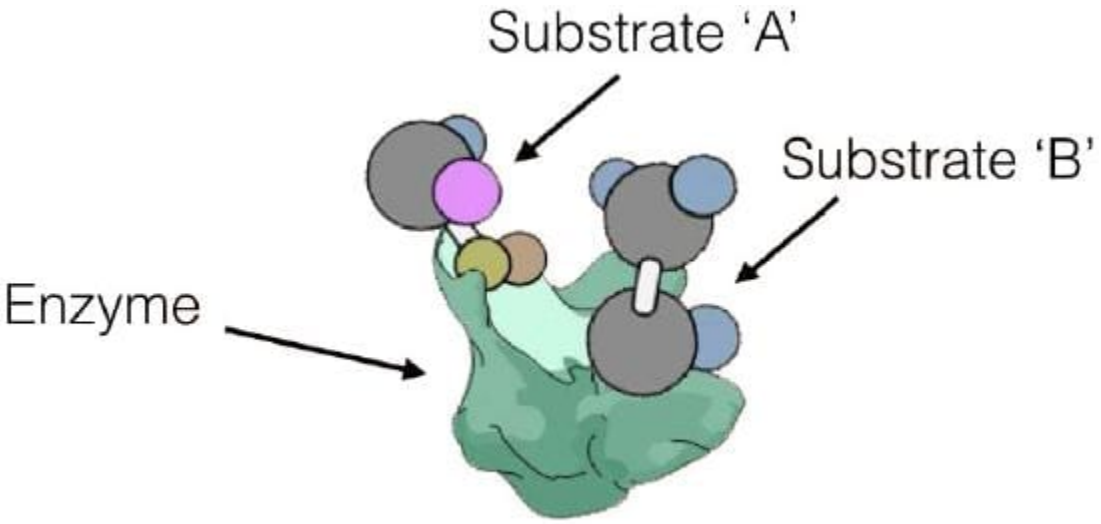


Figure 4.10 ES complex formation

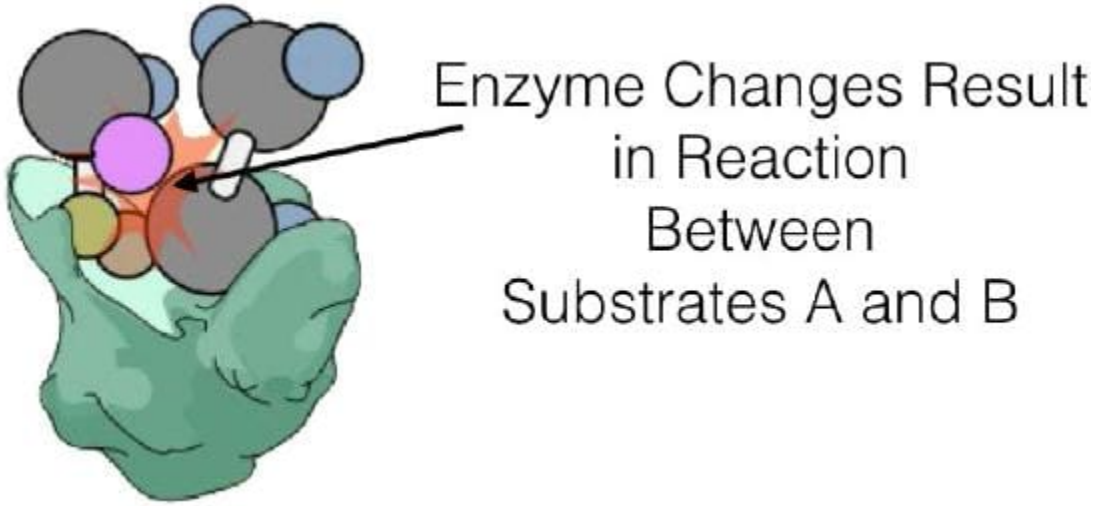


Figure 4.11 - ES* complex - The reaction occurs.

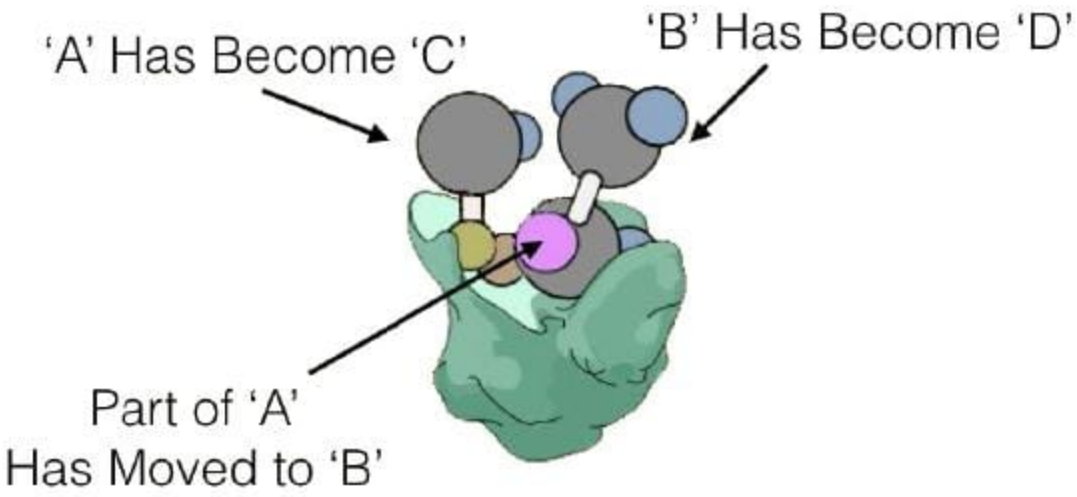


Figure 4.12 - EP complex - the reaction is complete.

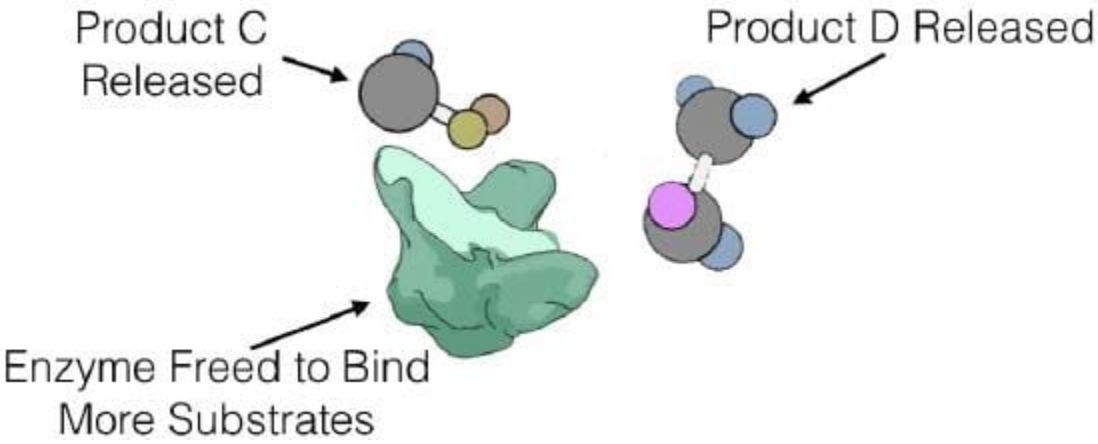


Figure 4.13 - E + P - the products are released

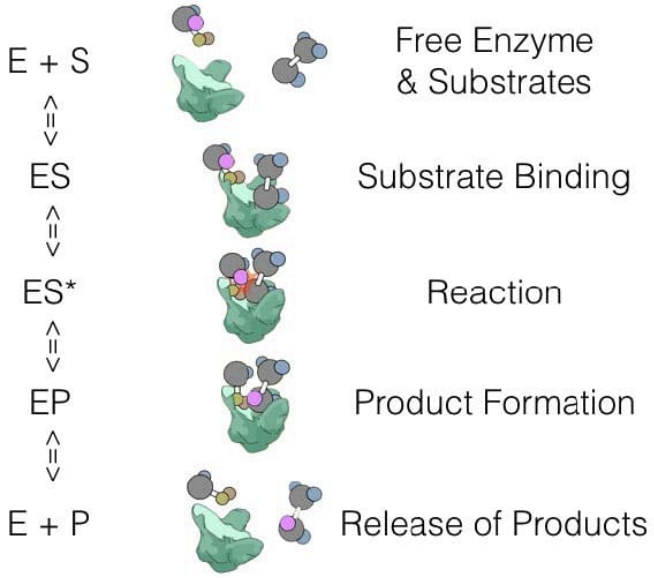


Figure 4.14 - Summary

Enzyme kinetics

To understand how an enzyme enhances the rate of a reaction, we must understand enzyme kinetics. We present a model

here proposed by Leonor Michaelis and Maud Menten. In order to understand the model, it is necessary to understand a few parameters.

First, we describe a reaction in simple terms proceeding as follows

E + S ⇄ ES -> E + P

where E is enzyme, S is substrate, and P is product. In this scheme, ES is the Enzyme-Substrate complex, which is simply the enzyme bound to its substrate.

We could define the ES state a bit further with

E + S ⇄ ES -> ES* -> EP -> E + P

where ES* is the activated state and EP is the enzyme-product complex before release of the product.

The first consideration we have is velocity. The velocity of a reaction is the rate of creation of product over time, measured as the concentration of product per time. The time is a critical consideration when measuring velocity. In a closed system (in which an enzyme operates), all reactions will advance towards equilibrium. Enzymatically catalyzed reactions are no different in the end result from non-enzymatic reactions, except that they get to equilibrium faster.

Equilibrium

At equilibrium, the ratio of product to reactant does not change. That is a property of equilibrium. Since the system is closed, the concentration of product over time will not change. The velocity will thus be zero under these conditions and we will have learned nothing about the reaction if we wait too long to study it.

Velocity

Enzyme	Turnover Number (per second)
Carbonic anhydrase	600,000
3-Ketoesteroid isomerase	280,000
Acetylcholinesterase	25,000
Penicillinase	2,000
Lactate dehydrogenase	1,000
Chymostrypsin	100
DNA Polymerase I	15
Tryptophan synthetase	2
Lysozyme	0.5

Figure 4.22 - Kcat/Km values for perfect enzymes. Image by Aleia Kim

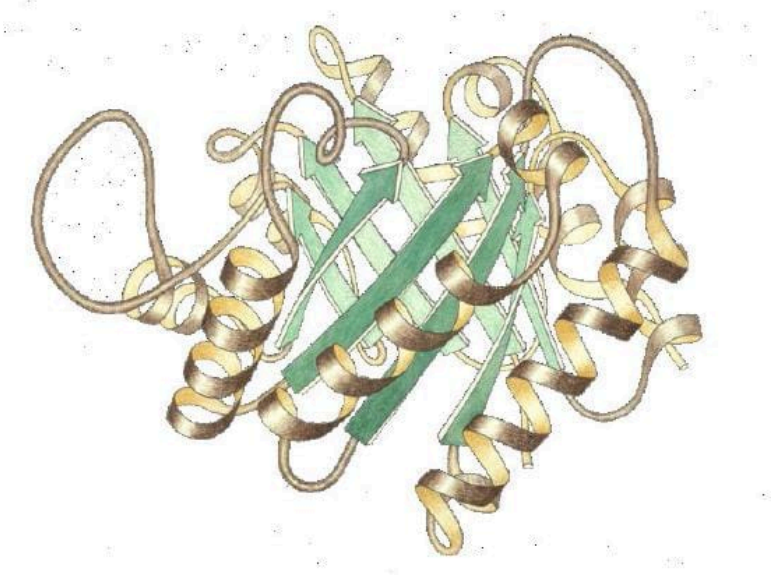


Figure 4.23 - Triose phosphate isomerase - A perfect enzyme. Wikipedia

Consequently, in Michaelis-Menten kinetics, velocity is measured as initial velocity (V0). This is accomplished by measuring the rate of formation of product early in the reaction before equilibrium is established and under these conditions, there is very little if any of the reverse reaction occurring.

The other two assumptions are related. First, we use conditions where there is much more substrate than enzyme. This makes sense. If the substrate is not in great excess, then the enzyme's conversion of substrate to product will occur much faster than the enzyme can bind substrate.

Waiting for substrate

Thus, the enzyme would "wait" for substrate to bind if there were not sufficient amounts of it to bind to the enzyme in a timely fashion (when substrate concentration is low). This would not give an accurate measure of velocity, since the enzyme would be inactive a good deal of the time. Because of this, we assume saturation of the enzyme with substrate will give a maximal velocity of the reaction.

Steady state

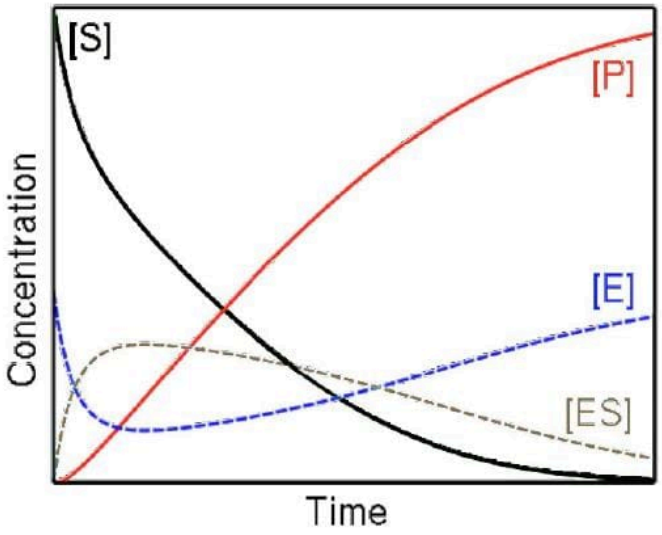


Figure 4.15 - Concentration of product (P), substrate (S), enzyme (E), and enzyme-substrate complex (ES) versus time for an enzymatic reaction

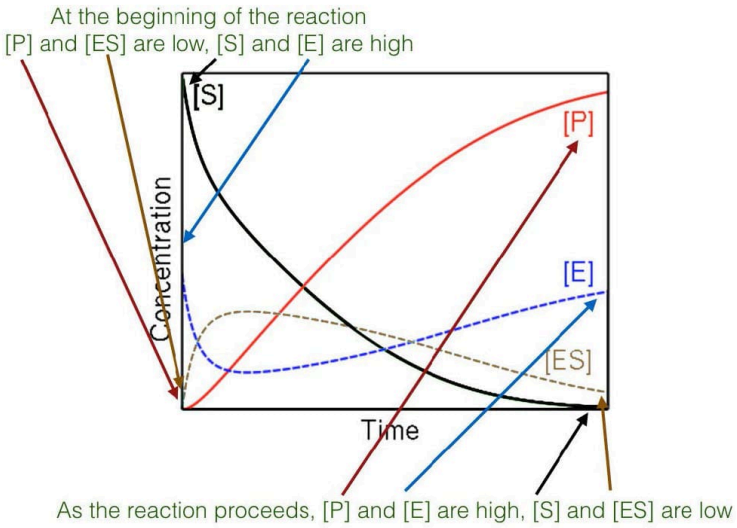


Figure 4.16 - Change in concentration of reaction materials over time

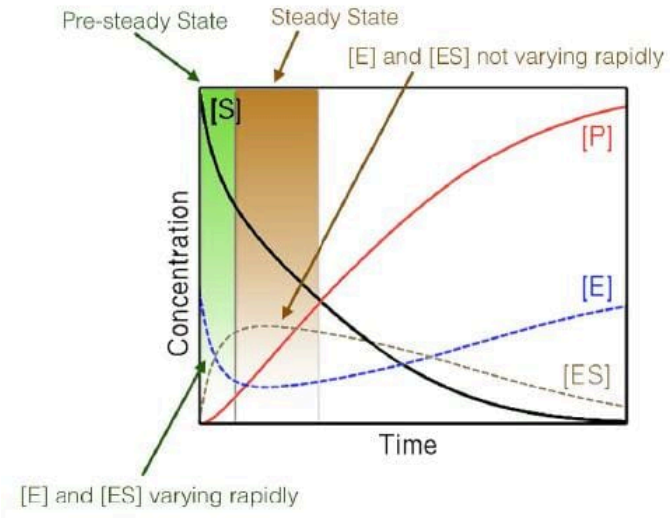


Figure 4.17 - Steady state versus non-steady state conditions

Last, the high concentration of substrate combined with measuring initial conditions results in studying reactions that are under so-called steady state conditions (Figure 4.17). When steady state occurs, the concentration of the ES complex over time is not significantly changing during the period of analysis.

Reiterating, the three assumptions for Michaelis-Menten kinetics are

1. Measurement of initial velocity of a reaction
2. Substrate in great excess compared to enzyme
3. Reaction conditions occurring under steady state

Experimental considerations

Now we turn our attention to how studies of the kinetic properties of an enzyme are conducted. To perform an analysis, one would do the following experiment - 20 different tubes would be set up with enzyme buffer (to keep the enzyme stable), the same amount of enzyme, and then a different amount of substrate in each tube, ranging from tiny amounts in the first tubes to very large amounts in the last tubes. The reaction would be allowed to proceed for a fixed, short amount of time and then the reaction

would be stopped and the amount of product contained in each tube would be determined.

The initial velocity (V0) of the reaction then would be the concentration of product found in each tube divided by the time that the reaction was allowed to run. Data from the experiment would be plotted on a graph using initial velocity (V0) on the Y-axis and the concentration of substrate on the X-axis, each tube, of course having a unique reaction velocity corresponding to a unique substrate concentration.

For an enzyme following Michaelis-Menten kinetics, a curve like that shown in Figure 4.18 or 4.19 would result. At low concentration of substrate, it is limiting and the enzyme converts it into product as soon as it can bind it. Consequently, at low concentrations of substrate, the rate of increase of [P] is almost linear with [S] (Figure 4.19).

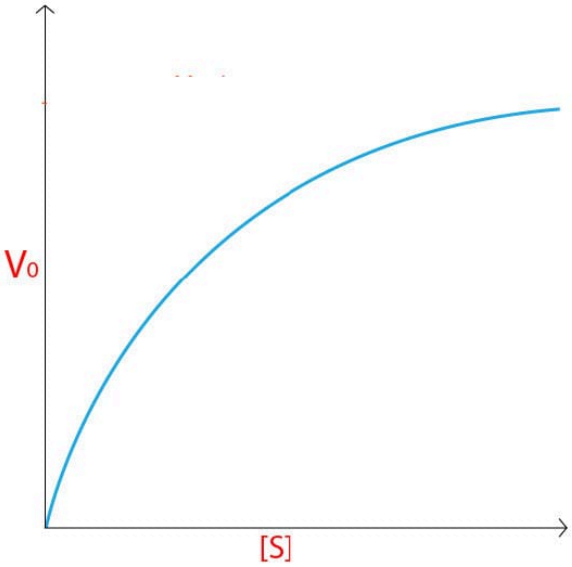


Figure 4.18 - Kinetics of an enzyme obeying Michaelis-Menten kinetics. Image by Aleia Kim

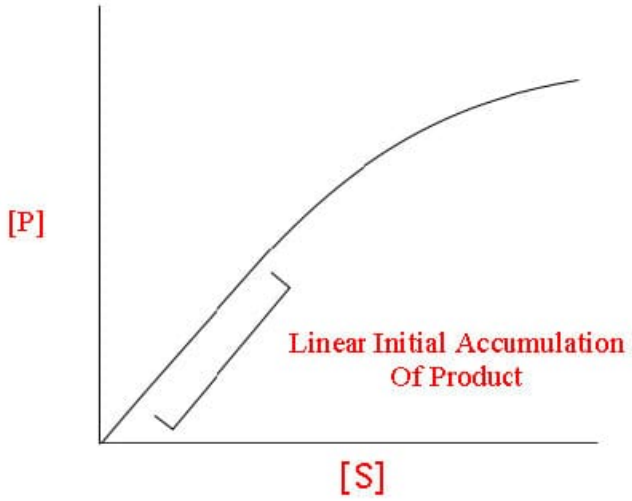


Figure 4.19 - Linear relationship between [P] and [S] at low [S]

Non-linear increase

As the substrate concentration increases, however, the velocity of the reaction in tubes with higher substrate concentration ceases to increase linearly and instead begins to flatten out, indicating that as the substrate concentration gets higher and higher, the enzyme has a harder time keeping up to convert the substrate to product.

Saturation

Not surprisingly, when the enzyme becomes completely saturated with substrate, it will not have to wait for substrate to diffuse to it and will therefore be operating at maximum velocity.

For an enzyme following Michaelis-Menten kinetics will have its velocity (v) at any given substrate concentration given by the following equation:

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

V_{max}

Two terms in the equation above require explanation. The first is V_{max}. It refers to the maximum velocity of an enzymatic reaction. Maximum velocity for a reaction occurs when an enzyme is saturated with substrate. Saturation is important because it means (per the assumption above) that none of the enzyme molecules are "waiting" for substrate after a product is released. Saturation ensures that another substrate is always instantly available. The unit of V_{max} is concentration of product per time = [P]/time.

On a plot of initial velocity versus substrate concentration (V0 vs. [S]), V_{max} is the value on the Y axis that the curve asymptotically approaches (dotted line in Figure 4.20). It should be noted that the value of V_{max} depends on the amount of enzyme used in a reaction. If you double the amount of enzyme used, you will double the V_{max}. If one wanted to compare the velocities of two different enzymes, it would be necessary to use the same amounts of enzyme in the reaction each one catalyzes.

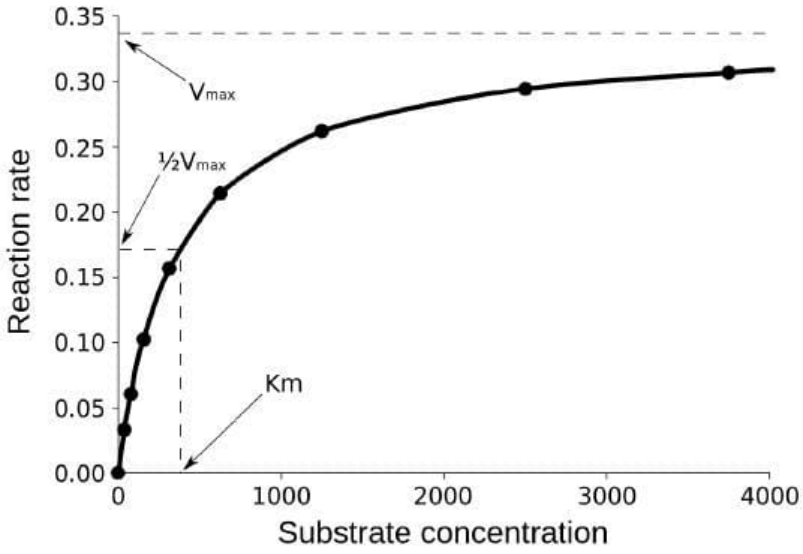


Figure 4.20 - V_{max}, $V_{max}/2$, and Km

Km

The second term is Km (also known as Ks). Referred to as the Michaelis constant, Km is the substrate concentration that causes the enzyme to work at half of maximum velocity (Vmax/2). What it measures, in simple terms, is the affinity an enzyme has for its substrate. The value of Km is inversely related to the affinity of the enzyme for its substrate. Enzymes with a high Km value will have a lower affinity for their substrate (will take more substrate to get to Vmax/2) whereas those with a low Km will have high affinity and take less substrate to get to Vmax/2. The unit of Km is concentration.

Affinities of enzymes for substrates vary considerably, so knowing Km helps us to understand how well an enzyme is suited to the substrate being used. Measurement of Km depends on the measurement of Vmax.

Common mistake

A common mistake students make in describing Vmax is saying that Km = Vmax /2. This is, of course, not true. Km is a substrate concentration and is the amount of substrate it takes for an enzyme to reach Vmax /2. On the other hand Vmax /2 is a velocity and a velocity certainly cannot equal a concentration.

Kcat

It is desirable to have a measure of velocity that is independent of enzyme concentration. Remember that Vmax depended on the amount of enzyme used. For this, we use the Kcat, also known as the turnover number. Kcat is a number that requires one to first determine Vmax for an enzyme and then divide the Vmax by the concentration of enzyme used to determine Vmax. Thus,

Kcat = Vmax /[Enzyme]

Since Vmax has units of concentration per time and [Enzyme] has units of concentration, the units on Kcat are time-1. While that might seem unintuitive, it means that the value of Kcat is the number of molecules of product made by each molecule of enzyme in the time given. So, a Kcat value of 1000/sec means each enzyme molecule in the reaction at Vmax is producing 1000 molecules of product per second. Note that since Kcat is a calculated value, it cannot be read from a V vs [S] graph as Vmax and Km can.

Enzyme	Turnover Number (per second)
Carbonic anhydrase	600,000
3-Ketoesteroid isomerase	280,000
Acetylcholinesterase	25,000
Penicillinase	2,000
Lactate dehydrogenase	1,000
Chymostrypsin	100
DNA Polymerase I	15
Tryptophan synthetase	2
Lysozyme	0.5

Figure 4.21 - Kcat (turnover number) values for several enzymes. Image by Aleia Kim

Amazing Kcat values

A Kcat value of 1000 molecules of product per enzyme per second might seem like a high value, but there are enzymes known (carbonic anhydrase, for example) that have a Kcat value of over 600,000/second (Figure 4.21). This astonishing value illustrates clearly why enzymes seem almost magical in their action. In contrast to V_{max}, which varies with the amount of enzyme used, Kcat is a constant for an enzyme under given conditions.

As seen earlier, enzymes that follow Michaelis-Menten kinetics produce hyperbolic plots of Velocity (V0) versus Substrate Concentration [S] (Figure 4.18). Not all enzymes, though, follow Michaelis-Menten kinetics. Many enzymes have multiple protein subunits and these sometimes interact differently upon binding of a substrate or an external molecule. See ATCase (HERE) for an example.

Perfect enzymes

Now, if we think about what an ideal enzyme might be, it would be one that has a very high velocity and a very high affinity for

its substrate. That is, it wouldn't take much substrate to get to Vmax/2 and the Kcat would be very high. Such enzymes would have values of Kcat / Km that are maximum. Interestingly, there are several enzymes that have this property and their maximal Kcat / Km values are all approximately the same. Such enzymes are referred to as being "perfect" because they have reached the maximum possible value.

Enzyme	K_{Cat} / K_M ($s^{-1}M^{-1}$)
Acetylcholinesterase	1.6×10^8
Carbonic anhydrase	8.3×10^7
Catalase	4.0×10^7
Crotonase	2.8×10^8
Fumarase	1.6×10^8
Triose phosphate isomerase	2.4×10^8
β-Lactamase	1.0×10^8
Superoxide dismutase	7.0×10^9

Figure 4.22 - Kcat/Km values for perfect enzymes. Image by Aleia Kim

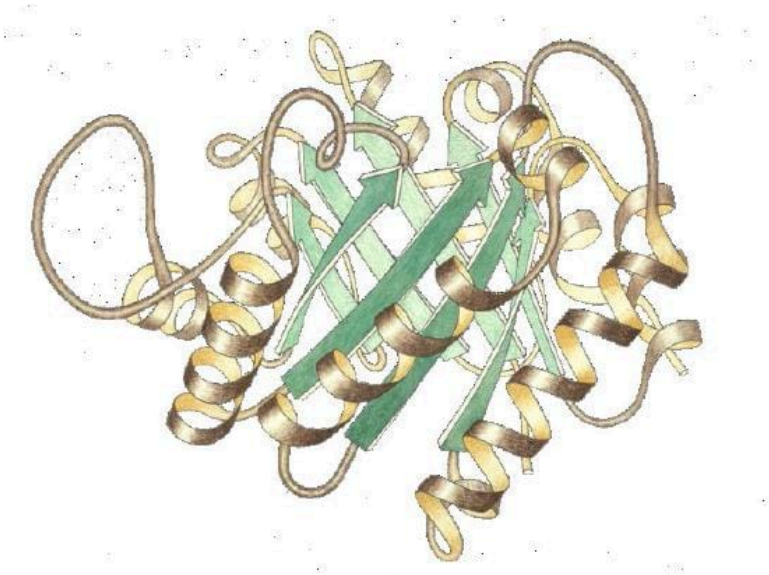


Figure 4.23 - Triose phosphate isomerase - A perfect enzyme. Wikipedia

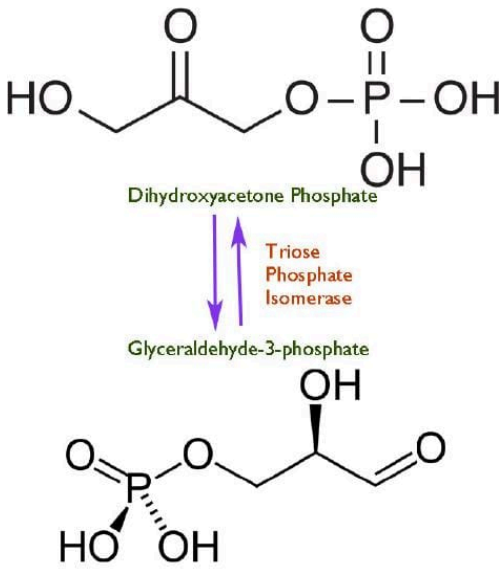


Figure 4.24 - Triose phosphate isomerase-catalyzed reaction

Diffusion limitation

Why should there be a maximum possible value of Kcat / Km? The answer is that movement of substrate to the enzyme becomes the limiting factor for perfect enzymes. Movement of substrate by diffusion in water has a fixed rate at any temperature and that limitation ultimately determines the maximum speed an enzyme can catalyze at. In a macroscopic world analogy, factories can't make products faster than suppliers can deliver materials. It is safe to say for a perfect enzyme that the only speed limit it has is the rate of substrate diffusion in water.

Given the "magic" of enzymes alluded to earlier, it might seem that all enzymes should have evolved to be "perfect." There are very good reasons why most of them have not.

Speed

Speed is a dangerous thing. The faster a reaction proceeds in catalysis by an enzyme, the harder it is to control. As we all know from learning to drive, speeding causes accidents. Just as drivers need to have speed limits for operating automobiles, so too must cells exert some control on the 'throttle' of their enzymes. In view of this, one might wonder then why any cells have evolved any enzymes to perfection. There is no single answer to the question, but a common one is illustrated by triose phosphate isomerase, which catalyzes a reaction in glycolysis shown in Figure 4.24.

The enzyme appears to have evolved this ability because at lower velocities, there is breakdown of an unstable enediol intermediate that then readily forms methyl glyoxal, a cytotoxic compound (Figure 4.25). Speeding up the reaction provides less opportunity for the unstable intermediate to accumulate and fewer undesirable byproducts to be made.

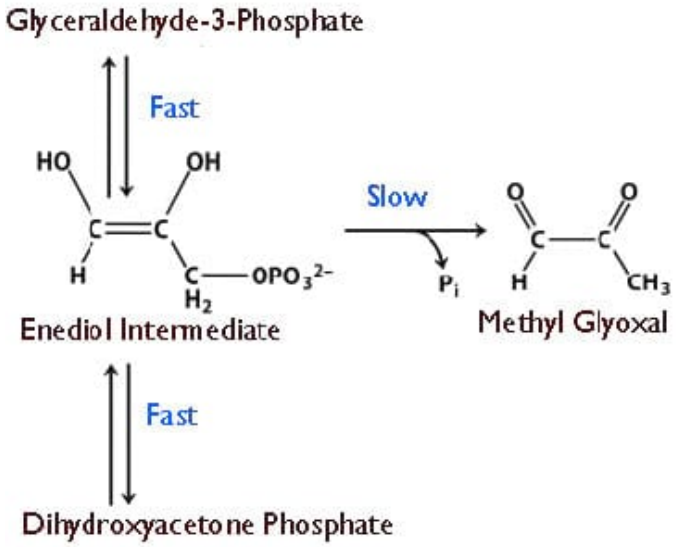


Figure 4.25 - A high speed reaction avoids production of methyl glyoxal. Image by Aleia Kim

Dissociation constant

In studying proteins and ligands, it is important to understand the “tightness” with which a protein (P) “holds onto” a ligand (L). This is measured with the dissociation constant (K_d). The formation of a ligand-protein complex LP occurs as

$$L + P \rightleftharpoons LP \tag{4.1.12}$$

The dissociation of the complex, therefore, is the reverse of this reaction, or

$$LP \rightleftharpoons L + P \tag{4.1.13}$$

so the corresponding dissociation constant is defined as

$$K_d = \frac{[L][P]}{[LP]} \tag{4.1.14}$$

where $[L]$, $[P]$, and $[LP]$ are the molar concentrations of the protein, ligand and the complex when they are joined together.

Smaller values of K_d indicate tight binding, whereas larger values indicate loose binding. The dissociation constant is the inverse of the association constant.

$$K_a = \frac{1}{K_d} \tag{4.1.15}$$

Where multiple molecules bond together, such as

$$J_x K_y \rightleftharpoons xJ + yK \tag{4.1.16}$$

The complex $J_x K_y$ is breaking down into x subunits of J and y subunits of K. The dissociation constant is then defined as

$$K_d = \frac{[J]^x [K]^y}{[J_x K_y]} \tag{4.1.17}$$

where $[J]$, $[K]$, and $[J_x K_y]$ are the concentrations of J, K, and the complex $J_x K_y$, respectively.

Lineweaver-Burk plots

The study of enzyme kinetics is typically the most math intensive component of biochemistry and one of the most daunting

aspects of the subject for many students. Although attempts are made to simplify the mathematical considerations, sometimes they only serve to confuse or frustrate students. Such is the case with modified enzyme plots, such as Lineweaver-Burk (Figure 4.26).

Indeed, when presented by professors as simply another thing to memorize, who can blame students? In reality, both of these plots are aimed at simplifying the determination of parameters, such as K_m and V_{max}. In making either of these modified plots, it is important to recognize that the same data is used as in making a V0 vs. [S] plot. The data are simply manipulated to make the plotting easier.

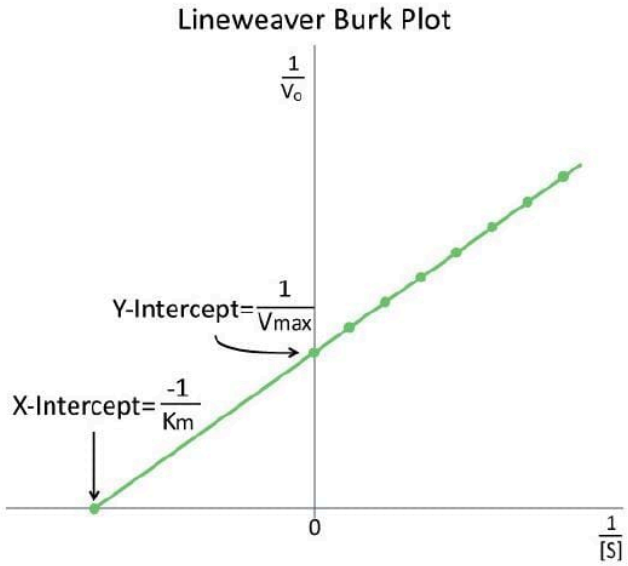


Figure 4.26 - A Lineweaver-Burk plot of $1/V_0$ vs $1/[S]$. Image by Aleia Kim

Double reciprocal

For a LineWeaver-Burk plot, the manipulation is using the reciprocal of the values of both the velocity and the substrate concentration. The inverted values are then plotted on a graph as 1/V0 vs. 1/[S]. Because of these inversions, Lineweaver-Burk plots are commonly referred to as 'double-reciprocal' plots. As can be seen in Figure 4.26, the value of Km on a Lineweaver Burk plot is easily determined as the negative reciprocal of the x-intercept , whereas the Vmax is the inverse of the y-intercept. Other related manipulation of kinetic data include Eadie-Hofstee diagrams, which plots V0 vs V0/[S] and gives Vmax as the Y-axis intercept with the slope of the line being -Km.

Molecularity of reactions

The term molecularity refers to the number of molecules that must come together in order for a reaction to take place. Reactions of the sort of A -> B (where 'A' is the reactant and 'B' is the product) are unimolecular, since A is directly changed into B. The rate of the reaction is related only to the concentration of reactant A. For a bimolecular reaction where A + B ⇄ C the reaction depends on the concentration of both A and B and its rate will be related to the product of the concentration of A and of B.

Coenzymes

Organic molecules that assist enzymes and facilitate catalysis are co-factors called coenzymes. The term co-factor is a broad category usually subdivided into inorganic ions and coenzymes. If the coenzyme is very tightly or covalently bound to the enzyme, it is referred to as a prosthetic group. Enzymes without their co-factors are inactive and referred to as apoenzymes. Enzymes containing all of their co-factors are called holoenzymes.

Cofactor	Enzyme
Coenzyme	
5′-Deoxyadenosyl cobalamin	Methylmalonyl mutase
Biotin	Pyruvate carboxylase
Coenzyme A (CoA)	Acetyl CoA carboxylase
Flavin adenine nucleotide	Monoamine oxidase
Nicotinamide adenine dinucleotide	Lactate dehydrogenase
Pyridoxal phosphate	Glycogen phosphorylase
Tetrahydrofolate	Thymidylate synthase
Thiamine pyrophosphate	Pyruvate dehydrogenase
Metal	
K^+	Propionyl CoA carboxylase
Mg^{2+}	Restriction Endonucleases; Hexokinase
Mn	Superoxide dismutase
Mo	Nitrate reductase
Ni^{2+}	Urease
Se	Glutathione peroxidase
Zn^{2+}	Carbonic anhydrase; Carboxypeptidase

Figure 4.27 - Enzyme cofactors. Image by Aleia Kim

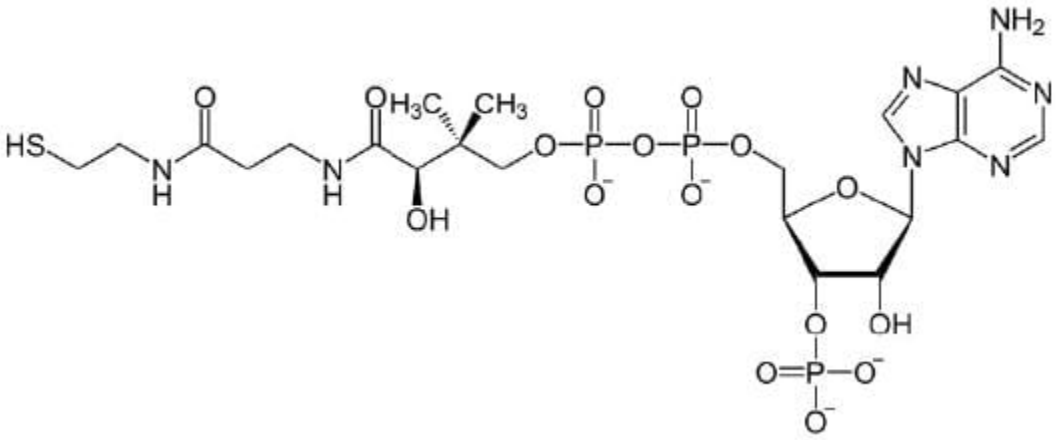


Figure 4.28 - Coenzyme A (CoA)

Pre-steady state kinetic studies

In the study of kinetic rates of enzymatic reactions, time zero is a very critical point. It establishes when the mixing of substrate with enzyme and measurement of formation of product begins. At time zero, there is no product. As shown in Figure 4.29, the appearance of product (on a short time scale) goes through an early burst phase with a steep slope for [product]/time and then changes.

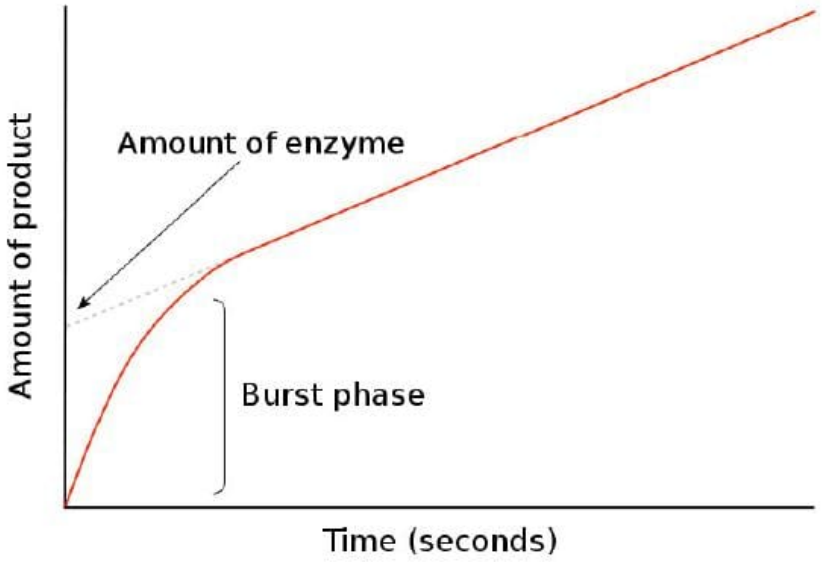


Figure 4.29 - Burst phase of product formation

This change occurs during a critical period in an enzymatic reaction and gives information about the rate of reaction cycles. The duration of the burst phase tells how long a single reaction turnover occurs, whereas the slow of the line post-burst phase tells the amount of “functional” enzyme performing the reaction.

After the burst phase, the slope of the line of the amount of product versus time decreases. This is due to the reaction entering conditions of steady state, used to study Michaelis-Menten kinetics. In steady state conditions, the amount of the enzyme-substrate complex (ES) is relatively constant over time. In simple terms, this occurs when the rate of formation of the ES complex equals the rate of conversion of the substrate to product by the enzyme with release of the product.

Earlier events

Events occurring prior to the conditions of steady state are referred to as pre-steady state. Depending on the enzyme, in as short as a few milliseconds, steady state conditions can be present meaning that if one hopes to study formation of reaction intermediates in pre-steady state, tools for this analysis must work very rapidly. One instrument commonly used for studying pre-steady state kinetics is called a stopped flow instrument.

It loads an enzyme solution and a substrate into separate syringes whose output is pointed into a mixing chamber. The solutions are rapidly mixed and measurements of product concentration begin. With a stopped flow instrument, dead times (time between mixing and detection) can be achieved of as small as 0.3 msec.

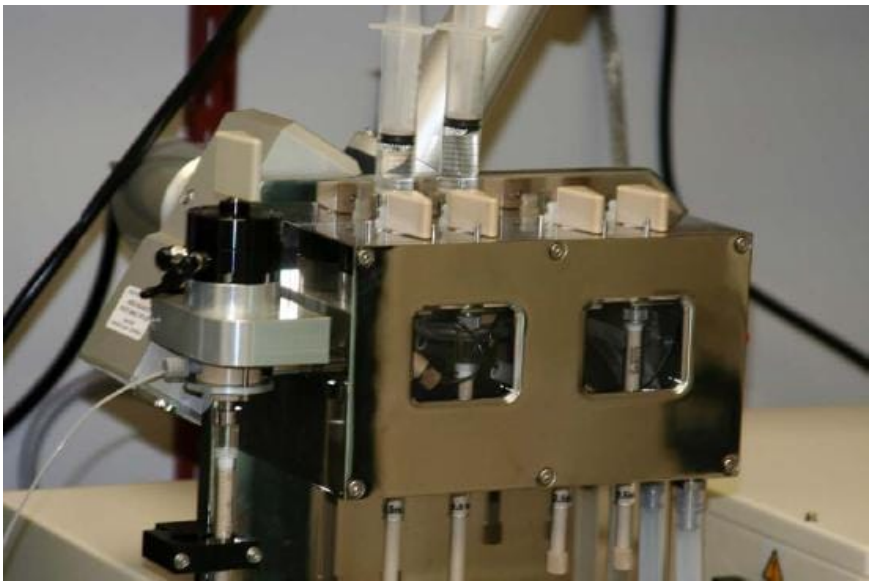
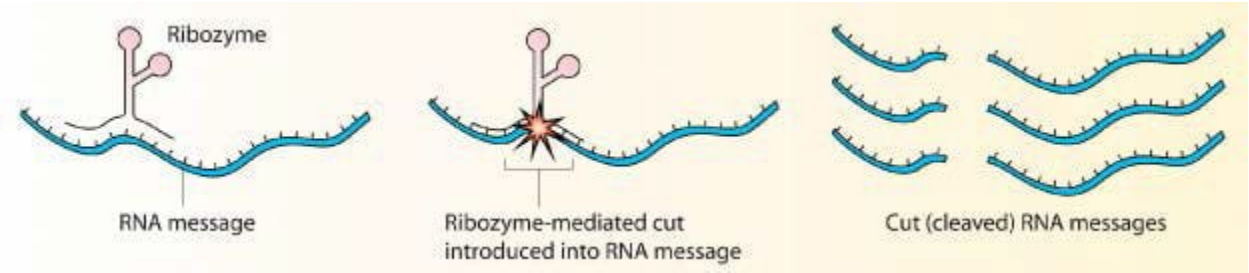


Figure 4.30 - Stopped flow instrument for studying pre-steady state kinetics. Wikipedia

Ribozymes

Proteins do not have a monopoly on acting as biological catalysts. Some RNA molecules are also capable of speeding reactions. The most famous of these molecules was discovered by Tom Cech in the early 1980s Studying excision of an intron in Tetrahymena, Cech was puzzled at his inability to find any proteins catalyzing the process. Ultimately, the catalysis was recognized as coming from the intron itself. It was a self-splicing RNA and since then, many other examples of catalytic RNAs have been found. Catalytic RNA molecules are known as ribozymes.



359

Figure 4.31 - Cleavage of an RNA by a ribozyme. Wikipedia

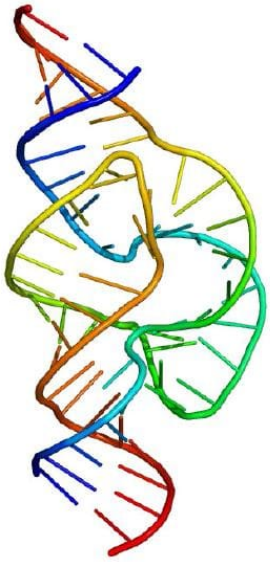


Figure 4.32 - Hammerhead ribozyme. Wikipedia

Not unusual

Ribozymes, however, are not rarities of nature. The protein-making ribosomes of cells are essentially giant ribozymes. The 23S rRNA of the prokaryotic ribosome and the 28S rRNA of the eukaryotic ribosome catalyze the formation of peptide bonds.

Ribozymes are also important in our understanding of the evolution of life on Earth. They have been shown to be capable via selection to evolve self-replication. Indeed, ribozymes actually answer a chicken/egg dilemma - which came first, enzymes that do the work of the cell or nucleic acids that carry the information required to produce the enzymes. As both carriers of genetic information and catalysts, ribozymes are likely both the chicken and the egg in the origin of life.

This page titled 4.1: Basic Principles of Catalysis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

4.2: Control of Enzymatic Activity

A printable version of this section is here: BiochemFFA_4_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Regulation of enzyme activity

Apart from their ability to greatly speed the rates of chemical reactions in cells, enzymes have another property that makes them valuable. This property is that their activity can be regulated, allowing them to be activated and inactivated, as necessary. This is tremendously important in maintaining homeostasis, permitting cells to respond in controlled ways to changes in both internal and external conditions.

Inhibition of specific enzymes by drugs can also be medically useful. Understanding the mechanisms that control enzyme activity is, therefore, of considerable importance.

Inhibition

We will first discuss four types of enzyme inhibition – competitive, non-competitive, uncompetitive, and suicide inhibition. Of these, the first three types are reversible. The last one, suicide inhibition, is not.

Competitive inhibition

Probably the easiest type of enzyme inhibition to understand is competitive inhibition and it is the one most commonly exploited pharmaceutically. Molecules that are competitive inhibitors of enzymes resemble one of the normal substrates of an enzyme. An example is methotrexate, which resembles the folate substrate of the enzyme dihydrofolate reductase (DHFR). This enzyme normally catalyzes the reduction of folate, an important reaction in the metabolism of nucleotides.

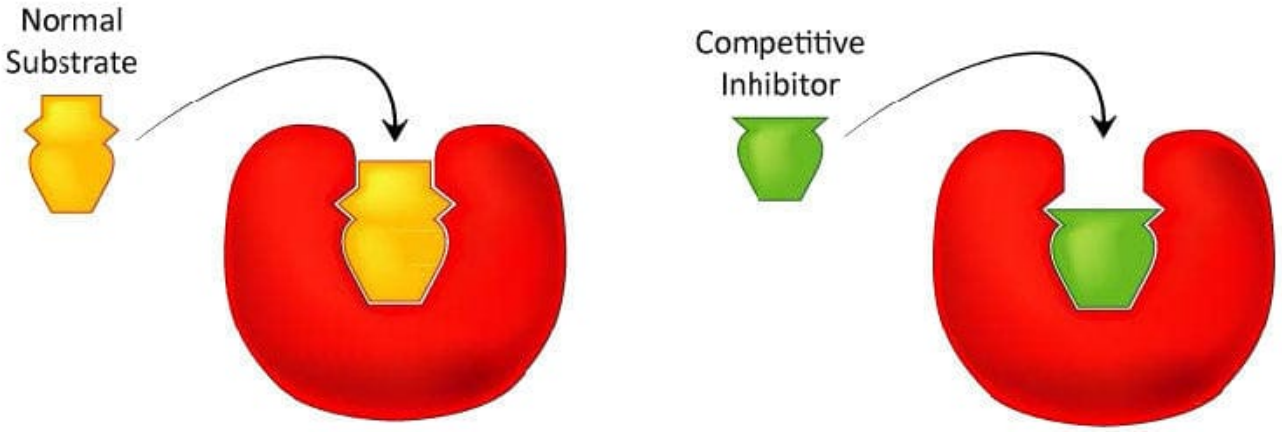


Figure 4.33 - Competitive inhibitors resemble the normal substrate and compete for binding at the active site. Image by Aleia Kim

Inhibitor binding

When the drug methotrexate is present, some of the DHFR enzyme binds to it, instead of to folate, and during the time methotrexate is bound, the enzyme is inactive and unable to bind folate. Thus, the enzyme is inhibited. Notably, the binding site on DHFR for methotrexate is the active site, the same place that folate would normally bind. As a result, methotrexate 'competes' with folate for binding to the enzyme. The more methotrexate there is, the more effectively it competes with folate for the enzyme's active site. Conversely, the more folate there is, the less of an effect methotrexate has on the enzyme because folate outcompetes

it.

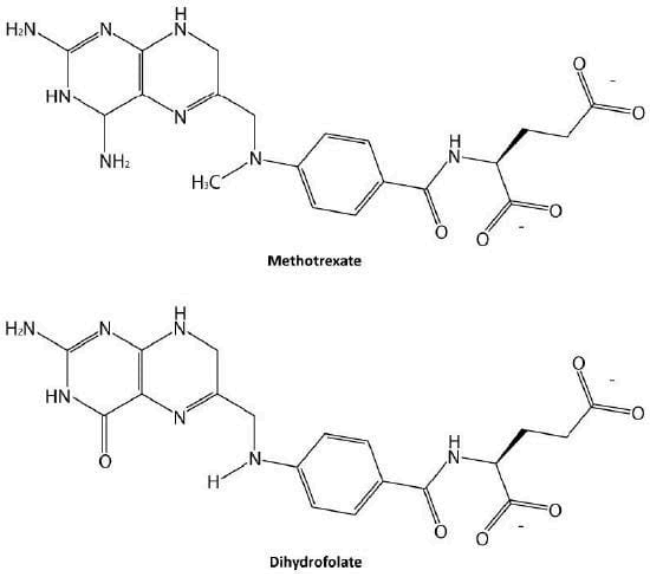


Figure 4.34 - Methotrexate and dihydrofolate. Image by Ben Carson

No effect on Vmax

How do we study competitive inhibition? It is typically done as follows. First, one performs a set of V0 vs. [S] reactions without inhibitor (20 or so tubes, with buffer and constant amounts of enzyme, varying amounts of substrate, equal reaction times). V0 vs. [S] is plotted (Figure 4.35 red line), as well as 1/V0 vs. 1/[S] (Figure 4.36 green line). Next, a second set of reactions is performed in the same manner as before, except that a fixed amount of the methotrexate inhibitor is added to each tube. At low concentrations of substrate, the methotrexate competes for the enzyme effectively, but at high concentrations of substrate, the inhibitor will have a much reduced effect, since the substrate outcompetes it, due to its higher concentration (remember that the inhibitor is at fixed concentration).

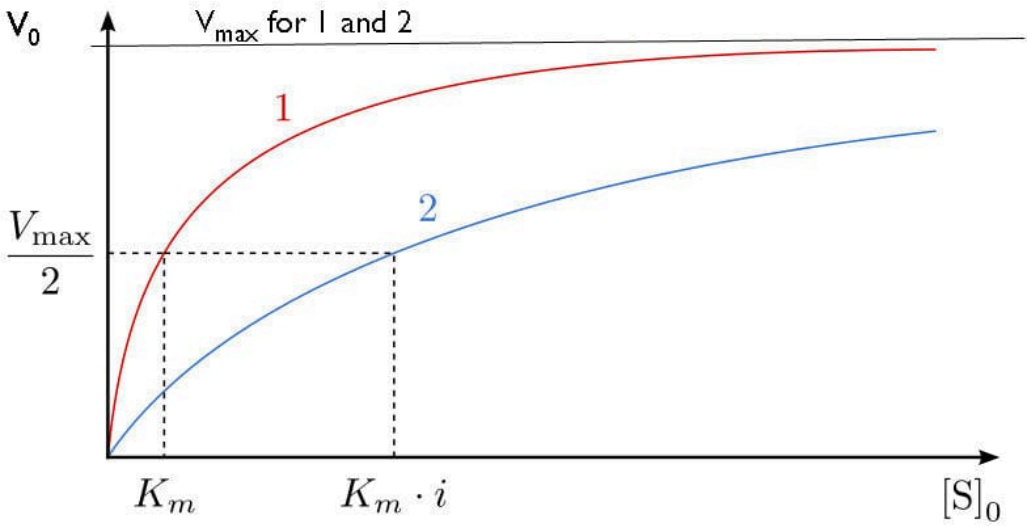


Figure 4.35 V0 vs [S] Plots for uninhibited reactions (red) and competitively inhibited reactions (blue). Both ultimately have same Vmax

Graphically, the results of these inhibitor experiments are shown in Figure 4.35 (blue line) and Figure 4.36 (orange line). Notice that at high substrate concentrations, the competitive inhibitor has essentially no effect, causing the V_{max} for the enzyme to remain unchanged. To reiterate, this is due to the fact that at high substrate concentrations, the inhibitor doesn't compete well. However, at lower substrate concentrations, it does.

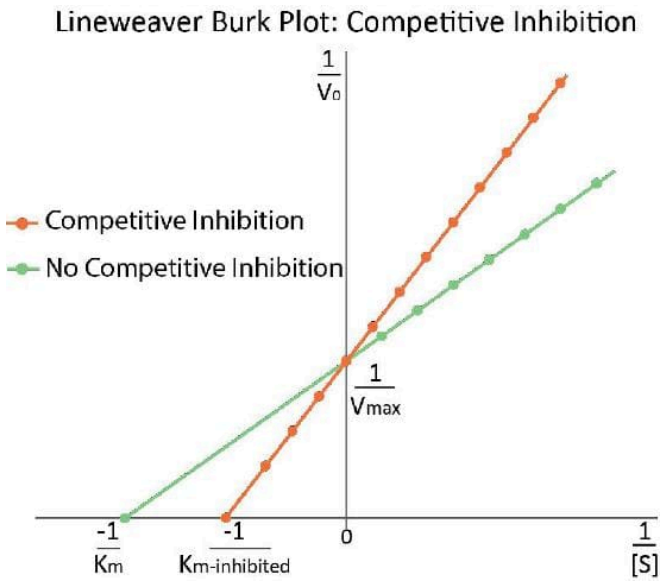


Figure 4.36 - Lineweaver-Burk plots - uninhibited reactions (green). Competitively inhibited reactions (orange). Lines cross on Y-axis at 1/Vmax Since Vmax is the same for both reactions. Image by Aleia Kim

Increased K_m

In competitively inhibited reactions, the apparent Km of the enzyme for the substrate increases ($-1/K_m$ gets closer to zero - red line in Figure 4.36) when the inhibitor is present compared to when the inhibitor is absent, thus illustrating the better competition of the inhibitor at lower substrate concentrations. It may not be obvious why we call the changed Km the apparent Km of the enzyme. The reason is that the inhibitor doesn't actually change the enzyme's affinity for the folate substrate. It only appears to do so. This is because of the way that competitive inhibition works. When the competitive inhibitor binds the enzyme, it is effectively 'taken out of action.' Inactive enzymes have NO affinity for substrate and no activity either. We can't measure Km for an inactive enzyme.

The enzyme molecules that are not bound by methotrexate can, in fact, bind folate and are active. Methotrexate has no effect on them and their Km values are unchanged. Why then, does Km appear higher in the presence of a competitive inhibitor? The reason is that the competitive inhibitor is having a greater effect of reducing the amount of active enzyme at lower concentrations of substrate than it does at higher concentrations of substrate. When the amount of enzyme is reduced, one must have more substrate to supply the reduced amount of enzyme sufficiently to get to Vmax/2.

It is worth noting that in competitive inhibition, the percentage of inactive enzyme changes drastically over the range of [S] values used. To start, at low [S] values, the greatest percentage of the enzyme is inhibited. At high [S], no significant percentage of enzyme is inhibited. This is not always the case, as we shall see in non-competitive inhibition.

Non-competitive inhibition

A second type of inhibition employs inhibitors that do not resemble the substrate and bind not to the active site, but rather to a separate site on the enzyme (Figure 4.37). The effect of binding a non-competitive inhibitor is significantly different from binding a competitive inhibitor because there is no competition. In the case of competitive inhibition, the effect of the inhibitor could be reduced and eventually overwhelmed with increasing amounts of substrate. This was because increasing substrate made increasing percentages of the enzyme active. With non-competitive inhibition, increasing the amount of substrate has no effect on the percentage of enzyme that is active. Indeed, in non-competitive inhibition, the percentage of enzyme inhibited remains the same through all ranges of [S].

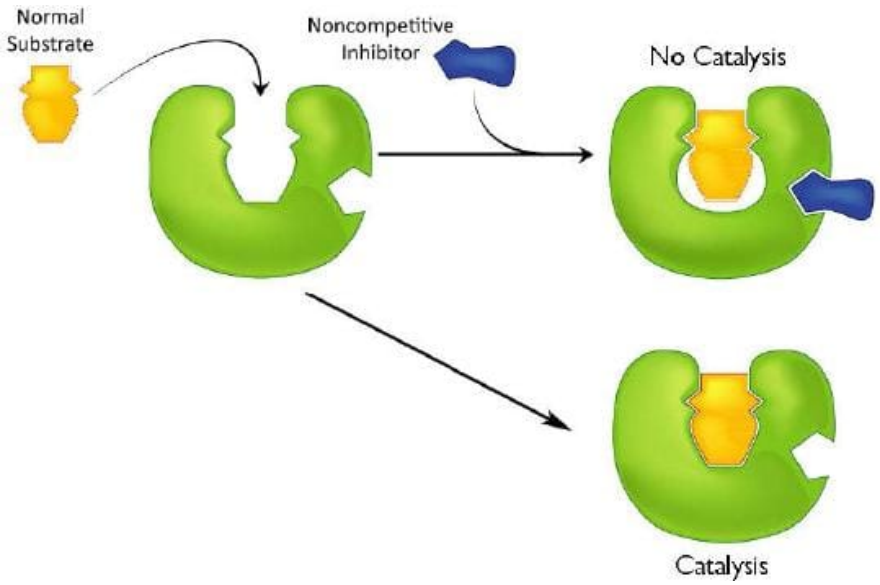


Figure 4.37 - Non-competitive inhibition - inhibitor does not resemble the substrate and binds to a site other than the active site. Image by Aleia Kim

This means, then, that non-competitive inhibition effectively reduces the amount of enzyme by the same fixed amount in a typical experiment at every substrate concentration used The effect of this inhibition is shown in Figure 4.38 & 4.39. As you can see, V_{max} is reduced in non-competitive inhibition compared to uninhibited reactions.

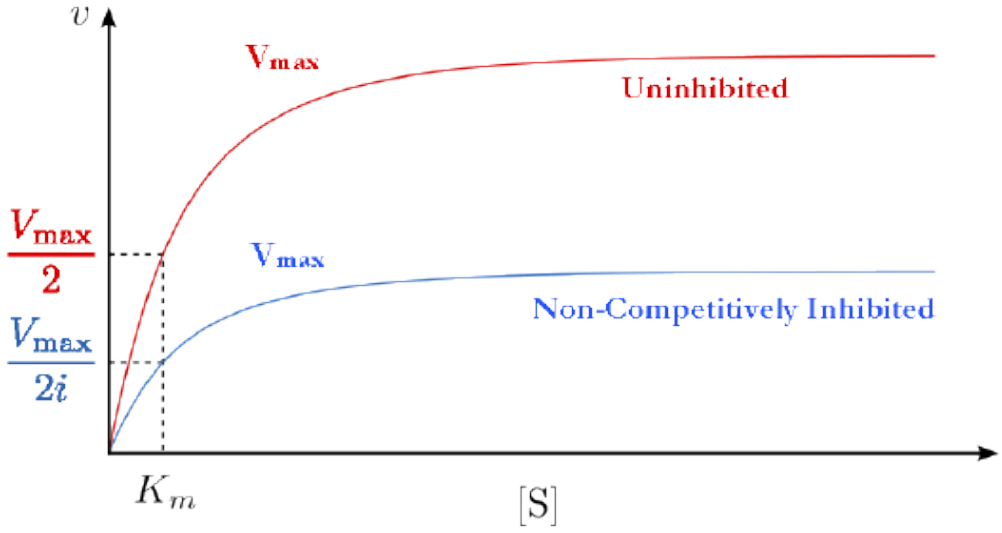


Figure 4.38 - V0 vs [S] plots of uninhibited reactions (red) and non-competitively inhibited reactions (blue). Vmax is reduced, but Km values are unchanged in non-competitively inhibited reactions

This makes sense if we remember that Vmax is dependent on the amount of enzyme present. Reducing the amount of enzyme present reduces V_{max}. In competitive inhibition, this doesn't occur detectably, because at high substrate concentrations, there is essentially 100% of the enzyme active and the V_{max} appears not to change. Additionally, Km for non-competitively inhibited reactions does not change from that of uninhibited reactions. This is because, as noted previously, one can only measure the K_m of active enzymes and K_m is a constant for a given enzyme.

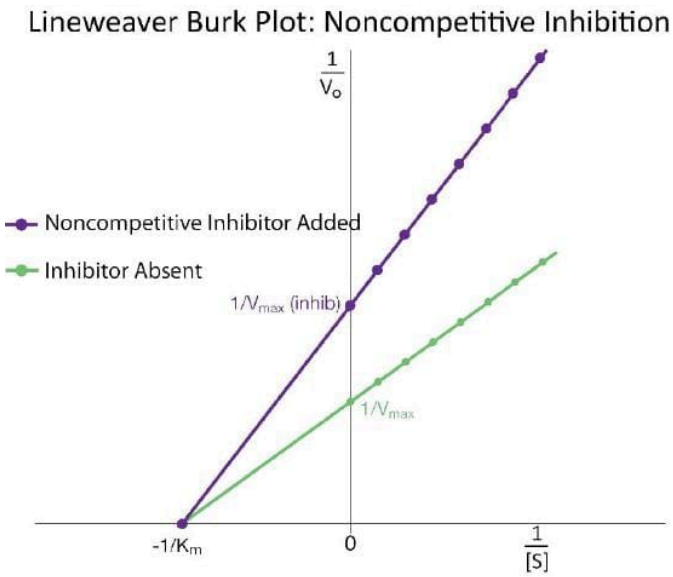


Figure 4.39 - Lineweaver Burk plots of uninhibited reactions (green) and non-competitively inhibited reactions (purple). Image by Aleia Kim

Uncompetitive inhibition

A third type of enzymatic inhibition is that of uncompetitive inhibition, which has the odd property of a reduced Vmax as well as a reduced Km. The explanation for these seemingly odd results is rooted in the fact that the uncompetitive inhibitor binds only to the enzyme-substrate (ES) complex (Figure 4.40). The inhibitor-bound complex forms mostly under concentrations of high substrate and the ES-I complex cannot release product while the inhibitor is bound, thus explaining the reduced V_{max}.

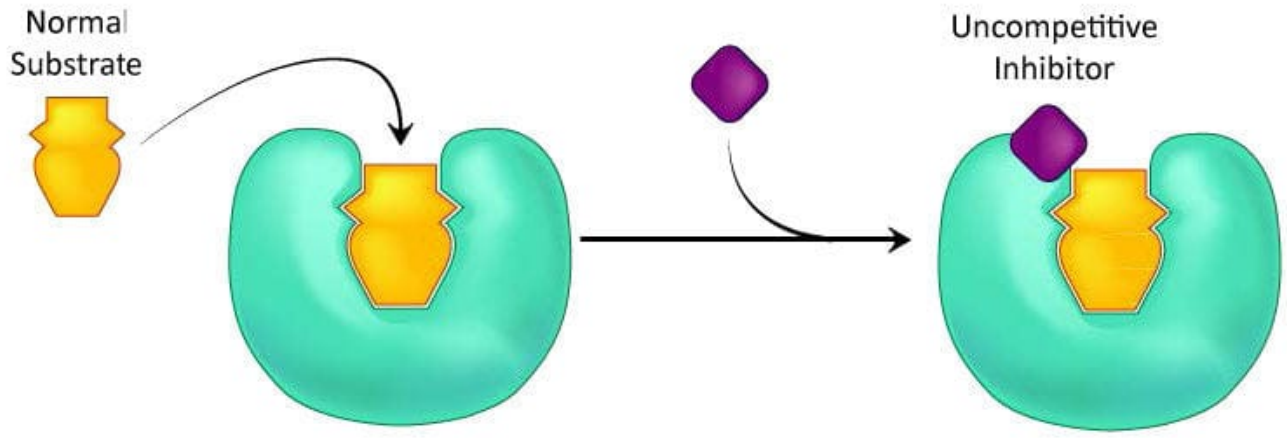


Figure 4.40 - Uncompetitive inhibition. Image by Aleia Kim

The reduced Km is a bit harder to conceptualize. The reason is that the inhibitor-bound complex effectively reduces the concentration of the ES complex. By Le Chatelier's Principle, a shift occurs to form additional ES complex, resulting in less free enzyme and more enzyme in the forms ES and ESI (ES with inhibitor). Decreases in free enzyme correspond to an enzyme with greater affinity for its substrate. Thus, paradoxically, uncompetitive inhibition both decreases V_{max} and increases an enzyme's affinity for its substrate (K_m - Figures 4.41 & 4.42).

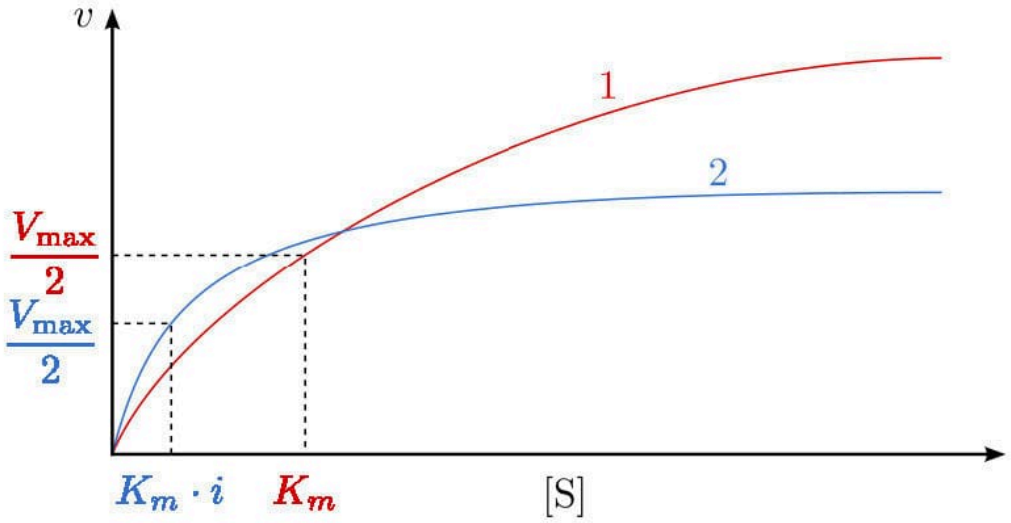


Figure 4.41 - V0 vs [S] plot for uncompetitive inhibition (blue) and uninhibited reactions (red)

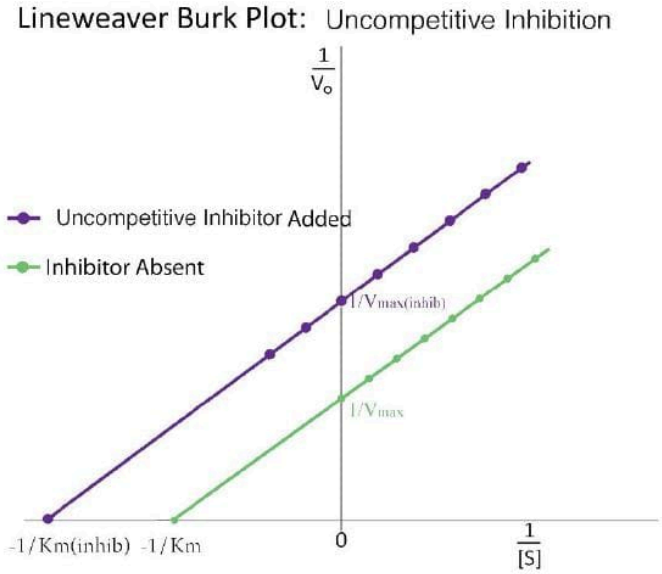


Figure 4.42 - Uncompetitive inhibition (purple) and uninhibited reactions (green). Image by Aleia Kim

Suicide inhibition

In contrast to the first three types of inhibition, which involve reversible binding of the inhibitor to the enzyme, suicide inhibition is irreversible, because the inhibitor becomes covalently bound to the enzyme during the inhibition. Suicide inhibition rather closely resembles competitive inhibition because the inhibitor generally resembles the substrate and binds to the active site of the enzyme. The primary difference is that the suicide inhibitor is chemically reactive in the active site and makes a bond with it that precludes its removal. Such a mechanism is that employed by penicillin (Figure 4.43), which covalently links to the bacterial enzyme, DD transpeptidase and stops it from functioning. Since the normal function of the enzyme is to make a bond necessary for the peptidoglycan complex of the bacterial cell wall, the cell wall cannot properly form and bacteria cannot reproduce.

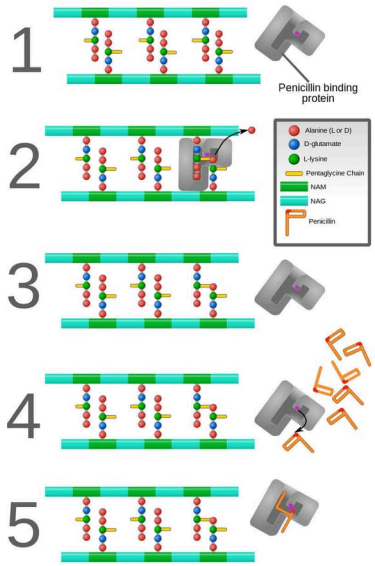


Figure 4.43 - Action of penicillin. DD-transpeptidase builds peptidoglycan layer of bacterial cell wall (1-3). Binding of penicillin by DD-transpeptidase stops peptidoglycan synthesis (4-5). Wikipedia

Control of enzymes

It is appropriate to talk at this point about mechanisms cells use to control enzymes. There are four general methods that are employed:

1. allosterism,
2. covalent modification,
3. access to substrate, and
4. control of enzyme synthesis/breakdown.

Some enzymes are controlled by more than one of these methods.

Allosterism

The term allosterism refers to the fact that the activity of certain enzymes can be affected by the binding of small molecules. Molecules causing allosteric effects come in two classifications. Ones that are substrates for the enzymes they affect are called **homotropic effectors** and those that are not substrates are called heterotropic effectors.

The homotropic effectors usually are activators of the enzymes they bind to and the results of their action can be seen in the conversion of the hyperbolic curve typical of a V0 vs. [S] plot for an enzyme (Figure 4.18), being converted to a sigmoidal plot (Figure 4.44). This is due to the conversion of the enzyme from the T-state to the R-state on binding the substrate/homotropic effector.

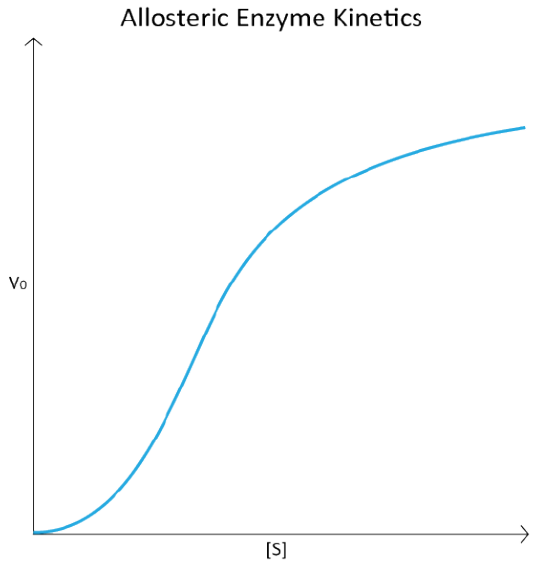


Figure 4.44 - Kinetic profile of an allosteric enzyme whose activity is controlled by a homotropic effector. Image by Aleia Kim

The V0 vs. [S] plot of allosteric enzyme reactions resembles the oxygen binding curve of hemoglobin (see Figure 2.83). Even though hemoglobin is not an enzyme and is thus not catalyzing a reaction, the similarity of the plots is not coincidental. In both cases, the binding of an external molecule is being measured – directly, in the hemoglobin plot, and indirectly by the V0 vs. [S] plot, since substrate binding is a factor in enzyme reaction velocity.

Allosteric inhibition

Allosterically, regulation of these enzymes works by inducing different physical states (shapes, as it were) that affect their ability to bind to substrate. When an enzyme is inhibited by binding an effector, it is converted to the T-state (T=tight), it has a reduced affinity for substrate and it is through this means that the reaction is slowed.

Allosteric activation

On the other hand, when an enzyme is activated by effector binding, it converts to the R-state (R=relaxed) and binds substrate much more readily. When no effector is present, the enzyme may be in a mixture of T- and R-states.

Feedback inhibition

An interesting kind of allosteric control is exhibited by HMG-CoA reductase, which catalyzes an important reaction in the pathway leading to the synthesis of cholesterol. Binding of cholesterol to the enzyme reduces the enzyme's activity significantly. Cholesterol is not a substrate for the enzyme, so it is therefore a heterotropic effector.

Notably, though, cholesterol is the end-product of the pathway that HMG-CoA reductase catalyzes a reaction in. When enzymes are inhibited by an end-product of the pathway in which they participate, they are said to exhibit feedback inhibition.

Feedback inhibition always operates by allosterism and further, provides important and efficient control of an entire pathway. By inhibiting an early enzyme in a pathway, the flow of materials (and ATP hydrolysis required for their processing) for the entire pathway is stopped or reduced, assuming there are not alternate supply methods.

Pathway control

In the cholesterol biosynthesis pathway, stopping this one enzyme has the effect of shutting off (or at least slowing down) the entire pathway. This is significant because after catalysis by HMG-CoA reductase, there are over 20 further reactions necessary to make cholesterol, many of them requiring ATP energy. Shutting down one reactions stops all of them. Another excellent example of allosteric control and feedback inhibition is the enzyme ATCase, discussed below.

ATCase

Another interesting example of allosteric control and feedback inhibition is associated with the enzyme Aspartate

Transcarbamoylase (ATCase). This enzyme, which catalyzes a step in the synthesis of pyrimidine nucleotides, has 12 subunits. These include six identical catalytic subunits and six identical regulatory subunits. The catalytic subunits bind to substrate and catalyze a reaction. The regulatory subunits bind to either ATP or CTP. If they bind to ATP, the enzyme subunits arrange themselves in the R-state.

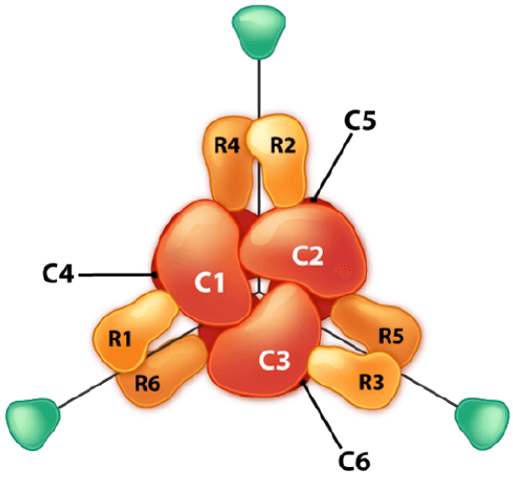


Figure 4.45 - Schematic structure of ATCase. Regulatory Units = R, Catalytic Units = C. Image by Aleia Kim

R-state

The R-state of ATCase allows the substrate to have easier access to the six active sites and the reaction occurs more rapidly. For the same amount of substrate, an enzyme in the R-state will have a higher velocity than the same enzyme that is not in the R-state. By contrast, if the enzyme binds to CTP on one of its regulatory subunits, the subunits will arrange in the T-state and in this form, the substrate will not have easy access to the active sites, resulting in a slower velocity for the same concentration of substrate compared to the R-state. ATCase is interesting in that it can also flip into the R-state when one of the substrates (aspartate) binds to an active site within one of the catalytic subunits.

Aspartate has the effect of activating the catalytic action of the enzyme by favoring the R-state. Thus, aspartate, which is a substrate of the enzyme is a homotropic effector and ATP and CTP, which are not substrates of the enzyme are heterotropic effectors of ATCase.

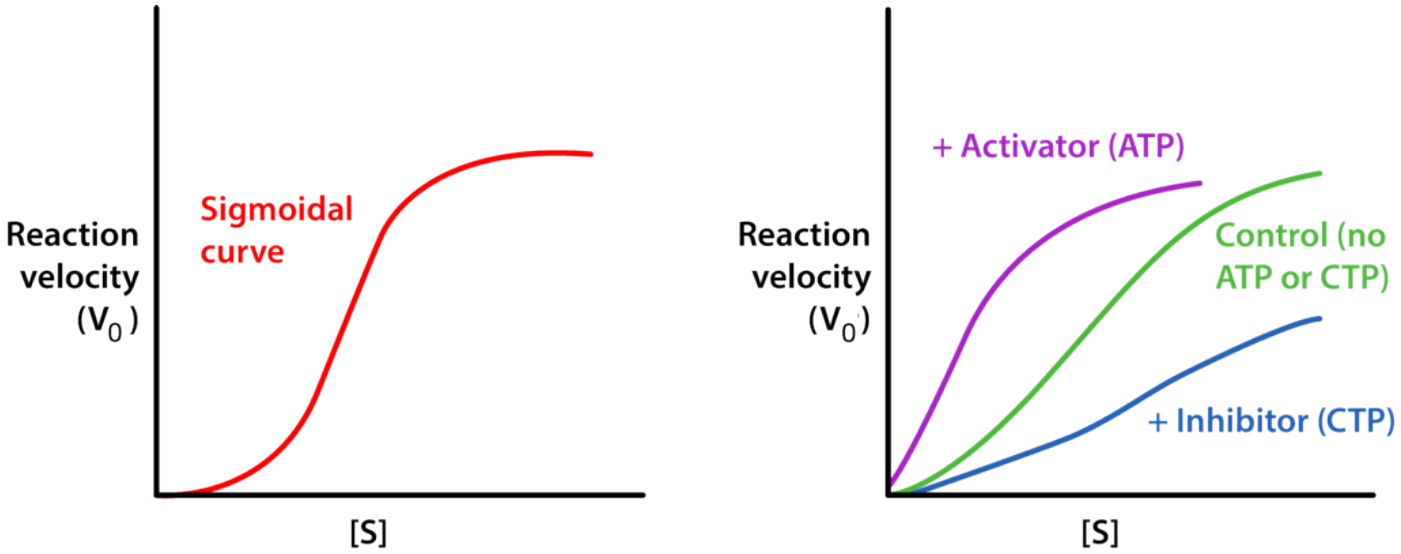


Figure 4.46 - Plots of V0 vs. [S] for ATCase. Left - Allosteric effect of aspartate. Right - Allosteric effects of ATP (activator) and CTP (inhibitor). Image by Pehr Jacobson

Allosteric models

There are three models commonly used to explain how allosterism regulates multi-subunit enzyme activity. They are known as

- the Monod-Wyman-Changeux (MWC) model (also known as the concerted model),
- the sequential model (also known as KNF),
- and the morpheein model.

All models describe a Tense (T) state that is less catalytically active and a Relaxed (R) state that is more catalytically active. The models differ in how the states change.

Sequential model

In the sequential model, binding of an allosteric effector by one subunit causes it to change from T to R state (or vice versa) and that change makes it easier for adjacent subunits to similarly change state. With this model , there is a cause/effect relationship between binding of an effector by one subunit and change of state by an adjacent subunit.

In hemoglobin, for example, binding of one oxygen by one unit of the complex may induce that unit to flip to the R-state and, through interactions with other subunits, cause them to favor adopting the R configuration before they bind to oxygen. In this way, binding of one subunit favors binding of others and cooperativity can be explained by the change in binding affinity as oxygen concentration changes.

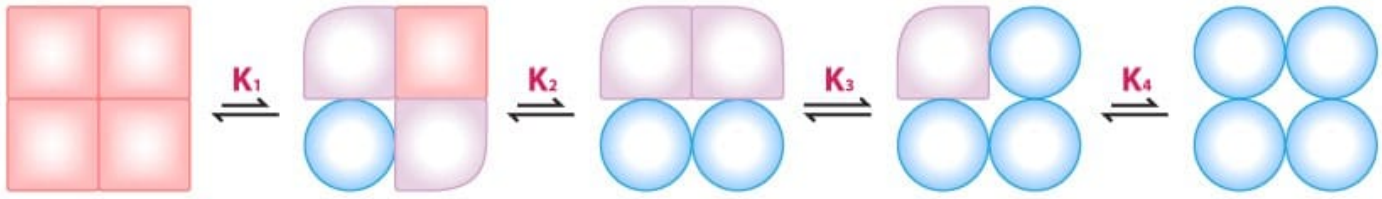


Figure 4.47 - Sequential model of allosteric regulation. Round = R-State. Square = T-state.. Image by Aleia Kim

MWC model

The MWC model is less intuitive. In it, the entire complex changes state from T to R (or vice versa) independently of the binding of effectors. Flipping between T-states and R-states is postulated to be in an equilibrium of states in the absence of effector (for example, a 50 to 1 ratio of T/R. This ratio is referred to as L, so L = T/R). Binding of effector by the enzyme complex has the tendency of “locking” the complex in a state. Binding of inhibitors will increase the ratio of T/R whereas binding of activators will increase R and thus decrease the ration of T/R.

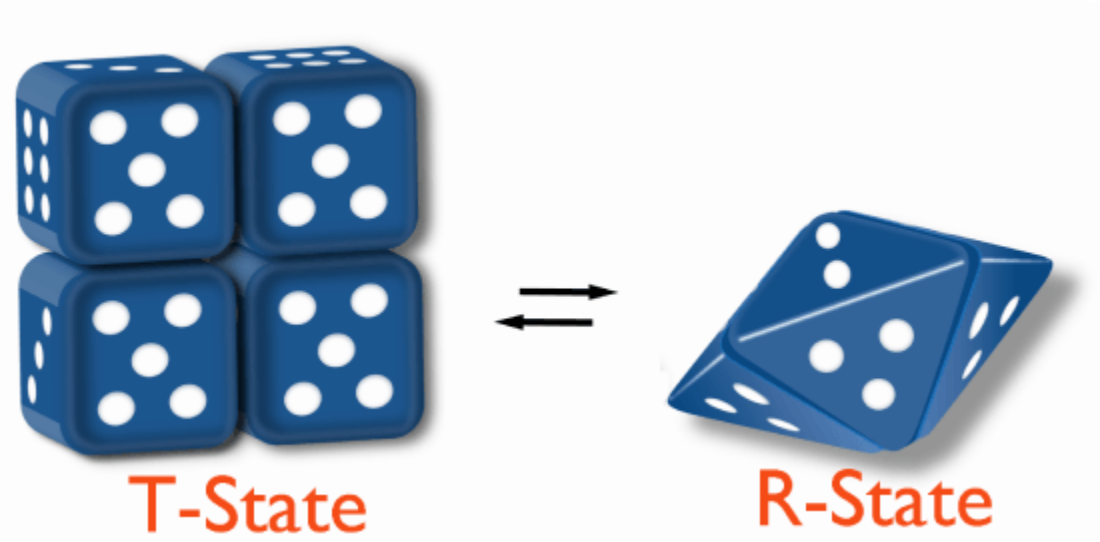


Figure 4.48 - The MWC Model - The multisubunit complex flips as a whole and is in equilibrium. Wikipedia

Morpheein model

The morpheein model is similar to the MWC model, but with an added step of dissociation of the subunits. The MWC model proposes that flipping between R and T states occurs by the complex as a whole and occurs on all units simultaneously. The morpheein model instead proposes that the multi-subunit enzyme breaks down to individual units which can then flip in structure and re-form the complex. In the morpheein model, only identically shaped units (all R, for example) can come together in the complex, thus explaining the “all-R-” or “all-T-” state found in the MWC model.

A large number of enzymes, including prominent ones like citrate synthase, acetyl-CoA carboxylase, glutamate dehydrogenase, ribonucleotide reductase and lactate dehydrogenase have behavior consistent with the morpheein model.

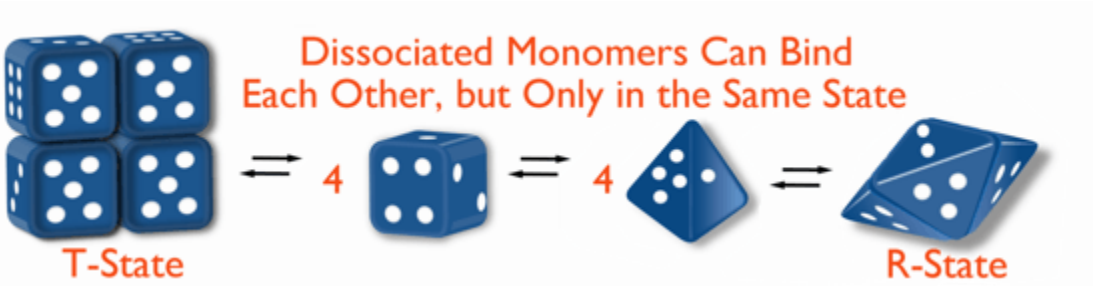


Figure 4.49 - Morpheein model of allosterism. States flip as monomers and aggregate only with same monomer. Wikipedia

Covalent control of enzymes

Some enzymes are synthesized in a completely inactive form and their activation requires covalent bonds in them to be cleaved. Such inactive forms of enzymes are called zymogens. Examples include the proteins involved in blood clotting and proteolytic enzymes of the digestive system, such as trypsin, chymotrypsin, pepsin, and others.

Synthesizing some enzymes in an inactive form makes very good sense when an enzyme’s activity might be harmful to the tissue where it is being made. For example, the painful condition known as pancreatitis arises when digestive enzymes made in the pancreas are activated too soon and end up attacking the pancreas.

Cascades

For both the blood clotting enzymes and the digestive enzymes, the zymogens are activated in a protease cascade. This occurs when activation of one enzyme activates others in a sort of chain reaction. In such a scheme the first enzyme activated proteolytically cleaves the second zymogen, causing it to be activated, which in turn activates a third and this may proceed through several levels of enzymatic action (Figure 4.50).

The advantage of cascades is that they allow a large amount of zymogens to become activated fairly quickly, since there is an amplification of the signal at each level of catalysis.

Zymogens are also abundant in blood. Blood clotting involves polymerization of a protein known as fibrin. Since random formation of fibrin is extremely hazardous because it can block the flow of blood, potentially causing heart attack/stroke, the body synthesizes fibrin as a zymogen (fibrinogen) and its activation results from a “cascade” of activations of proteases that arise when a signal is received from a wound. Similarly, the enzyme catalyzing removal of fibrin clots (plasmin) is also synthesized as a zymogen (plasminogen), since random clot removal would also be hazardous (see below also).

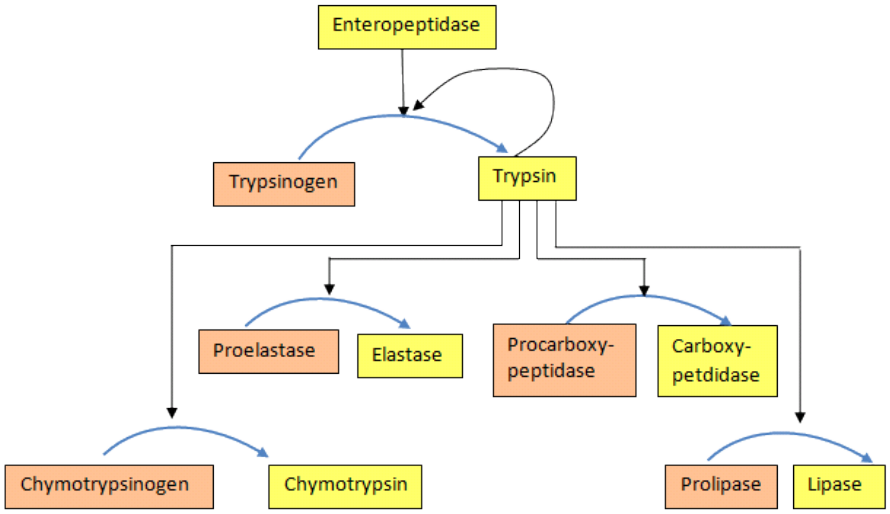


Figure 4.50 - Protease activation scheme. Wikipedia

Phosphorylation/dephosphorylation

Another common mechanism for control of enzyme activity by covalent modification is phosphorylation. The phosphorylation of enzymes (on the side chains of serine, threonine or tyrosine residues) is carried out by protein kinases. Enzymes activated by phosphorylation can be regulated by the addition of phosphate groups by kinases or their removal by phosphatases. Thus, this type of covalent modification is readily reversible, in contrast to proteolytic cleavage.

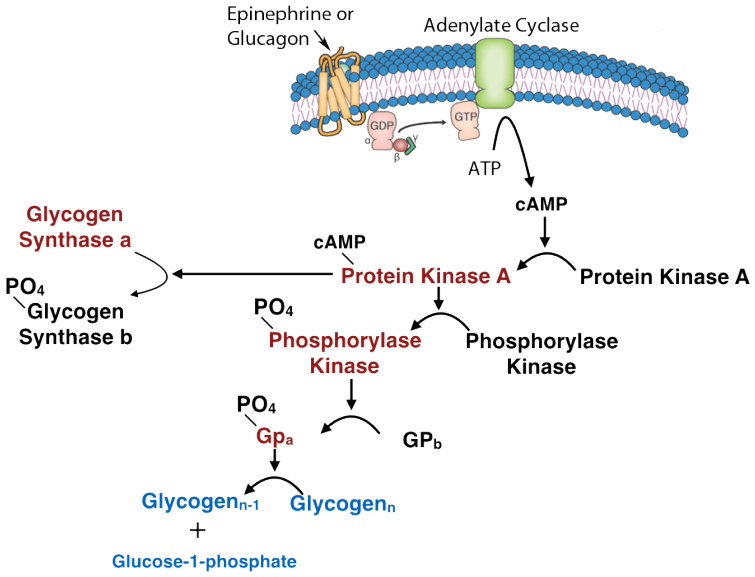


Figure 4.51 - Regulation by covalent modification of glycogen catabolism enzymes

Reduction/oxidation

An interesting covalent control of enzymes using reduction/oxidation is exhibited in photosynthetic plants. In the light phase of photosynthesis, electrons are excited by light and flow through carriers to NADP+, forming NADPH. Thus, in the light, the NADPH concentration is high. When NADPH concentration is high, the concentration of reduced ferredoxin (a molecule donating electrons to NADP+) is also high.

Reduced ferredoxin can transfer electrons to thioredoxin, reducing it. Reduced thioredoxin can, in turn, transfer electrons to proteins to reduce their disulfide bonds. Four enzymes related to the Calvin cycle can receive electrons from thioredoxin and become activated, as a result.

These include sedoheptulose 1,7-bisphosphatase, ribulose-5-phosphate kinase, fructose 1,6-bisphosphatase, and glyceraldehyde 3-phosphate dehydrogenase. Thus, in the light, electrons flow, causing NADPH to accumulate and ferredoxin to push electrons in the direction of these enzymes above, activating them and favoring the Calvin cycle. In the dark, the concentration of reduced NADPH, reduced ferredoxin, and reduced thioredoxin fall, resulting in loss of electrons by the Calvin cycle enzymes (oxidations that re-form disulfide bonds) and the Calvin cycle inactivates.

Other enzyme control mechanisms

Other means of controlling enzymes relate to access to substrate (substrate-level control) and control of enzyme synthesis. Hexokinase is an enzyme that is largely regulated by availability of its substrate, glucose. When glucose concentration is low, the product of the enzyme's catalysis, glucose-6-phosphate, inhibits the enzyme's function.

Regulation of enzymes by controlling their synthesis is covered later in the book in the discussion relating to control of gene expression.

This page titled 4.2: Control of Enzymatic Activity is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

4.3: Mechanisms of Catalysis

A printable version of this section is here: BiochemFFA_4_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

The magic of enzymes, as noted, is in their ability to create electronic environments conducive to initiation of a reaction. There are more mechanisms of reaction than we could ever hope to cover in a book like this, and comprehensive discussion of these is not our aim. Instead, we will cite some examples and go into detail on one of them - the mechanism of action of serine proteases.

Chymotrypsin

We will begin with mechanism of action of one enzyme - chymotrypsin. Found in our digestive system, chymotrypsin's catalytic activity is cleaving peptide bonds in proteins and it uses the side chain of a serine in its mechanism of catalysis. Many other protein-cutting enzymes employ a very similar mechanism and they are known collectively as serine proteases (Figure 4.52).

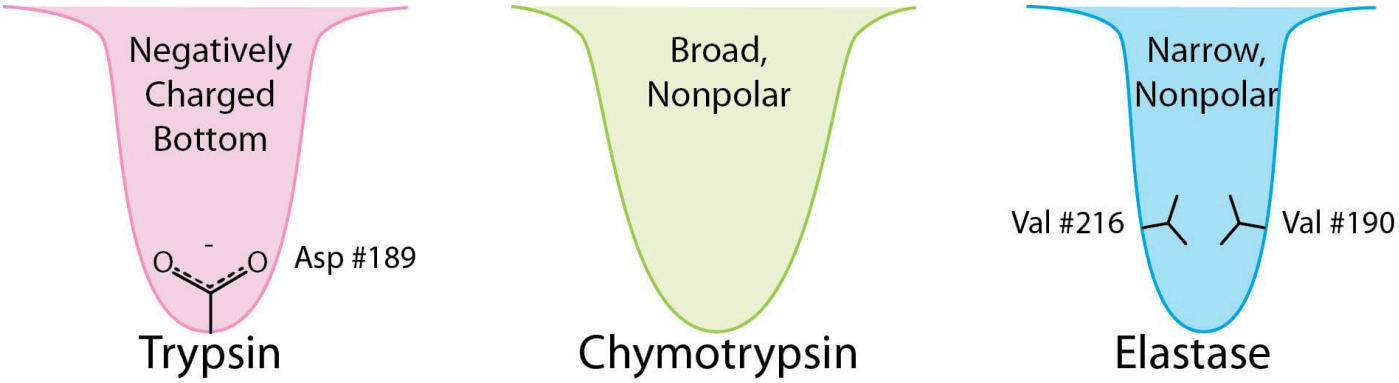


Figure 4.52 - Substrate binding sites (S1 pockets) of three serine proteases. Image by Aleia Kim

These enzymes are found in prokaryotic and eukaryotic cells and all use a common set of three amino acids in the active site called a catalytic triad (Figure 4.53). It consists of aspartic acid, histidine, and serine. The serine is activated in the reaction mechanism to form a nucleophile in these enzymes and gives the class their name. With the exception of the recognition that occurs at the substrate binding site, the mechanism shown here for chymotrypsin would be applicable to any of the serine proteases.

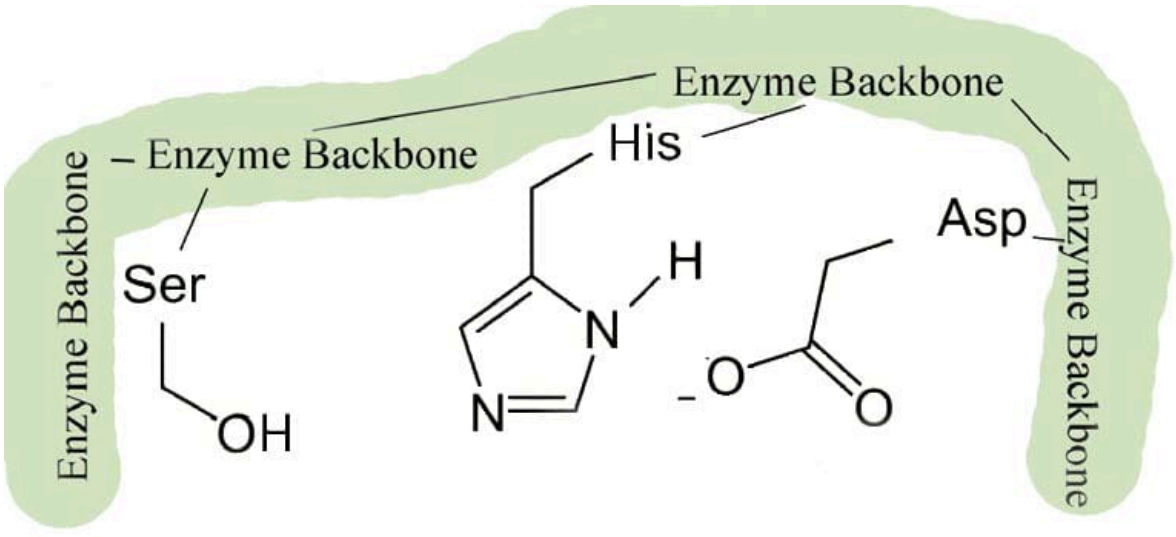


Figure 4.53 - 1. Active site of chymotrypsin showing the catalytic triad of serine - histidine-aspartic acid

Specificity

As a protease, chymotrypsin acts fairly specifically, cutting not all peptide bonds, but only those that are adjacent to relatively non-polar amino acids in the protein. One of the amino acids it cuts adjacent to is phenylalanine. The enzyme's action occurs in two phases – a fast phase that occurs first and a slower phase that follows. The enzyme has a substrate binding site that includes a region of the enzyme known as the S1 pocket. Let us step through the mechanism by which chymotrypsin cuts adjacent to phenylalanine.

Substrate binding

The process starts with the binding of the substrate in the S1 pocket (Figure 4.54). The S1 pocket in chymotrypsin has a hydrophobic hole in which the substrate is bound. Preferred substrates will include amino acid side chains that are bulky and hydrophobic, like phenylalanine. If an ionized side chain, like that of glutamic acid binds in the S1 pocket, it will quickly exit, much like water would avoid an oily interior.

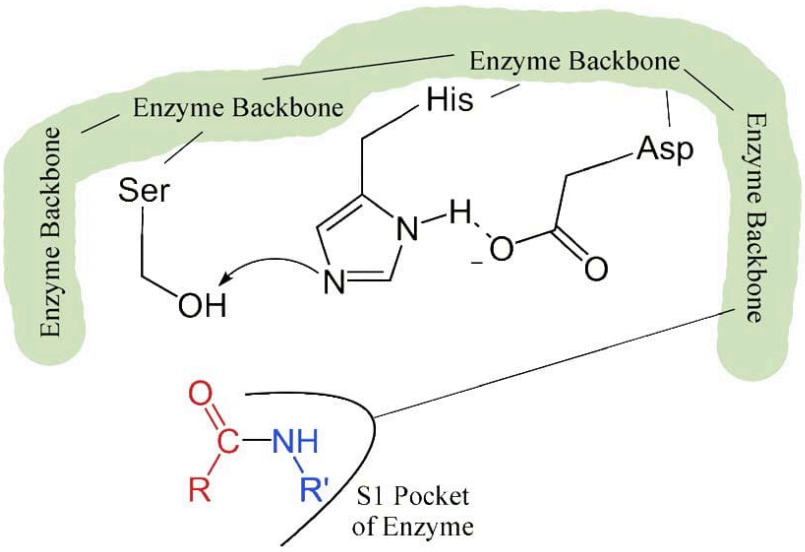


Figure 4.54 - 2. Binding of substrate to S1 pocket in the active site

Shape change on binding

When the proper substrate binds in the S1 pocket, its presence induces an ever so slight change in the shape of the enzyme. This subtle shape change on the binding of the proper substrate starts the steps of the catalysis. Since the catalytic process only starts when the proper substrate binds, this is the reason that the enzyme shows specificity for cutting at specific amino acids in the target protein. Only amino acids with the side chains that interact well with the S1 pocket start the catalytic wheels turning.

The slight changes in shape involve changes in the positioning of three amino acids (aspartic acid, histidine, and serine) in the active site known as the *catalytic triad*.

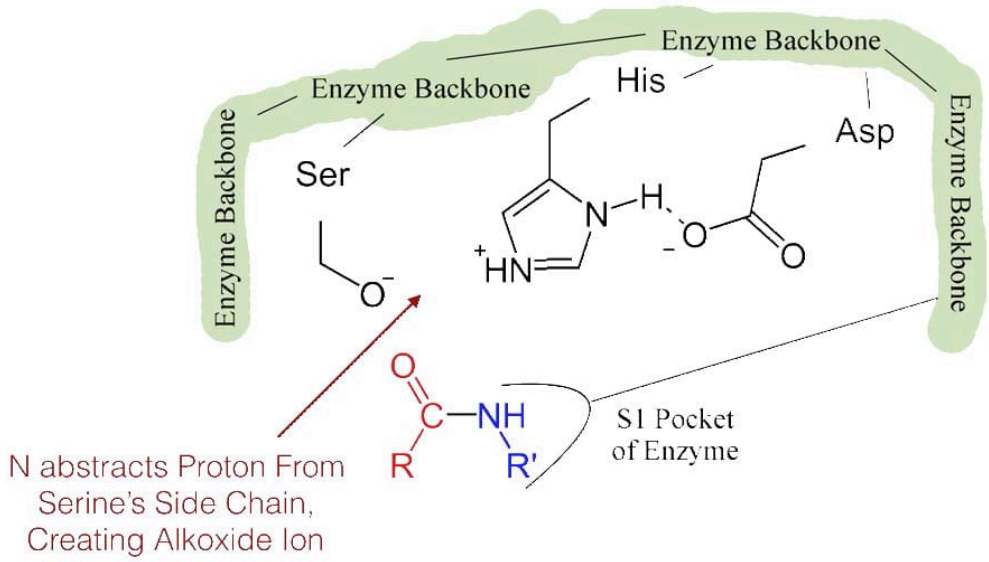


Figure 4.55 - 3. Formation of alkoxide ion

The shift of the negatively charged aspartic acid towards the electron rich histidine ring favors the abstraction of a proton by the histidine from the hydroxyl group on the side chain of serine, resulting in production of a very reactive alkoxide ion in the active site (Figure 4.55).

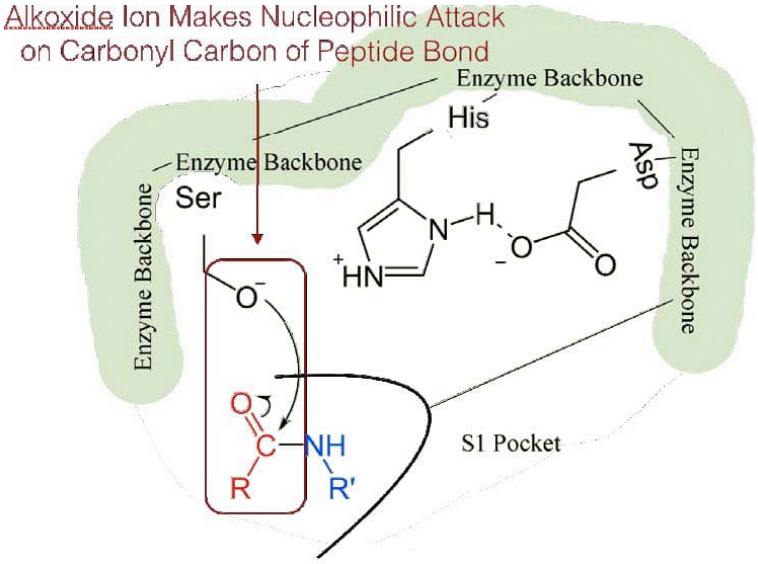


Figure 4.56 - 4. Nucleophilic attack

Since the active site at this point also contains the polypeptide chain positioned with the phenylalanine side chain embedded in the S1 pocket, the alkoxide ion performs a nucleophilic attack on the peptide bond on the carboxyl side of phenylalanine sitting in the S1 pocket (Figure 4.56). This reaction breaks the peptide bond (Figure 4.57) and causes two things to happen.

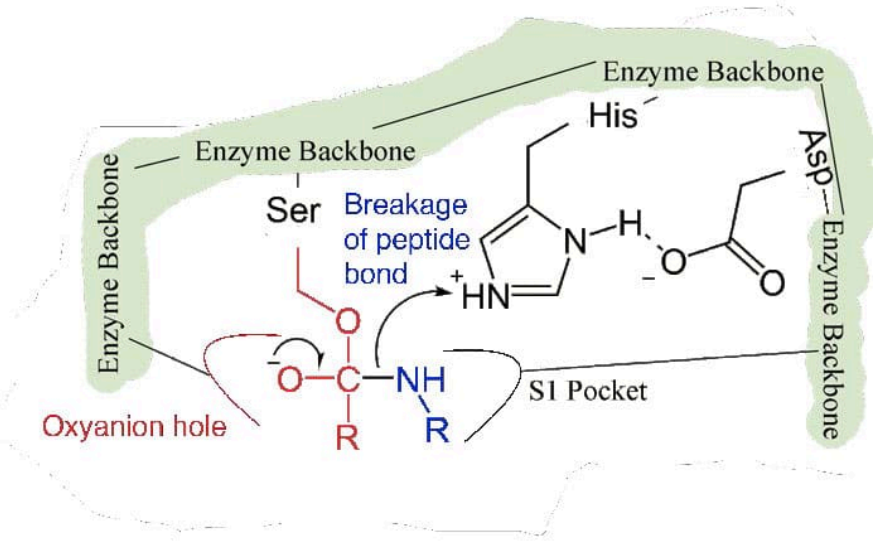


Figure 4.57 - 5. Stabilization by oxyanion hole. Breakage of peptide bond.

First, one end of the original polypeptide is freed and exits the active site (Figure 4.58). The second is that the end containing the phenylalanine is covalently linked to the oxygen of the serine side chain. At this point we have completed the first (fast) phase of the catalysis.

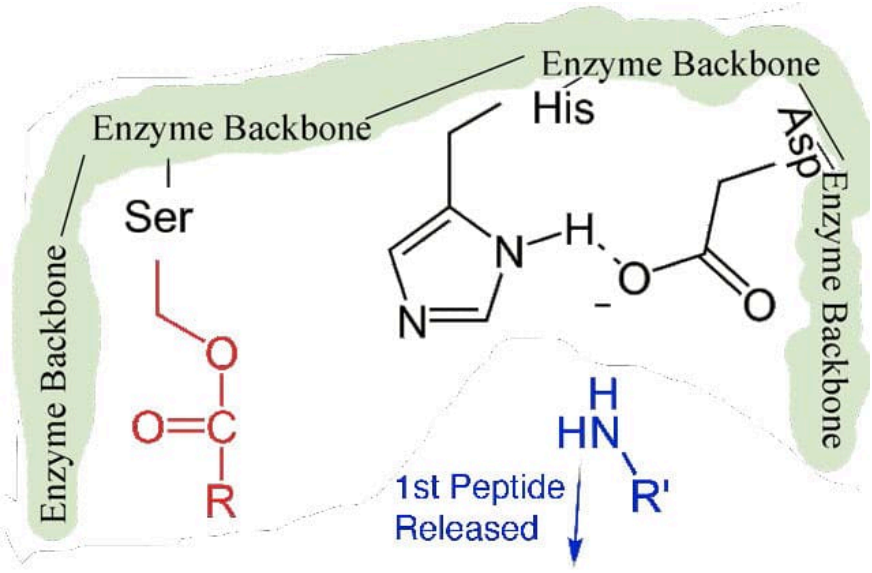


Figure 4.58 - 6. First peptide released. Other half bonded to serine.

Slower second phase

The second phase of the catalysis by chymotrypsin is slower. It requires that the covalent bond between phenylalanine and serine's oxygen be broken so the peptide can be released and the enzyme can return to its original state. The process starts with entry of water into the active site. Water is attacked in a fashion similar to that of the serine side chain in the first phase, creating a reactive hydroxyl group (Figure 4.59) that performs a nucleophilic attack on the phenylalanine-serine bond (Figure 4.60), releasing it and replacing the proton on serine. The second peptide is released in the process and the reaction is complete with the enzyme back in its original state (Figure 4.61).

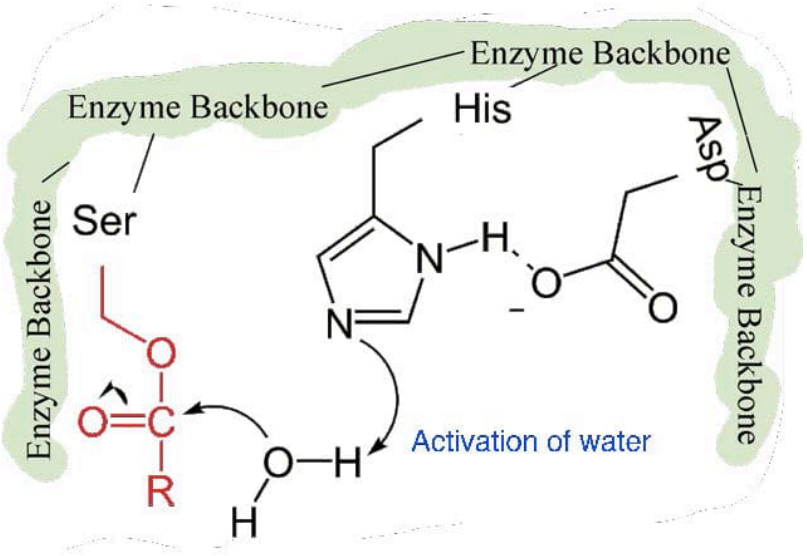


Figure 4.59 - 7. Activation of water by histidine

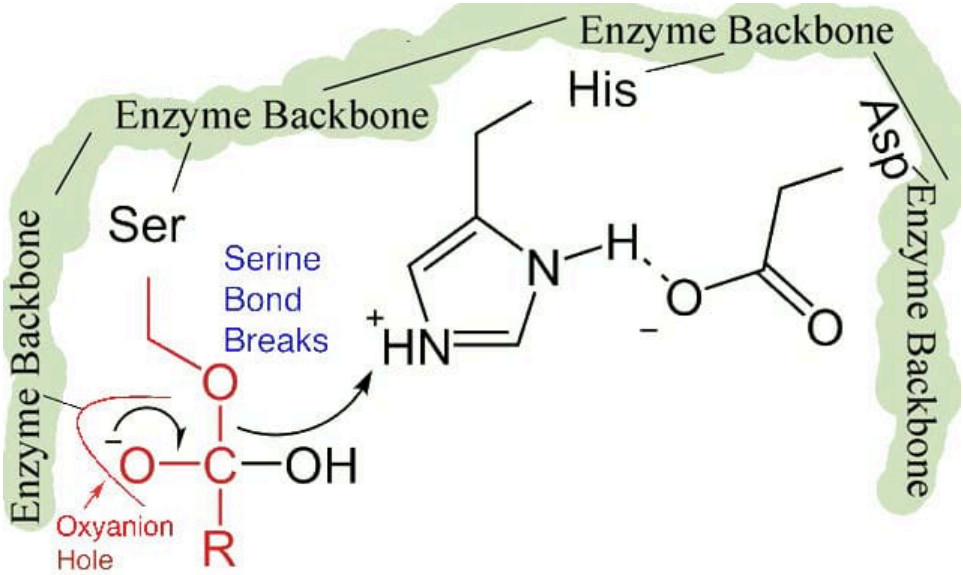


Figure 4.60 - 8. Nucleophile attack by hydroxyl creates tetrahydryl intermediate stabilized by oxyanion hole. Bond to serine breaks.

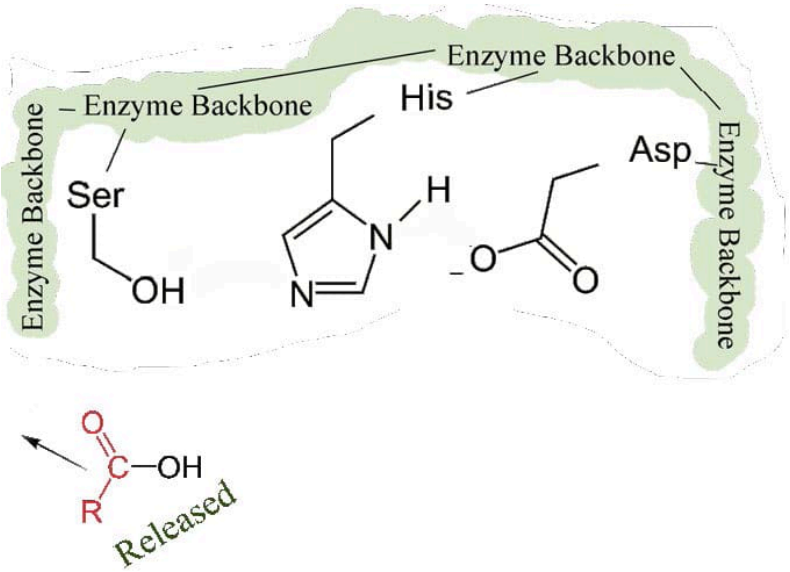


Figure 4.61 - 9. Second half of peptide released. Enzyme active site restored.

Serine proteases

The list of serine proteases is quite long. They are grouped in two broad categories - 1) those that are chymotrypsin-like and 2) those that are subtilisin-like. Though subtilisin-type and chymotrypsin-like enzymes use the same mechanism of action, including the catalytic triad, the enzymes are otherwise not related to each other by sequence and appear to have evolved independently. They are, thus, an example of convergent evolution - a process where evolution of different forms converge on a structure to provide a common function.

The serine protease enzymes cut adjacent to specific amino acids and the specificity is determined by the size/shape/charge of amino acid side chain that fits into the enzyme's S1 binding pocket (Figure 4.62).

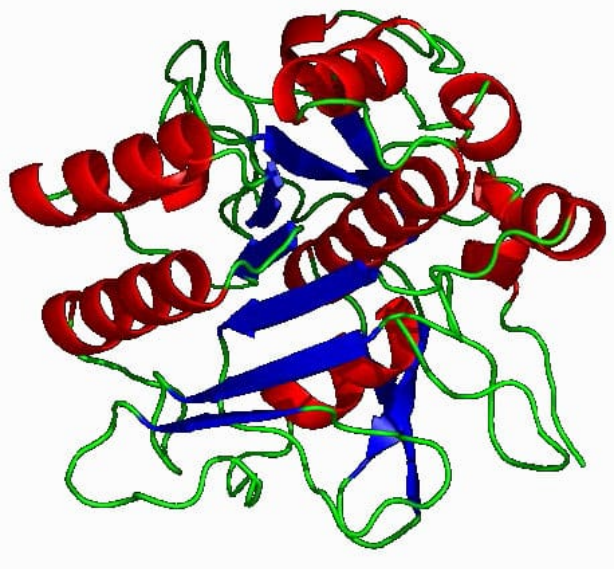


Figure 4.62 - Subtilisin - A serine protease

Examples of serine proteases include trypsin, chymotrypsin, elastase, subtilisin, signal peptidase I, and nucleoporin. Serine proteases participate in many physiological processes, including blood coagulation, digestion, reproduction, and the immune response.

Cysteine proteases

Cysteine proteases (also known as thiol proteases) catalyze the breakdown of proteins by cleaving peptide bonds using a nucleophilic thiol from a cysteine (Figure 4.63). The cysteine is typically found in a catalytic dyad or triad also involving histidine and (sometimes) aspartic acid (very much like serine proteases). The sulfhydryl group of cysteine proteases is more acidic than the hydroxyl of serine proteases, so the aspartic acid of the triad is not always needed.

The mechanism of action is very similar to that of serine proteases. Binding of proper substrate results in activation of the thiol (removal of the proton by the histidine group). The activated thiol acts as a nucleophile, attacking the peptide bond and causing it break. One peptide is released and the other peptide becomes covalently linked to the sulfur. Hydrolysis by water releases the second peptide and completes the cycle. Examples of cysteine proteases include papain, caspases, hedgehog protein, calpain, and cathepsin K.

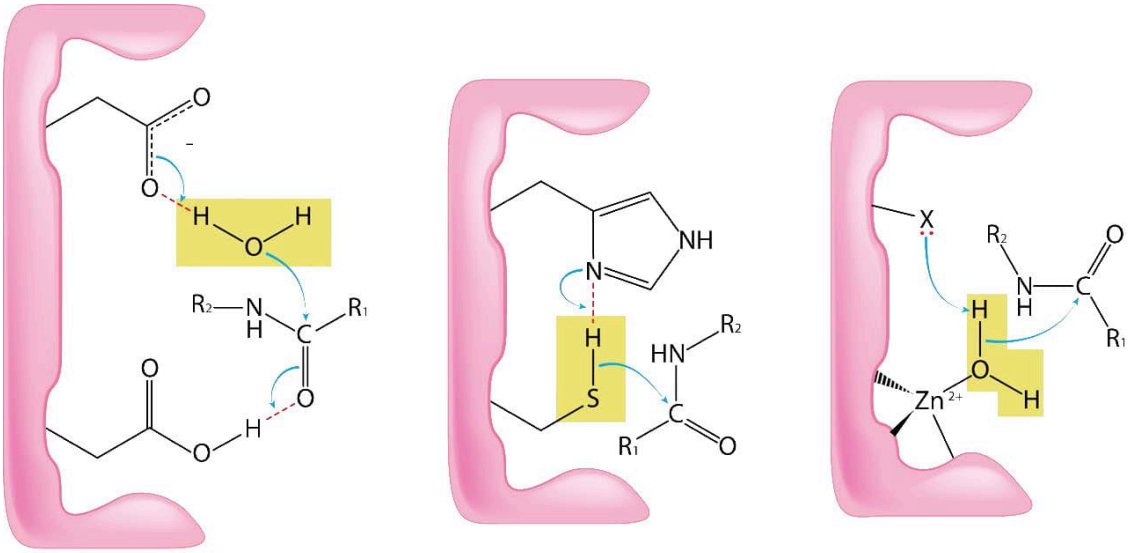


Figure 4.63 - Mechanism of action of proteases. In each case, a nucleophile is created - hydroxyl (aspartyl proteases), thiol (cysteine proteases), and hydroxyl (metalloproteases). Image by Aleia Kim

Caspases

Caspases (Cysteine-ASPartic ProteASEs) are a family of cysteine proteases that play important roles in the body. At the cellular level they function in apoptosis and necrosis and in the body, they are involved in inflammation and the immune system. Maturation of lymphocytes is one such role. They are best known, however, for their role in apoptosis, which has given rise to descriptions of them as “executioner” proteins or “suicide proteases” that dismantle cells in programmed cell death.

There are 12 known human caspases. The enzymes are synthesized as pro-caspase zymogens with a prodomain and two other subunits. The prodomain contains regions that allow it to interact with other molecules that regulate the enzyme’s activity. The caspases come in two forms. The initiator caspases, when activated, activate the effector caspases. The effector caspases cleave other proteins in the cell. Targets for effector caspase cleavage action include the nuclear lamins (fibrous proteins providing structural integrity to the nucleus), ICAD/DFF45 (an inhibitor of DNAse), PARP (flags areas where DNA repair needed), and PAK2 (apoptotic regulation).

The caspase activation cascade can itself be activated by granzyme B (a serine protease secreted by natural killer cells and cytotoxic T-cells), cellular death receptors, and the apoptosome (large protein structure in apoptotic cells stimulated by release of cytochrome C from the mitochondria). Each of these activators is responsible for activating different groups of caspases.

Metalloproteases

Metalloproteases (Figure 4.64) are enzymes whose catalytic mechanism for breaking peptide bonds involves a metal. Most metalloproteases use zinc as their metal, but a few use cobalt, coordinated to the protein by three amino acid residues with a

labile water at the fourth position. A variety of side chains are used - histidine, aspartate, glutamate, arginine, and lysine. The water is the target of action of the metal which, upon binding of the proper substrate, abstracts a proton to create a nucleophilic hydroxyl group that attacks the peptide bond, cleaving it (Figure 4.64). Since the nucleophile here is not attached covalently to the enzyme, neither of the cleaved peptides ends up attached to the enzyme during the catalytic process. Examples of metalloproteases include carboxypeptidases, aminopeptidases, insulinases and thermolysin.



Figure 4.64 - Carboxypeptidase - A metalloprotease

Aspartyl proteases

As the name suggests, aspartyl proteases use aspartic acid in their catalytic mechanism (Figures 4.63 & 4.65). Like the metalloproteases, aspartyl proteases activate a water to create a nucleophile for catalysis (Figure 4.65). The activated water attacks the peptide bond of the bound substrate and releases the two pieces without the need to release a bound intermediate, since water is not covalently attached to the enzyme. Common aspartyl proteases include pepsin, signal peptidase II, and HIV-1 protease.

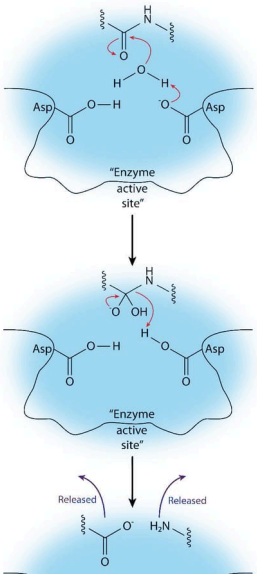


Figure 4.65 - Activation of an aspartyl protease - aspartate side chain abstracts proton (top). Hydroxyl attacks peptide bond (middle). Broken peptide pieces released (bottom). Image by Pehr Jacobson

Threonine proteases

Though threonine has an R-group with a hydroxyl like serine, the mechanism of action of this class of proteases differs somewhat from the serine proteases. There are some similarities. First, the threonine's hydroxyl plays a role in catalysis and that is to act as a nucleophile. The nucleophile is created, however, not by a catalytic triad, but rather as a result of threonine's own α-amine group abstracting a proton.

Because of this, the nucleophilic threonine in a threonine protease must be at the n-terminus of the enzyme. Nucleophilic attack of the peptide bond in the target protease results in breakage of the bond to release one peptide and the other is covalently attached to serine, like the serine proteases. Also, as with the serine proteases, water must come in to release the covalently linked second peptide to conclude the catalytic mechanism.

Examples

Examples of threonine proteases include the catalytic subunits of the proteasome. Some acyl transferases (such as ornithine acyltransferase) have evolved the same catalytic mechanism by convergent evolution. The latter enzymes use ornithine instead of water to break the enzyme-substrate covalent bond, with the result that the acyl-group becomes attached to ornithine, instead of water.

Protease inhibitors

Molecules which inhibit the catalytic action of proteases are known as protease inhibitors. These come in a variety of forms and have biological and medicinal uses. Many biological inhibitors are proteins themselves. Protease inhibitors can act in several ways, including as a suicide inhibitor, a transition state inhibitor, a denaturant, and as a chelating agent. Some work only on specific classes of enzymes. For example, most known aspartyl proteases are inhibited by pepstatin. Metalloproteases are sensitive to anything that removes the metal they require for catalysis. Zinc-containing metalloproteases, for example, are very sensitive to EDTA, which chelates the zinc ion.

One category of proteinaceous protease inhibitors is known as the serpins. Serpins inhibit serine proteases that act like chymotrypsin. 36 of them are known in humans.

Serpins are unusual in acting by binding to a target protease irreversibly and undergoing a conformational change to alter the active site of its target. Other protease inhibitors act as competitive inhibitors that block the active site.

Serpins can be broad in their specificity. Some, for example, can block the activity of cysteine proteases. One of the best known biological serpins is α-1-anti-trypsin (A1AT - Figure 4.66) because of its role in lungs, where it functions to inhibit the elastase protease. Deficiency of A1AT leads to emphysema. This can arise as a result of genetic deficiency or by cigarette smoking. Reactive oxygen species produced by cigarette smoking can oxidize a critical methionine residue (#358 of the processed form) in A1AT, rendering it unable to inhibit elastase. Uninhibited, elastase can attack lung tissue and cause emphysema. Most serpins work extracellularly. In blood, for example, serpins like antithrombin can help to regulate the clotting process.

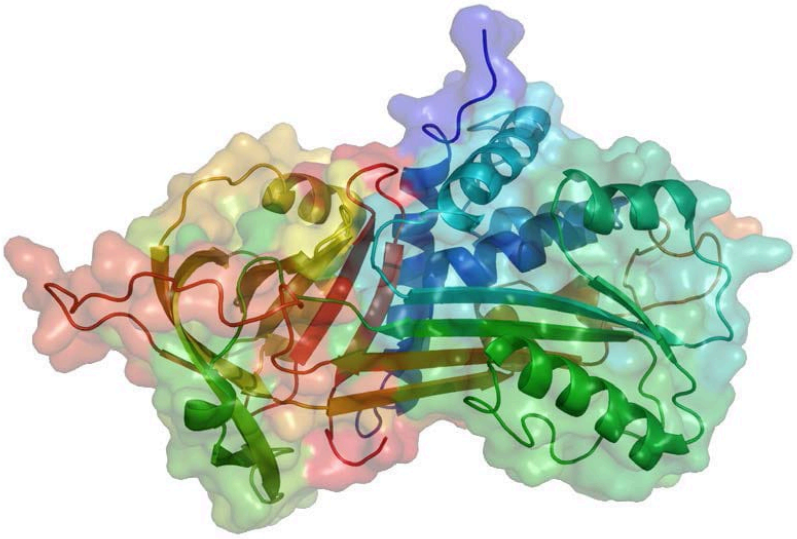


Figure 4.66 - α-1-antitrypsin

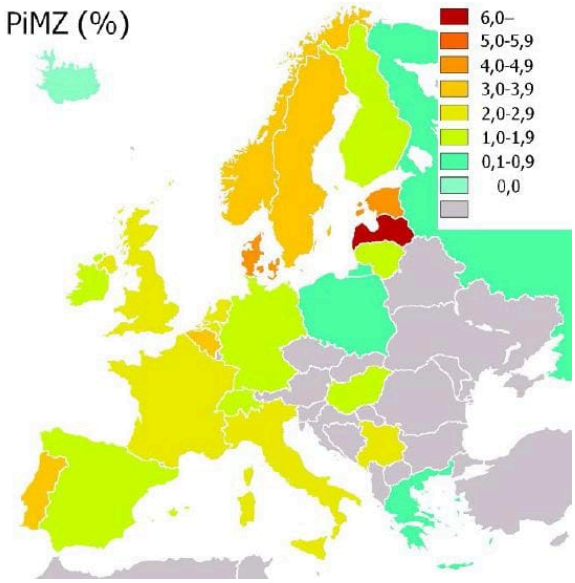


Figure 4.67 - Incidence of α-1-antitrypsin (PiMZ) deficiency in Europe by percent. Wikipedia

Anti-viral Agents

Protease inhibitors are used as anti-viral agents to prohibit maturation of viral proteins - commonly viral coat proteins.

They are part of drug “cocktails” used to inhibit the spread of HIV in the body and are also used to treat other viral infections, including hepatitis C. They have also been investigated for use in treatment of malaria and may have some application in anti-cancer therapies as well.

This page titled 4.3: Mechanisms of Catalysis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

4.4: Blood Clotting

A printable version of this section is here: BiochemFFA_4_4.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Clotting is a process in which liquid blood is converted into a gelatinous substance that eventually hardens. The aim is to stop the flow of blood from a vessel. The formation of a clot is the result of a series of enzymatic reactions that are triggered upon injury. The process involves:

1. a step of activation (wounding) followed by
2. a cellular response (aggregation of blood platelets) and
3. a molecular response (polymerization of the protein called fibrin to create a meshwork that hardens).

Factors released in the cellular response help activate the molecular response. The process is highly conserved across species.

Cellular Response

Injury to the epithelial lining of a blood vessel begins the process of coagulation almost instantly. The cellular response has an initial action followed by an amplification step. In the cellular response (Figure 4.68), the platelets bind directly to collagen using Ia/IIa collagen-binding surface receptors and glycoprotein VI to form a plug. The signal to the platelets to take this action is exposure of the underlying collagen, something that would not happen in the absence of a wound. Upon injury, platelet integrins get activated and bind tightly to the extracellular matrix to anchor them to the site of the wound.

The von Willebrand factor (see below also) assists by forming additional links between the platelets' glycoprotein Ib/IX/V and the fibrils of the collagen.

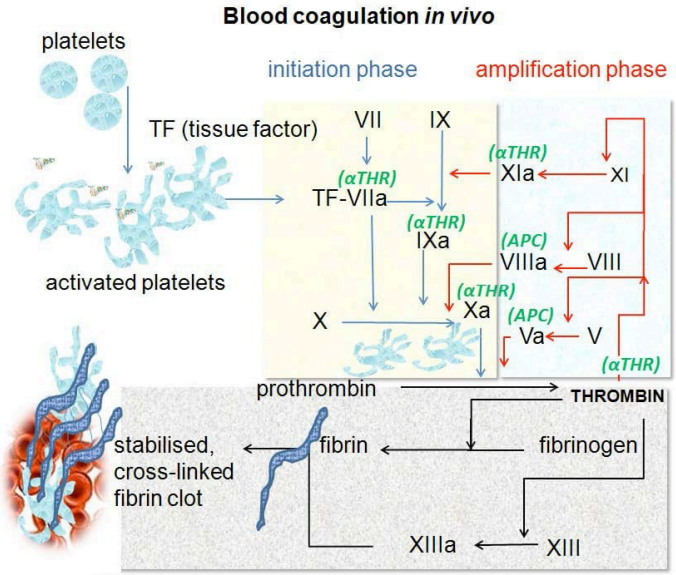


Figure 4.68 - Blood coagulation scheme. The cellular response is shown in white on upper left. Everything else is part of the molecular response. Wikipedia

Amplification

In the amplification part of the cellular response, the activated platelets release a large number of factors, including platelet factor 4 (a cytokine stimulating inflammation and moderating action of the heparin anticoagulant) and thromboxane A2, The latter has the effect of increasing the "stickiness" of platelets, favoring their aggregation. In addition, a a Gq-protein linked receptor cascade is activated, resulting in release of calcium from intracellular stores. This will play a role in the molecular response.

Molecular response

The molecular response results in the creation of a web comprised of polymers of fibrin protein. Like the cellular pathway, the molecular pathway begins with an initiation phase and continues with an amplification phase. Polymerization of fibrin results from convergence of two cascading catalytic pathways. They are the intrinsic pathway (also called the contact activation pathway) and the extrinsic pathway (also referred to as the tissue factor pathway). Of the two pathways, the tissue factor pathway has recently been shown to be the more important.

Serine Protease Cascade

In both pathways, a series of zymogens of serine proteases are sequentially activated in rapid succession. The advantage of such a cascading system is tremendous amplification of a small signal. At each step of the cascade, activation of a zymogen causes the production of a considerable amount of an active serine protease, which is then able to activate the next zymogen which, in turn, activates an even larger amount of the next zymogen in the system. This results in the ultimate activation of a tremendous amount more fibrin than could be achieved if there were only a single step where an enzyme activated fibrinogen to fibrin.

Nomenclature

The zymogen factors in the molecular response are generally labeled with Roman numerals. A lowercase, subscripted 'a' is used to designate an activated form.

The tissue factor pathway functions to create a thrombin burst, a process in which thrombin is activated very quickly. This is the initiation phase. It is fairly straightforward because it has one focus - activation of thrombin. Thrombin, which converts fibrinogen into the fibrin of the clot, is central also to the amplification phase, because it activates some of the factors that activate it, creating an enormous increase in signal and making a lot of thrombin active at once.

Initiation phase

The initiation phase of the molecular response begins when Factor VII (the letter 'F' before the Roman numeral is often used as an abbreviation for 'factor') gets activated to FVIIa after damage to the blood vessel (Figure 4.69 & 4.70). This happens as a result of its interaction with Tissue Factor (TF, also called coagulation Factor III) to make a TF-FVIIa complex. The combined efforts of TF-FVIIa, FIXa, and calcium (from the cellular response) inefficiently convert FX to FXa. FXa, FV, and calcium inefficiently convert prothrombin (zymogen) to thrombin (active). A tiny amount of thrombin has been activated at the end of the initiation phase.

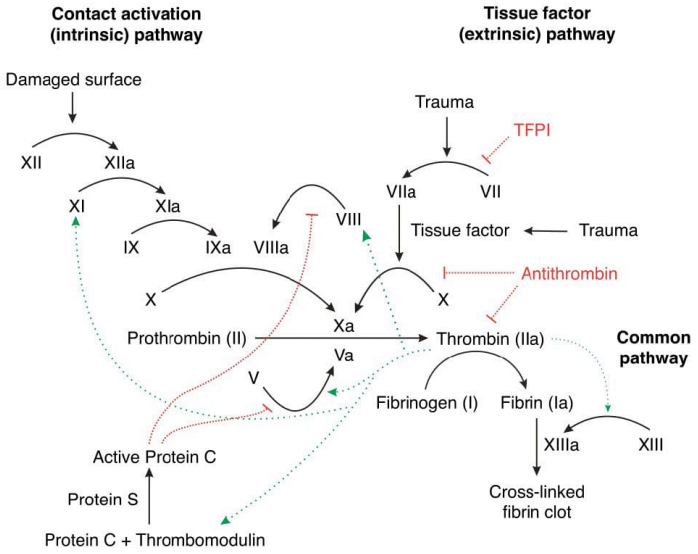


Figure 4.69 - Intrinsic and extrinsic pathways of blood coagulation. The aim is making a fibrin clot (lower right). Wikipedia

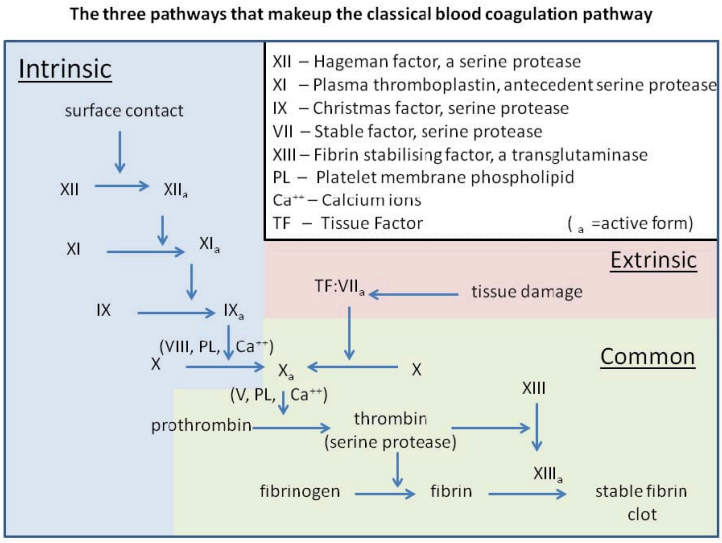


Figure 4.70 - Another view of the molecular response of the blood clotting pathway. Wikipedia

Amplification phase

To make sufficient thrombin to convert enough fibrinogen to fibrin to make a clot, thrombin activates other factors (FV, FXI, FVIII) that help to make more thrombin. This is the amplification phase of the molecular process and is shown in the light blue portion in the upper right part of Figure 4.68. The amplification phase includes factors in both the intrinsic and extrinsic pathways. FVIII is normally bound in a complex with the von Willebrand factor and is inactive until it is released by action of thrombin. Activation of FXI to FXIa helps favor production of more FIXa. FIXa plus FVIIIa stimulate production of a considerable amount of FXa. FVa joins FXa and calcium to make a much larger amount of thrombin. Factors FVa and FVIIIa are critical to the amplification process. FVIIIa stimulates FIXa's production of FXa by 3-4 orders of magnitude. FVa helps to stimulate FXa's production of thrombin by about the same magnitude. Thus, thrombin stimulates activation of factors that, in turn, stimulate activation of more thrombin.

Transglutaminase

In addition to helping to amplify product of itself and conversion of fibrinogen to fibrin, thrombin catalyzes the activation of FXIII to FXIIIa. FXIIIa is a transglutaminase that helps to "harden" the clot (Figure 4.71 & 4.73). It accomplishes this by catalyzing formation of a covalent bond between adjacent glutamine and lysine side chains in the fibrin polymers.

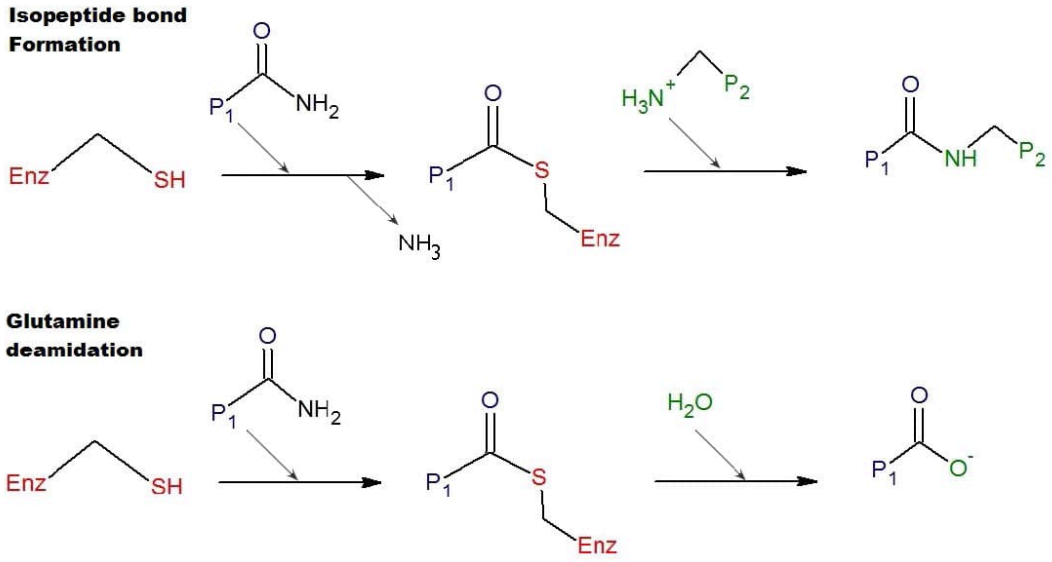


Figure 4.71 - Catalytic action of transglutaminase (top) and breaking of transglutamide bonds by hydrolysis (bottom)

Not all of the factors involved in the clotting process are activated by the pathway, nor are all factors serine proteases. This includes FVIII and FV which are glycoproteins, and FXIII, which is the transglutaminase described above.

The blood clotting process must be tightly regulated. Formation of clots in places where no damage has occurred can lead to internal clots (thrombosis) cutting off the flow of blood to critical regions of the body, such as heart or brain. Conversely, lack of clotting can lead to internal bleeding or, in severe cases, death due to unregulated external bleeding. Such is a danger for people suffering from hemophilia.

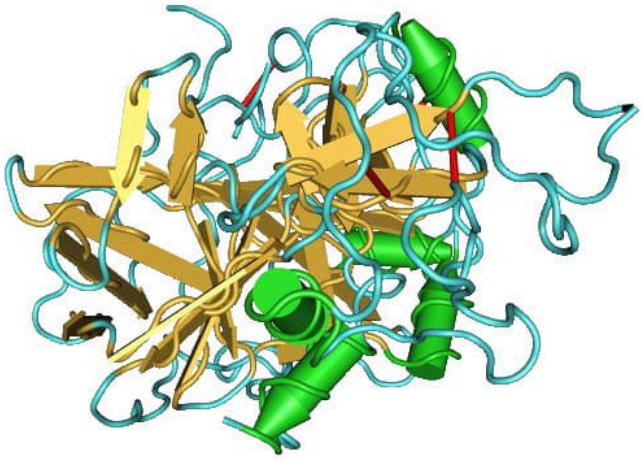


Figure 4.72 - α-thrombin. Wikipedia

Diseases of Blood Clotting: Hemophilia

Hemophilia is a hereditary genetic disorder affecting the blood clotting process in afflicted individuals. The disease is X-linked and thus occurs much more commonly in males. Deficiency of FVIII leads to Hemophilia A (about 1 in 5000 to 10,000 male births) and deficiency of FIX produces Hemophilia B (about 1 in 20,000 to 35,000 male births).

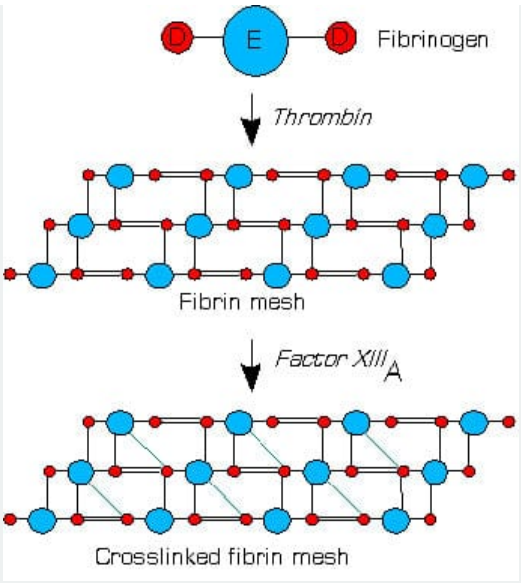


Figure 4.73 - Product of transglutaminase action - cross links. Wikipedia

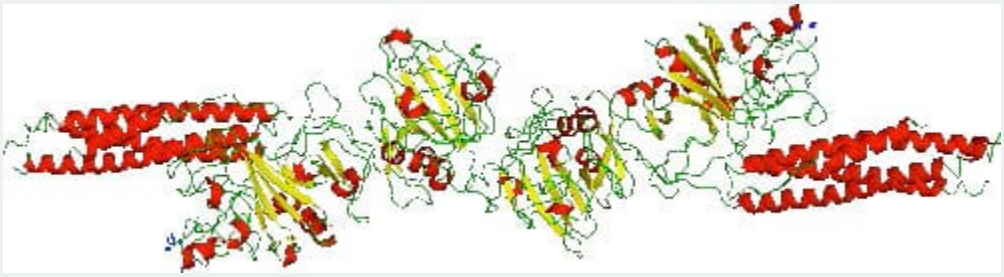


Figure 4.74 - Fibrin dimer - basic unit of a fibrin clot. Wikipedia

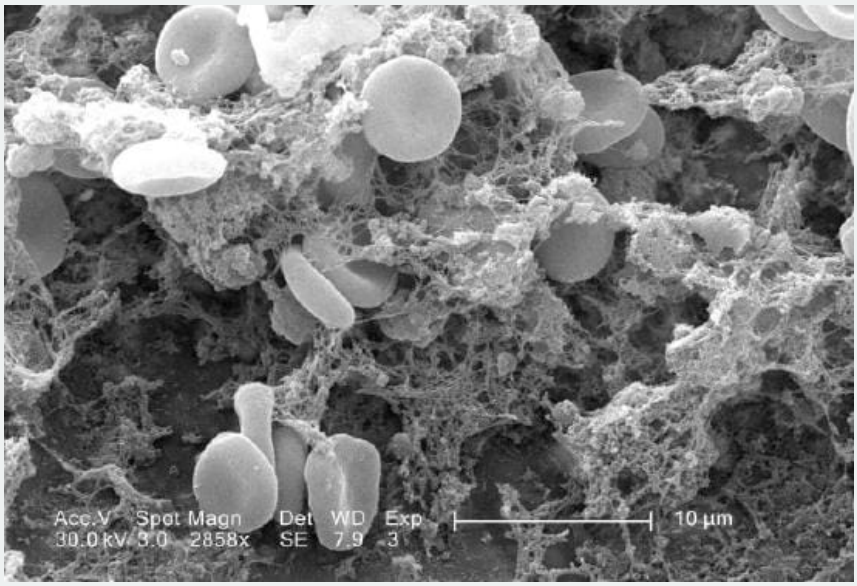


Figure 4.75 - Blood cells enmeshed in a fibrin clot

Hemophilia B spread through the royal families of Europe, beginning with Queen Victoria's son, Leopold. Three of the queen's grandsons and six of her great-grandsons suffered from the disease. Hemophilia is treated by exogenous provision of missing clotting factors and has improved life expectancy dramatically. In 1960, the life expectancy of a hemophiliac was about 11 years. Today, it is over 60.



Figure 4.76 - Queen Victoria, whose descendants suffered from hemophilia B

Diseases of Blood Clotting: von Willebrand's disease

A related disease to hemophilia that is also genetically linked is von Willebrand's Disease. The von Willebrand factor plays a role in both the cellular and the molecular responses in blood clotting. First, the factor is a large multimeric glycoprotein present in blood plasma and also is produced in the endothelium lining blood vessels.

The von Willebrand factor helps to anchor platelets near the site of the wound in the cellular response. It binds to several things. First, it binds to platelets' Ib glycoprotein. Second, it binds to heparin and helps moderate its action. Third, it binds to collagen and fourth, the factor binds to FVIII in the molecular response, playing a protective role for it. In the absence of the von Willebrand factor, FVIII is destroyed. Fifth, the von Willebrand factor binds to integrin of platelets, helping them to adhere together and form a plug. Defects of the von Willebrand factor lead to various various bleeding disorders.

Blood "thinners"

The clotting of blood is essential for surviving wounds that cause blood loss. However, some people have conditions that predispose them to the formation of clots that can lead to stroke, heart attack, or other problems, like *pulmonary embolism*. For these people, anti-clotting agents (commonly called blood thinners) are used to reduce the likelihood of undesired clotting.

The first, and more common of these is aspirin. Aspirin is an inhibitor of the production of prostaglandins. Prostaglandins are molecules with 20 carbons derived from arachidonic acid that have numerous physiological effects. Metabolically, the prostaglandins are precursors of a class of molecules called the thromboxanes. Thromboxanes play roles in helping platelets to stick together in the cellular response to clotting. By inhibiting the production of prostaglandins, aspirin reduces the production of thromboxanes and reduces platelet stickiness and the likelihood of clotting.

Vitamin K action

Another approach to preventing blood clotting is one that interferes with an important molecular action of Vitamin K. A pro-clotting factor found in the blood, vitamin K is necessary for an important modification to prothrombin and other blood clotting proteins. Vitamin K serves as an enzyme cofactor that helps to catalyze addition of an extra carboxyl group onto the side chain of glutamic acid residues of several clotting enzymes (see HERE). This modification gives them the ability to bind to calcium (Figure 4.77), which is important for activating the serine protease cascade. During the reaction that adds carboxyl groups to glutamate, the reduced form of vitamin K becomes oxidized. In order for vitamin K to stimulate additional carboxylation reactions to occur, the oxidized form of vitamin K must be reduced by the enzyme vitamin K epoxide reductase.

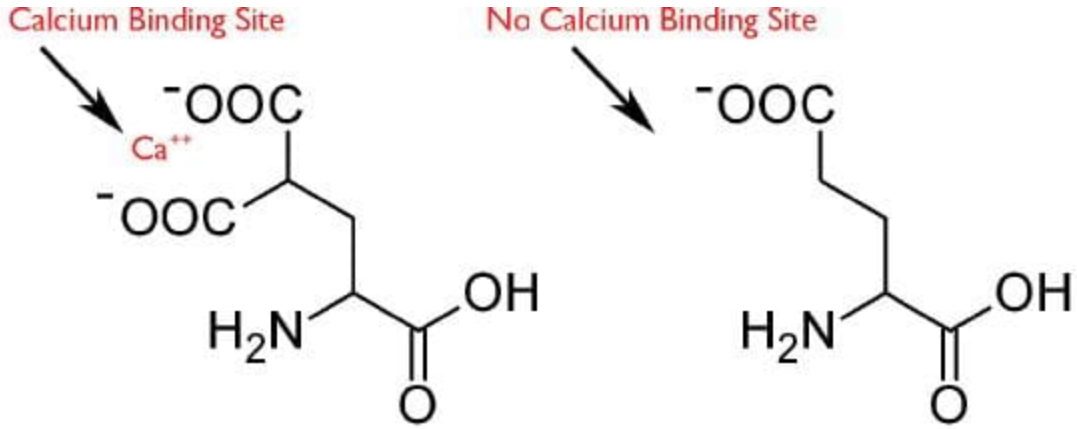


Figure 4.77 - γ-carboxylglutamic acid (left) has a calcium binding Site. Unmodified glutamic acid (right) does not.

Warfarin blocks reduction

The compound known as warfarin (brand name = coumadin - Figure 4.78) interferes with the action of vitamin K epoxide reductase and thus, blocks recycling of vitamin K. As a consequence, fewer prothrombins (and other blood clotting proteins) get carboxylated, and less clotting occurs.

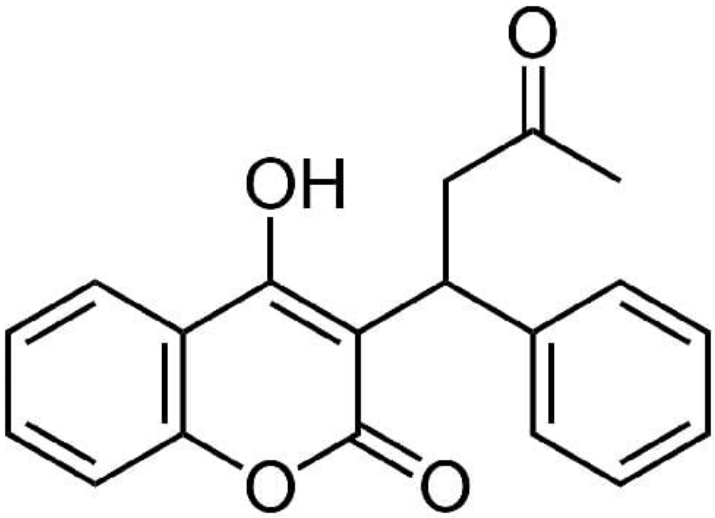


Figure 4.78 - Warfarin

Vitamin K-mediated carboxylation of glutamate occurs on the γ carbon of the amino acid's side chain, for 16 different proteins, 7 of which are involved in blood clotting, including prothrombin. When the carboxyl group is added as described, the side chain is able to efficiently bind to calcium ions. In the absence of the carboxyl group, the side chain will not bind to calcium. Calcium released near the site of the wound in the cellular response to clotting helps to stimulate activation of proteins in the serine protease cascade of the molecular response.

Vitamin K comes in several forms. It is best described chemically as a group of 2-methyl-1,4-naphthoquinone derivatives. There are five different forms recognized as vitamin Ks (K1, K2, K3, K4, and K5). Of these, vitamins K1 and K2 come from natural sources and the others are synthetic. Vitamin K2, which is made from vitamin K1 by gut microorganisms, has several forms, with differing lengths of of isoprenoid side-chains. The various forms are commonly named as MK-X, where X is a number, and MK stands for menaquinone, which is the name given to this form of vitamin K. Figure 4.79 shows a common form known as MK-4 (menatetrenone).

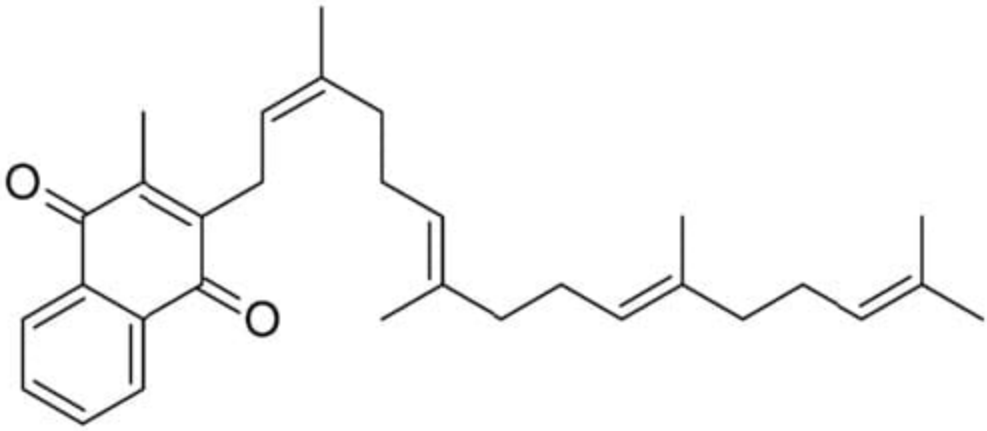


Figure 4.79 - MK-4 (menatetrenone)

Hemorrhaging danger

It is very critical that the proper amount of warfarin be given to patients. Too much can result in hemorrhaging. Patients must have their clotting times checked regularly to ensure that they are taking the right dose of anti-coagulant medication. Diet and the metabolism of Vitamin K in the body can affect the amount of warfarin needed. Vitamin K is synthesized in plants and plays a role in photosynthesis. It can be found in the highest quantities in vegetables that are green and leafy. Patients whose diet is high in these vegetables may require a different dose than those who rarely eat greens. Dietary vitamin K is also, as mentioned earlier, metabolized by bacteria in the large intestine, where they convert vitamin K1 into vitamin K2.

Plasmin

Clots, once made in the body, do not remain there forever. Instead, a tightly regulated enzyme known as plasmin is activated, when appropriate, to break down the fibrin-entangled clot. Like many of the enzymes in the blood clotting cascade, plasmin is a serine protease. It is capable of cleaving a wide range of proteins. They include polymerized fibrin clots, fibronectin, thrombospondin, laminin, and the von Willebrand factor.

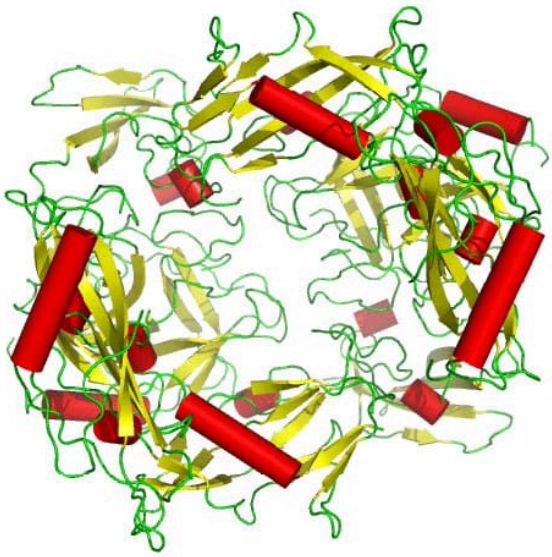


Figure 4.81 - Plasmin

Plasmin plays a role in activating collagenases and in the process of ovulation by weakening the wall of the Graafian follicle in the ovary. Plasmin is made in the liver as the zymogen known as plasminogen. Several different enzymes can activate it.

Tissue plasminogen activator (tPA), using fibrin as a co-factor, is one. Others include urokinase plasminogen activator (using urokinase plasminogen activator receptor as a co-factor), kallikrein (plasma serine protease with many forms and blood functions), and FXIa and FXIIa from the clotting cascade.

Plasmin inhibition

Plasmin's activity can also be inhibited. Plasminogen activator inhibitor, for example, can inactivate tPA and urokinase. After plasmin has been activated, it can also be inhibited by α2-antiplasmin and α2-macroglobulin (Figure 4.80). Thrombin also plays a role in plasmin's inactivation, stimulating activity of thrombin activatable fibrinolysis inhibitor. Angiostatin is a sub-domain of plasmin produced by auto-proteolytic cleavage. It blocks the growth of new blood vessels and is being investigated for its anti-cancer properties.

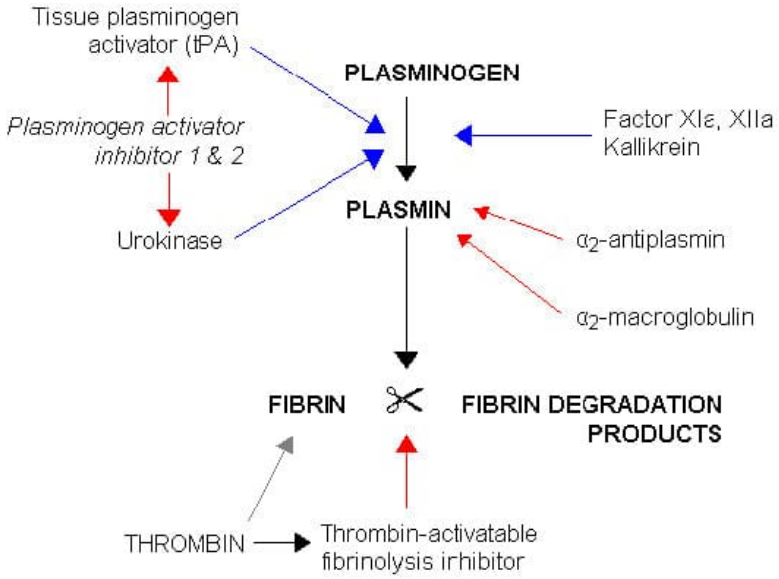


Figure 4.80 - Regulation of fibrin breakdown. Activators in blue. Inhibitors in red. Wikipedia

Fibronectin

Fibronectin is a large (440 kDa) glycoprotein found in the extracellular matrix that binds to integral cellular proteins called integrins and to extracellular proteins, including collagen, fibrin, and heparan sulfate. It comes in two forms. The soluble form is found in blood plasma and is made by the liver. It is found in high concentration in the blood stream (300 µg/ml). The insoluble form is found abundantly in the extracellular matrix.

The protein is assembled in the extracellular matrix and plays roles in cellular growth, adhesion, migration, and differentiation. It is very important in wound healing.

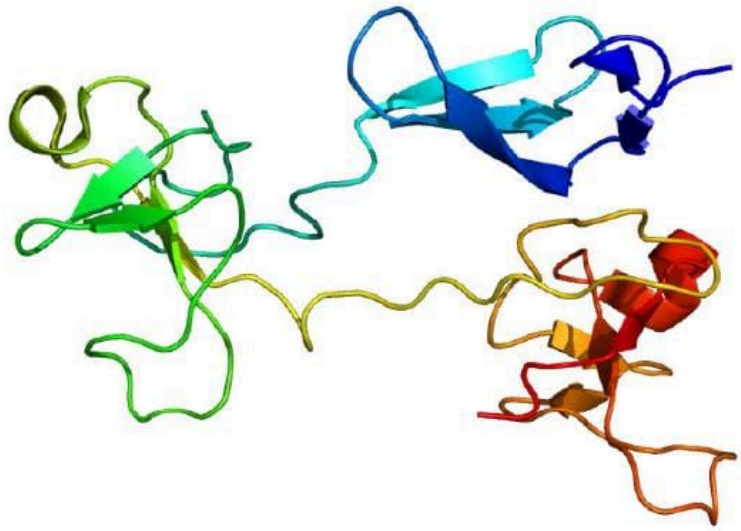


Figure 4.82 - Fibronectin 1. Wikipedia

Assists in blood clot formation

Fibronectin from the blood plasma is localized to the site of the wound, assisting in formation of the blood clot to stop bleeding. In the initial stages of wound healing, plasma fibronectin interacts with fibrin in clot formation. It also protects tissue surrounding the wound. Later in the repair process, remodeling of the damaged area begins with the action of fibroblasts and endothelial cells at the wound site. Their task is to degrade proteins of the blood clot matrix, replacing them with a new matrix like the undamaged, surrounding tissue.

Fibroblasts act on the temporary fibronectin-fibrin matrix, remodeling it to replace the plasma fibronectin with cellular fibronectin. This may cause the phenomenon of wound contraction, one of the steps in wound healing. Secretion of cellular fibronectin by fibroblasts is followed by fibronectin assembly and integration with the extracellular matrix.

Embryogenesis

Fibronectin is essential for embryogenesis. Deleting the gene in mice causes lethality before birth. This is likely due to its role in migration and guiding the attachment of cells as the embryo develops. Fibronectin also has a role in the mouth. It is found in saliva and is thought to inhibit colonization of the mouth by pathogenic bacteria.

Platelet activating factor

Platelet Activating Factor (PAF) is a compound (Figure 4.83) produced primarily in cells involved in host defense. These include platelets, macrophages, neutrophils, and monocytes, among others. It is produced in greater quantities in inflammatory cells upon proper stimulation. The compound acts like a hormone and mediates platelet aggregation/degranulation, inflammation, and anaphylaxis. It can transmit signals between cells to trigger and amplify inflammatory and clotting cascades.

When unregulated, signaling by PAF can cause severe inflammation resulting in sepsis and injury. Inflammation in allergic reactions arises partly as a result of PAF and is an important factor in bronchoconstriction in asthma. In fact, at a concentration of only 10 picomolar, PAF can cause asthmatic inflammation of the airways that is life threatening.

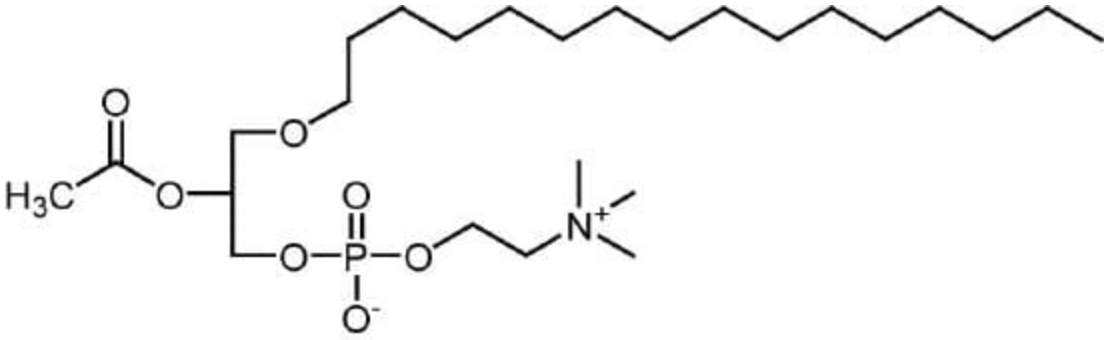


Figure 4.83 - Platelet Activating Factor. Wikipedia

I'm feeling so sad

'Cuz I cut . . . . myself bad

Now I'm all worried 'bout . . . . consequences

It's starting to bleed

There's some clo . . . . sure I need

So the body kicks . . . . in its defenses

It's happened all so many times before

The blood flows out and then it shuts the door

Thank goodness my blood is clotting

Enmeshing the fibrin chains

Thank goodness my blood is clotting

The zymogens

Are activating and all is well

So I'll stop bleeding again

The vitamin K's

Help to . . . . bind to cee-ays

Adding C-O-. . . . O-H to amend things

Um-m-um-um-um-um

It hardens and stays

When a glu. . . . taminase

Creates co. . . . valent bonds . . . . for cementing

In just a moment, things are good to go

The clot's in place and it has stopped the flow

But what about clot dissolving?

Untangling fibrin chains?

This calls for some problem solving

There is a way

Just activate up some t-PA

Get plasmin active in veins

Oh, oh, oh.

And thanks to the dis-enclotting'

As part of repairin' veins

It's part of my body's plotting

The wound is gone

I'm back where I started and

Nothing's wrong

My blood flow is normal again.

Thank Goodness My Blood is Clotting

To the tune of "Don't Sleep in the Subway Darling"

Metabolic Melodies Website HERE

Recording by Liz Bacon and David Simmons

Lyrics by Kevin Ahern
Recording by Liz Bacon and David Simmons Lyrics by Kevin Ahern

This page titled 4.4: Blood Clotting is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

5: ENERGY

5.1: Basics of Energy
5.3: Energy - Photophosphorylation
5.2: Electron Transport and Oxidative Phosphorylation

This page titled 5: Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

5.1: Basics of Energy

Source: BiochemFFA_5_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy



Living organisms are made up of cells, and cells contain a horde of biochemical components. Living cells, though, are not random collections of these molecules. They are extraordinarily organized or "ordered". By contrast, in the nonliving world, there is a universal tendency to increasing disorder. Maintaining and creating order in cells takes the input of energy. Without energy, life is not possible.

Oxidative Energy

The primary mechanism used by non-photosynthetic organisms to obtain energy is oxidation and carbon is the most commonly oxidized energy source. The energy released during the oxidative steps is “captured” in ATP and can be used later for energy coupling. The more reduced a carbon atom is, the more energy can be realized from its oxidation. Fatty acids are highly reduced, whereas carbohydrates are moderately so. Complete oxidation of both leads to carbon dioxide, which has the lowest energy state. Conversely, the more oxidized a carbon atom is, the more energy it takes to reduce it.

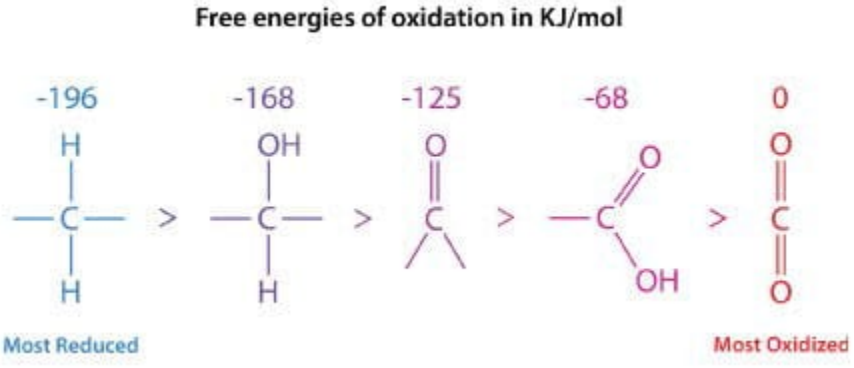


Figure 5.1.1: Five oxidation states of carbon. Image by Pehr Jacobson

In the series shown in Figure 5.1.1, the most reduced form of carbon is on the left. The energy of oxidation of each form is shown above it. The reduction states of fatty acids and carbohydrates can be seen by their formulas.

- Palmitic acid: $C_{16}H_{34}O_2$
- Glucose: $C_6H_{12}O_6$

Palmitic acid only contains two oxygens per sixteen carbons, whereas glucose has six oxygen atoms per six carbons. Consequently, when palmitic acid is fully oxidized, it generates more ATP per carbon (128/16) than glucose (38/6). It is because of this that we use fat (contains fatty acids) as our primary energy storage material.

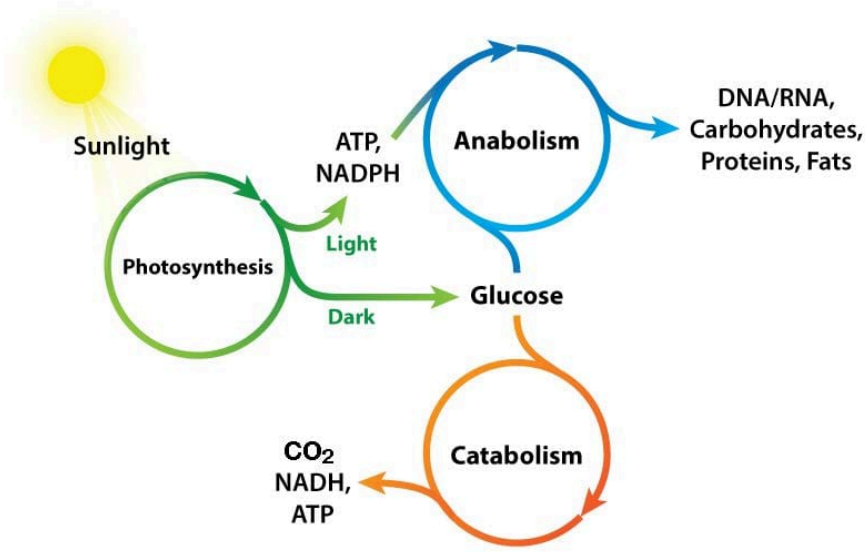


Figure $5.1.2$*: Photosynthesis: The primary source of biological energy. Image by Aleia Kim*

Oxidation vs. Reduction in Metabolism

Biochemical processes that break things down from larger to smaller are called *catabolic processes*. Catabolic processes are often oxidative in nature and energy releasing. Some, but not all, of that energy is captured as ATP. If not all of the energy is captured as ATP, what happens to the rest of it? The answer is simple. It is released as heat and it is for this reason we get hot when we exercise.

By contrast, synthesizing large molecules from smaller ones (for example, making proteins from amino acids) is referred to as anabolism. Anabolic processes are often reductive in nature (Figures 5.3 & 5.4) and require energy input. By themselves, they would not occur, as they are reversing oxidation and decreasing entropy (making many small things into a larger one). To overcome this energy barrier, cells must expend energy. For example, if one wishes to reduce CO_2 to carbohydrate, energy must be used to do so. Plants do this during the dark reactions of photosynthesis (Figure $5.1.3$). The energy source for the reduction is ultimately the sun. The electrons for the reduction come from water, and the CO_2 is removed from the atmosphere and gets incorporated into a sugar.

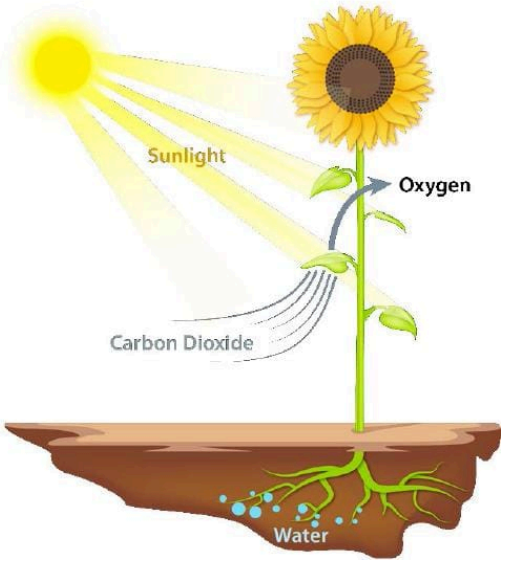


Figure 5.1.3: Movement of biological energy. Image by Aleia Kim

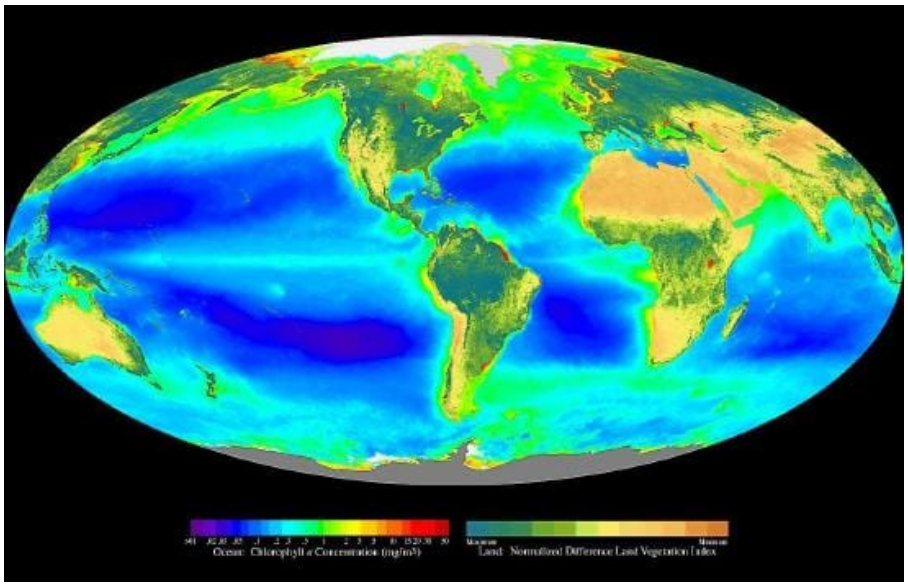


Figure 5.1.4: Photosynthesis as measured by chlorophyll concentration

Energy Coupling

The synthesis of the many molecules needed by cells needs the input of energy to occur. Cells overcome this energy obstacle by using ATP to “drive” the reaction (Figure 5.1.6). The energy needed to drive reactions is harvested in very controlled conditions in enzymes. This involves a process called ‘coupling’. Coupled reactions rely on linking an energetically favorable reaction (i.e., one with a negative $\Delta G^{\circ\prime}$) with the reaction requiring an energy input, which has a positive $\Delta G^{\circ\prime}$. As long as the overall $\Delta G^{\circ\prime}$ of the two reactions together is negative, the reaction can proceed. Hydrolysis of ATP is a very energetically favorable reaction that is commonly linked to many energy requiring reactions in cells. Without the hydrolysis of ATP (or GTP, in some cases), the reaction would not be feasible.

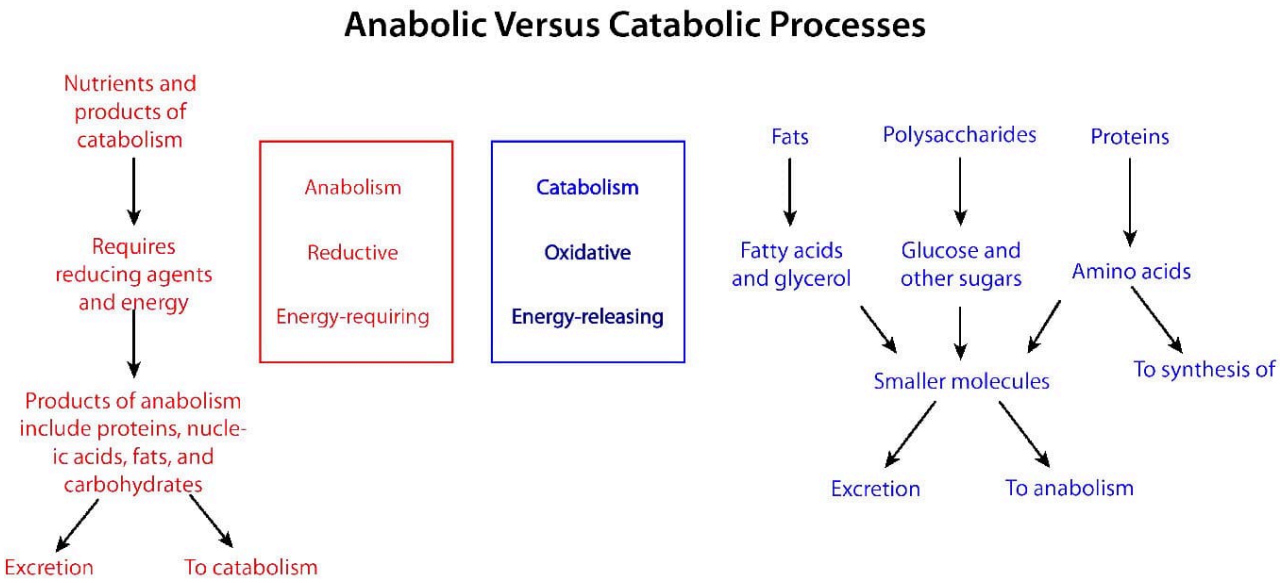


Figure 5.1.5: Synthesis and breakdown pathways in metabolism. Image by Pehr Jacobson

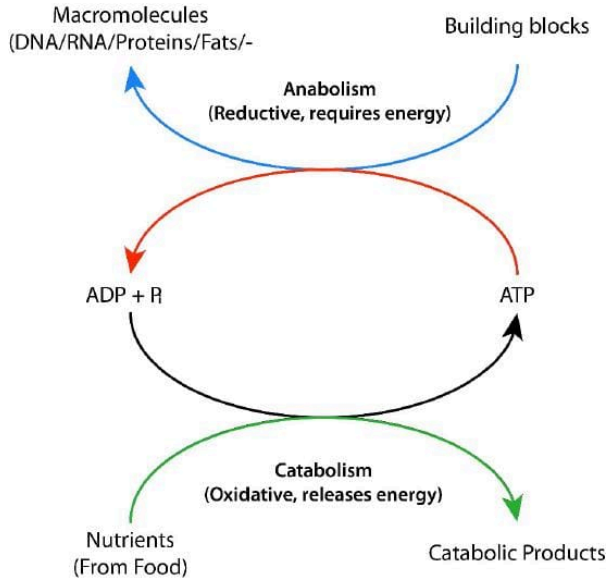


Figure 5.1.6: Cycling of biological energy via ADP and ATP. Image by Pehr Jacobson

Entropy and energy

Most students who have had some chemistry know about the Second Law of Thermodynamics with respect to increasing disorder of a system. Cells are very organized or ordered structures, leading some to mistakenly conclude that life somehow violates the second law. In fact, that notion is incorrect. The second law doesn't say that entropy always increases, just that, left alone, it tends to do so, in an isolated system. Cells are not isolated systems, though, in that they obtain energy, either from the sun, if they are autotrophic, or food, if they are heterotrophic.

To counter the universal tendency towards disorder on a local scale requires energy. As an example, take a fresh deck of cards which is neatly aligned with Ace-King-Queen . . . . 4,3,2 for each suit. Throw the deck into the air, letting the cards scatter. When you pick them up, they will be more disordered than when they started. However, if you spend a few minutes (and expend a bit of energy), you can reorganize the same deck back to its previous, organized state. If entropy always increased everywhere, you could not do this. However, with the input of energy, you overcame the disorder. This illustrates an important concept: the cost of fighting disorder is energy.

Biological energy

There are, of course, other reasons that organisms need energy. Muscular contraction, synthesis of molecules, neurotransmission, signaling, thermoregulation, and subcellular movements are examples. Where does this energy come from? The currencies of energy are generally high-energy phosphate-containing molecules. ATP is the best known and most abundant, but GTP is also an important energy source (energy source for protein synthesis). CTP is involved in synthesis of glycerophospholipids and UTP is used for synthesis of glycogen and other sugar compounds. In each of these cases, the energy is in the form of potential chemical energy stored in the multi-phosphate bonds. Hydrolyzing those bonds releases the energy in them.

Of the triphosphates, ATP is the primary energy source, acting to facilitate the synthesis of the others by action of the enzyme NDPK. ATP is made by three distinct types of phosphorylation – oxidative phosphorylation (in mitochondria), photophosphorylation (in chloroplasts of plants), and substrate level phosphorylation (in enzymatically catalyzed reactions).

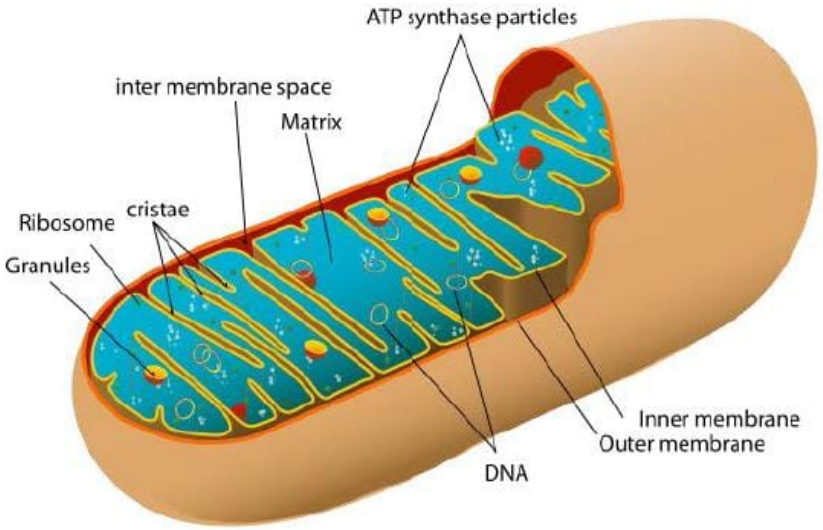


Figure 5.1.7: Mitochondrion

Gibbs free energy in Biology

ATP is generally considered the “storage battery” of cells (See also ‘Molecular Battery Backups for Muscles HERE). In order to understand how energy is captured, we must first understand Gibbs free energy and in doing so, we begin to see the role of energy in determining the directions chemical reactions take.

Gibbs free energy may be thought of as the energy available to do work in a thermodynamic system at constant temperature and pressure. Mathematically, the Gibbs free energy is given as:

$$G = H–TS \tag{5.1.1}$$

where H is the enthalpy, T is the temperature in Kelvin, and S is the entropy. At standard temperature and pressure, every system seeks to achieve a minimum of free energy. Thus, increasing entropy, S, will reduce Gibbs free energy. Similarly, if excess heat is available (reducing the enthalpy, H), the free energy can also be reduced.

Cells must work within the laws of thermodynamics, as noted, so all of their biochemical reactions, too, are ruled by these laws. Now we shall consider energy in the cell. The change in Gibbs free energy (ΔG) for a reaction is crucial, for it, and it alone, determines whether or not a reaction goes forward.

$$\Delta G = \Delta H–T\Delta S. \tag{5.1.2}$$

There are three cases

- ΔG < 0: the reaction proceeds as written

- ΔG = 0: the reaction is at equilibrium
- ΔG > 0: the reaction runs in reverse

For a reaction

$$a\,\mathrm{A} \rightleftharpoons b\,\mathrm{B} \tag{5.1.3}$$

(where 'a' and 'b' are integers and A and B are molecules) at pH 7, ΔG can be determined by the following equation,

$$\Delta G = \Delta G^{\circ\prime} + RT\ln(\frac{[B]^b}{[A]^a}) \tag{5.1.4}$$

For multiple substrate reactions, such as

$$a\,\mathrm{A} + c\,\mathrm{C} \rightleftharpoons b\,\mathrm{B} + d\,\mathrm{D} \tag{5.1.5}$$

$$\Delta G = \Delta G^{\circ\prime} + RT\ln(\frac{[B]^b[D]^d}{[A]^a[C]^c}) \tag{5.1.6}$$

The ΔG°' term is called the change in Standard Gibbs Free energy, which is the change in energy that occurs when all of the products and reactants are at standard conditions and the pH is 7.0. It is a constant for a given reaction.

In simple terms, we can collect all of the terms of the numerator together and call them {Products} and all of the terms of the denominator together and call them {Reactants},

$$\Delta G = \Delta G^{\circ\prime} + RT\ln(\frac{\{\mathrm{Products}\}}{\{\mathrm{Reactants}\}}) \tag{5.1.7}$$

For most biological systems, the temperature, T, is a constant for a given reaction. Since ΔG°' is also a constant for a given reaction, the ΔG is changed almost exclusively as the ratio of {Products}/{Reactants} changes.

Importance of ΔG°'

If one starts out at standard conditions, where everything except protons is at 1M, the RTln({Products}/{Reactants}) term is zero, so the ΔG°' term equals the ΔG, and the ΔG°' determines the direction the reaction will take (only under those conditions). This is why people say that a negative ΔG°' indicates an energetically favorable reaction, whereas a positive ΔG°' corresponds to an unfavorable one.

Increasing the ratio of {Products}/{Reactants} causes the value of the natural log (ln) term to become more positive (less negative), thus making the value of ΔG more positive. Conversely, as the ratio of {Products}/{Reactants} decreases, the value of the natural log term becomes less positive (more negative), thus making the value of ΔG more negative.

System response to stress

Intuitively, this makes sense and is consistent with Le Chatelier's Principle – a system responds to stress by acting to alleviate the stress. If we examine the ΔG for a reaction in a closed system, we see that it will always move to a value of zero (equilibrium), no matter whether it starts with a positive or negative value.

Another type of free energy available to cells is that generated by electrical potential. For example, mitochondria and chloroplasts partly use Coulombic energy (based on charge) from a proton gradient across their membranes to provide the necessary energy for the synthesis of ATP. Similar energies drive the transmission of nerve signals (sodium and potassium gradients) and the movement of some molecules in secondary active transport processes across membranes (e.g., H+ differential driving the movement of lactose). From the Gibbs free energy change equation,

$$\Delta G = \Delta H – T\Delta S \tag{5.1.8}$$

it should be noted that an increase in entropy will help contribute to a decrease in ΔG. This happens, for example when a large molecule is being broken into smaller pieces or when the rearrangement of a molecule increases the disorder of molecules around it. The latter situation arises in the hydrophobic effect, which helps drive the folding of proteins.

Chemical and electrical potential

It is said that absence makes the heart grow fonder. We won't tackle that philosophical issue here, but we will say that separation provides potential energy that cells can and do harvest. The lipid bilayer of cell and (in eukaryotic cells) organelle membranes provide the necessary barrier for separation.

Impermeable to most ions and polar compounds, biological membranes are essential for processes that generate cellular energy. Consider Figure 5.8. A lipid bilayer separates two solutions with different concentrations of a solute. There is a greater concentration of negative ions in the bottom and a greater concentration of positive ions on the top.

Whenever there is a difference in concentration of molecules across a membrane, there is said to be a concentration gradient across it. A difference in concentration of ions across a membrane also creates a charge (or electrical) gradient. Because there is a difference in both the chemical concentration of the ions and in the charge on the two sides of the membrane, this is described as an electrochemical gradient (Figures 5.8 -5.10).

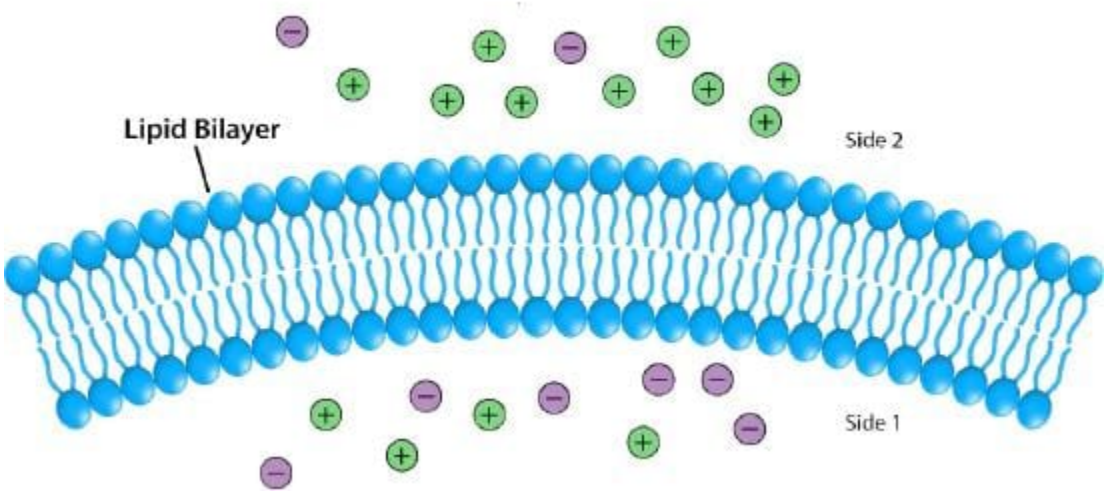


Figure 5.8: Differences in ion concentration across a membrane give rise to chemical and electrical gradients. Image by Pehr Jacobson

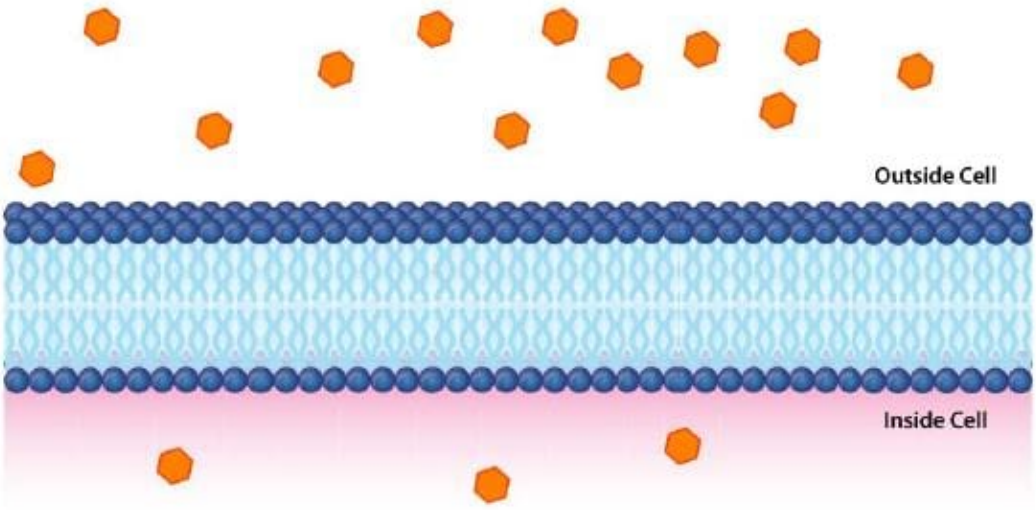


Figure 5.1.9: A chemical gradient. Image by Aleia Kim

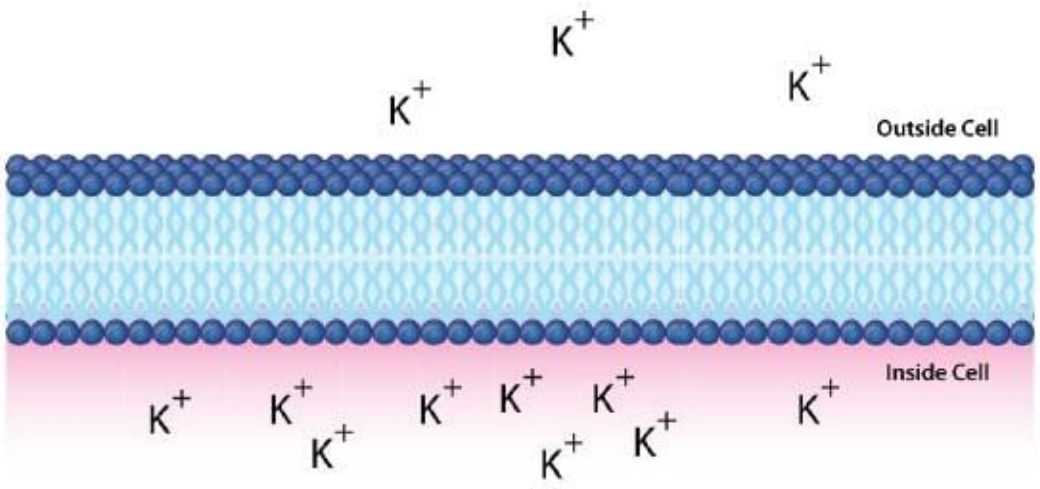


Figure 5.1.10: An electrochemical gradient of potassium ions. Image by Aleia Kim

Potential energy

Such gradients function like batteries and contain potential energy. When the potential energy is harvested by cells, they can create ATP, transmit nerve signals, pump molecules across membranes, and more. It is important, therefore, to understand how to calculate the potential energy of electrochemical gradients.

First, we consider chemical (solute) gradients. In Figure 5.9, two concentrations of glucose are separated by a lipid bilayer. Let's assume C2 be the concentration of glucose inside the cell (bottom) and C1 be the glucose concentration outside (top). The Gibbs free energy associated with moving glucose in the direction of C2 (into the cell) is given by

ΔG = RTln[C2/C1]

To move it in the direction of C1 (to the outside of the cell) the expression would be

$$\Delta G = RT \ln[C_1/C_2] \tag{5.1.9}$$

Since C2 is smaller than C1 (i.e., there are fewer glucose molecules inside the cell) then the ΔG is negative and diffusion would be favored into the cell, if the glucose could traverse the bilayer.

Conversely, if C2 was greater than C1 (more glucose was in the cell than outside) the ΔG would be positive, so movement in the direction of C2 would not be favored and instead the glucose would tend to move towards C1 , that is, out of the cell.

If C2 = C1, with the same concentration of glucose inside and outside, then the ΔG would be zero and there would be no net movement, as the system would be at equilibrium.

In the example above, we considered glucose, which is an uncharged molecule. When ions are involved, their charges must also be taken into consideration. Figure 5.1.10 depicts a similar situation across a lipid bilayer. In this case, a difference of concentration and charge exists. There are more positive charges inside the cell than outside.

Using C2 to indicate the concentration of materials inside the cell and C1 for the concentration outside the cell (as before), then the free energy for movement of an ion from top to bottom is given by the following equation

$$\Delta G = RT \ln[C_2/C_1] + ZF\Delta\psi \tag{5.1.10}$$

Note here that this equation must take into consideration both the concentration differences and the charge differences. Z refers to the charge of the transported species, F is the Faraday constant (96,485 Coulombs/mol), and Δψ is the electrical potential difference (voltage difference) across the membrane.

If we were to calculate the ΔG for movement of the potassium ion from top to bottom, it would be positive, since [C2/C1] is greater than 1 (making for a positive ln term), and the ZFΔψ is positive because positively charged ions (Z) are moving against a positive charge gradient given by Δψ (greater concentration at the target (bottom) than the starting point (top)). If we were to calculate the concentration of ions moving from bottom to top, then the ln term would be negative (C2

Reduction Potential

In discussing chemical potential, we must also consider reduction potential. Reduction potential measures the tendency of a chemical to be reduced by electrons. It is also designated by several other names/variables. These include redox potential, oxidation/reduction potential, ORP, pE, ε, E, and Eh.

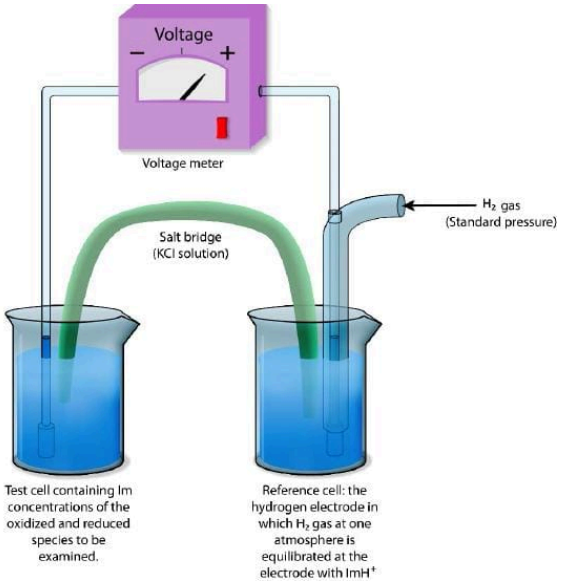


Figure 5.1.11: Reference electrode for measuring reduction potentials. Image by Pehr Jacobson

Reduction potential is measured in volts, or millivolts. A substance with a higher reduction potential will have a greater tendency to accept electrons and be reduced. If two substances are mixed in an aqueous solution, the one with the greater (more positive) reduction potential will tend to take electrons away, thus being reduced, from the one with the lower reduction potential, which becomes oxidized.

Relative measures

Absolute reduction potentials are difficult to measure, so reduction potentials are typically defined relative to a reference electrode. In aqueous solutions, reduction potentials are measured as the potential difference between an inert sensing electrode (typically platinum) in contact with the test solution and a stable reference electrode (measured as a Standard Hydrogen Electrode: SHE) as shown in Figure 5.1.11. The standard of reference for measurement is the half-reaction

$$H^+ + e^- \rightarrow ½\ H_2$$

The electrode where this reaction occurs (referred to as a half-cell) is given the value of E° (Standard Reduction Potential) of 0.00 volts. The hydrogen electrode is connected via an external circuit to another half cell containing a mixture of the reduced and oxidized species of another molecule (for example, Fe^{++} and Fe^{+++}) at 1M each and standard conditions of temperature (25°C) and pressure (1 atmosphere).

Direction and voltage measured

The direction and magnitude of electron movement is then measured. If the test mixture takes electrons from the hydrogen electrode, the sign of the voltage is positive and if the direction is reversed, the voltage is negative.

Thus, compounds which have greater affinity for electrons than hydrogen will register a positive voltage and negative voltages correspond to compounds with lesser affinity for electrons than hydrogen.

Movement of electrons

Under standard conditions, electrons will move from compounds generating lower voltages to ones generating higher (more positive) voltages. Just as the standard Gibbs free energy change is the Gibbs free energy change under standard conditions, so, too, is the standard reduction potential E° the reduction potential E under standard conditions.

The actual reduction potential of a half cell will vary with the concentration of each chemical species in the cell. The relationship between the reduction potential E and the standard reduction potential E° is given by the following equation (also called the Nernst equation)

$$E = E^\circ + (RT/nF)*\ln\left(\frac{[\text{reduced molecule}]}{[\text{oxidized molecule}]}\right)$$

where F is the Faraday constant (96,480 J/(Volts*moles), R is the gas constant (8.315 J/(moles*K), n is the number of moles of electrons being transferred, and T is the absolute temperature in Kelvin.

At 25°C, this equation becomes

$$E = E^\circ + (0.026V/n)*\ln\left(\frac{[\text{reduced molecule}]}{[\text{oxidized molecule}]}\right)$$

As for Gibbs free energy, it is useful to measure values at conditions found in cells. This means doing measurements at pH = 7, which differs from having all species at 1M.

Adjustment

Because of this adjustment, a slightly different standard reduction potential is defined and we designate it by E°', just as we defined a special standard Gibbs free energy change at pH 7 as ΔG°'.

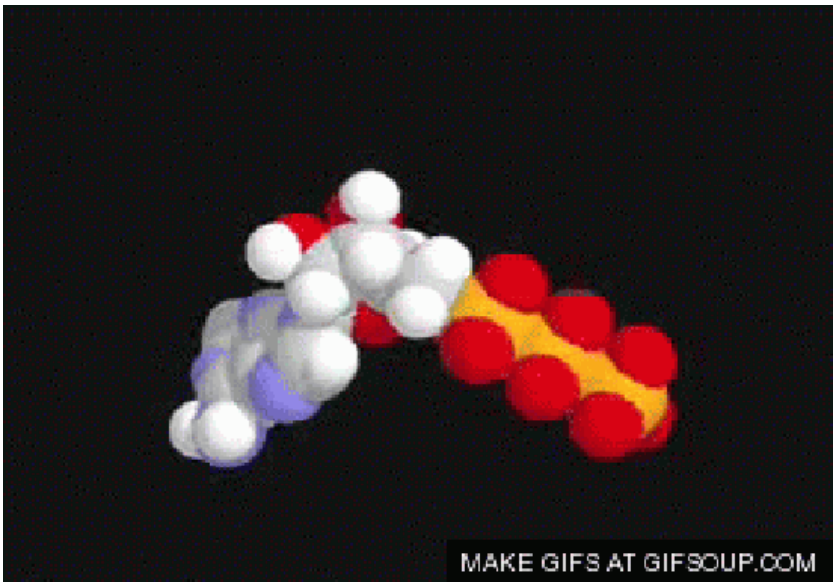
There is a relationship between the change in Gibbs free energy ΔG and the change in reduction potential (ΔE). It is

$$\Delta G = -nF\Delta E \tag{5.1.11}$$

Similarly, the relation between the change in standard Gibbs free energy and the change in standard reduction potential is

\]ΔG°' = -nFΔE°'\]

Energy Storage in Triphosphates



Movie 5.1: ATP: The fuel of the cell

Formation of triphosphates, like ATP, is essential to meeting the cell's energy needs for synthesis, motion, and signaling. In a given day, an average human body makes and breaks down more than its weight in triphosphates. This is especially remarkable considering that there is only about 250 g of the molecule present in the body at any given time. Energy in ATP is released by hydrolysis of a phosphate from the molecule.

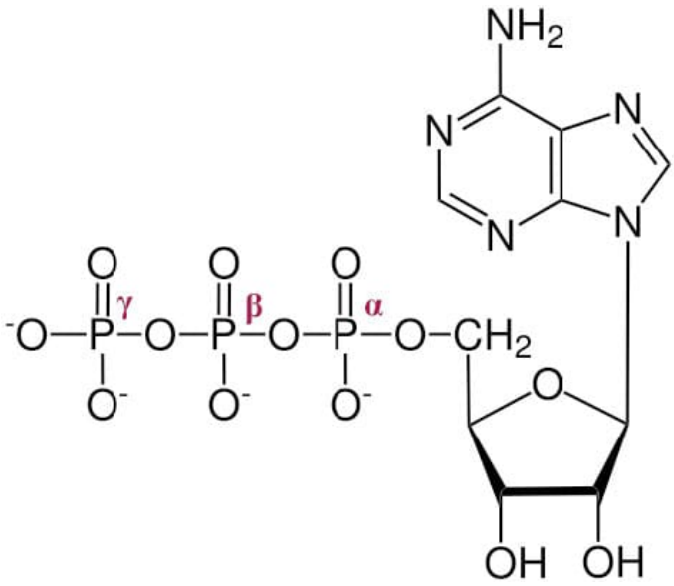


Figure 5.1.12 ATP showing α, β and γ phosphates

The three phosphates, starting with the one closest to the sugar are referred to as α, β, and γ (Figure 5.1.12). It is the γ phosphate that is cleaved in hydrolysis and the product is ADP. In a few reactions, the bond between the α and β is cleaved. When this happens, a pyrophosphate (β linked to γ) is released and AMP is produced. This latter reaction to produce AMP releases more energy ($\Delta G^{\circ\prime}$ = -45.6 kJ/mol) than the first reaction which produces ADP ($\Delta G^{\circ\prime}$ = -30.5 kJ/mol).

Since triphosphates are the "currency" that meet immediate needs of the cell, it is important to understand how triphosphates are made. There are three phosphorylation mechanisms – 1) substrate level; 2) oxidative; and 3) photophosphorylation. We consider them here individually.

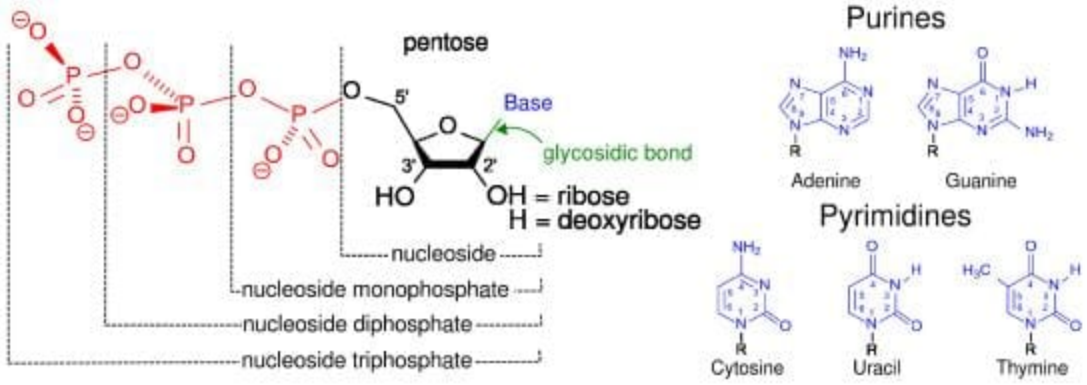


Figure 5.1.13: Nucleotides, nucleosides, and bases

Substrate level phosphorylation

The easiest type of phosphorylation to understand is that which occurs at the substrate level. This type of phosphorylation involves the direct synthesis of ATP from ADP and a high energy intermediate, typically a phosphate-containing molecule. Substrate level phosphorylation is a relatively minor contributor to the total synthesis of triphosphates by cells. An example substrate phosphorylation comes from glycolysis.

Phosphoenolpyruvate (PEP) + ADP ⇌ Pyruvate + ATP

This reaction has a very negative ΔG°' (-31.4 kJ/mol), indicating that the PEP contains more energy than ATP, thus tending to energetically favor ATP's synthesis. Other triphosphates can be made by substrate level phosphorylation, as well. For example, GTP can be synthesized by the following citric acid cycle reaction.

Succinyl-CoA + GDP + Pi ⇌ Succinate + GTP + CoA-SH

Triphosphates can be interchanged readily in substrate level phosphorylations catalyzed by the enzyme Nucleoside Diphosphate Kinase (NDPK). A generalized form of the reactions catalyzed by this enzyme is as follows:

XTP + YDP ⇌ XDP + YTP

where X = adenosine, cytidine, uridine, thymidine, or guanosine and Y can be any of these as well. Further, XTP and YDP can be any of the deoxynucleotides as well.

Last, an unusual way of synthesizing ATP by substrate level phosphorylation is via the reaction catalyzed by adenylate kinase

2 ADP ⇌ ATP + AMP

ATP source

This reaction is an important means of generating ATP when the cell doesn't have other sources of energy. Accumulation of AMP resulting from this reaction activates enzymes, such as phosphofructokinase, of glycolysis, which will catalyze reactions to give the cell additional, needed energy.

It is important to note that enzymes cannot make reactions happen that are energetically unfavorable. Enzymes speed reactions, but do not change their direction. Cells are thus bound by the rules of Gibbs free energy. So, how do energetically unfavorable reactions happen in a cell?

Reaction coupling

Reactions that are energetically unfavorable, can be made favorable by coupling them with the hydrolysis of ATP, a very energetically favorable reaction. There are numerous parallels in the "real world." Movement of automobiles is energetically unfavorable, but coupling movement of the automobile to oxidation of gasoline makes an unfavorable process favorable. Another

approach to making an unfavorable reaction favorable is to manipulate the concentration of reactants and products. Consider the reaction below, which occurs in pyrimidine nucleotide metabolism:

orotate + PRPP ⇌ OMP + PP_i

The ΔG°' for this reaction is -0.8 kJ/mol, meaning that if one starts with equal concentrations of reactants and products, at equilibrium, there will be a small excess of products. In the cell, however, this reaction moves strongly to the right (ΔG = very negative). Given that the ΔG°' is very close to zero, a very negative ΔG can only occur if the concentrations of reactants and products are altered, since

$$\Delta G = \Delta G^{\circ\prime} + RT\ln(\frac{[\mathrm{OMP}][\mathrm{PP_i}]}{[\mathrm{Orotate}][\mathrm{PRPP}]}) \tag{5.1.12}$$

Manipulation is exactly what happens here. The key item whose concentration is adjusted in this reaction is the pyrophosphate (PPi). This is possible because cells contain an enzyme called pyrophosphorylase that catalyzes the following reaction

$PP_i + H_2O$ ⇌ 2 P_i

Hydrolysis of pyrophosphate is very energetically favored, causing the PPi produced in the reaction to be quickly hydrolyzed. As a result, the concentration of PPi in the cell is kept very low. A low concentration of a product (PPi) causes the natural log (ln) term of the orotate equation to become more negative, driving the ΔG term for the overall reaction to become much more negative.

Pushing and pulling

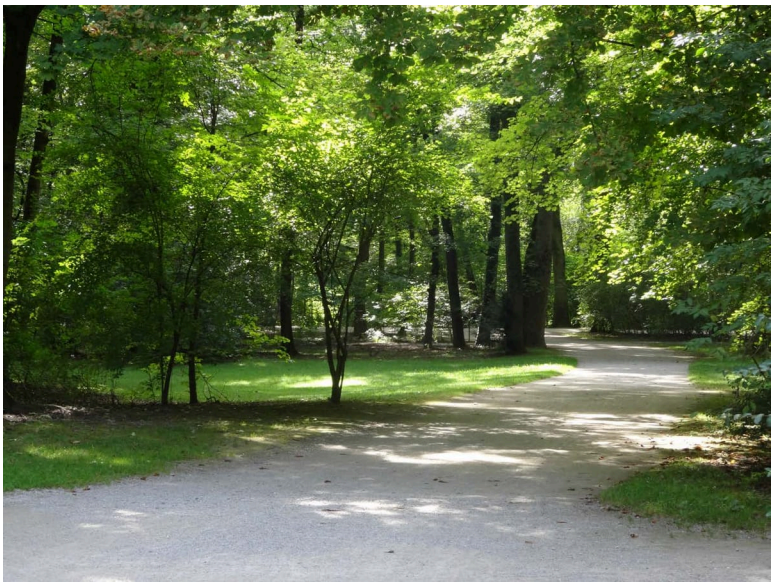
Reactions that yield pyrophosphate as a product are produced in the synthesis of DNA and RNA, as well as many other molecules. As shown in the previous example, this pyrophosphate is rapidly hydrolyzed, causing the overall reaction to move in the direction of pyrophosphate production. When reactants are removed/reduced in a metabolic reaction to decrease the concentration of a product, we say that the reaction is “pulled”, to represent the increase in the forward reaction as a result of product depletion.

Pushing happens when reactants in a reaction are added/increased. This too has the effect of reducing the ΔG of a reaction and making it more favorable because the ratio of [Products]/[Reactants] is decreased with increasing [Reactants]. Pushing and pulling of reactions are additional tools for cells to overcome energy barriers, just like coupling energetically favorable processes to energetically unfavorable ones.

This page titled 5.1: Basics of Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

5.3: Energy - Photophosphorylation

Source: BiochemFFA_5_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy



Photophosphorylation

The third type of phosphorylation to make ATP is found only in cells that carry out photosynthesis. This process is similar to oxidative phosphorylation in several ways. A primary difference is the ultimate source of the energy for ATP synthesis. In oxidative phosphorylation, the energy comes from electrons produced by oxidation of biological molecules. In photosynthesis, the energy comes from the light of the sun. Photons from the sun interact with chlorophyll molecules in reaction centers in the chloroplasts (Figures $5.3.1$ and $5.3.2$) of plants or membranes of photosynthetic bacteria.

The similarities of photophosphorylation to oxidative phosphorylation include:

- a membrane associated electron transport chain
- creation of a proton gradient
- harvesting energy of the proton gradient by making ATP with the help of an ATP synthase.

Some of the differences include :

- the source of the electrons – H2O for photosynthesis versus NADH/FADH2 for oxidative phosphorylation
- direction of proton pumping – into the thylakoid space of the chloroplasts versus outside the matrix of the mitochondrion
- movement of protons during ATP synthesis – out of the thylakoid space in photosynthesis versus into the mitochondrial matrix in oxidative phosphorylation
- nature of the terminal electron acceptor – NADP+ in photosynthesis versus O2 in oxidative phosphorylation.

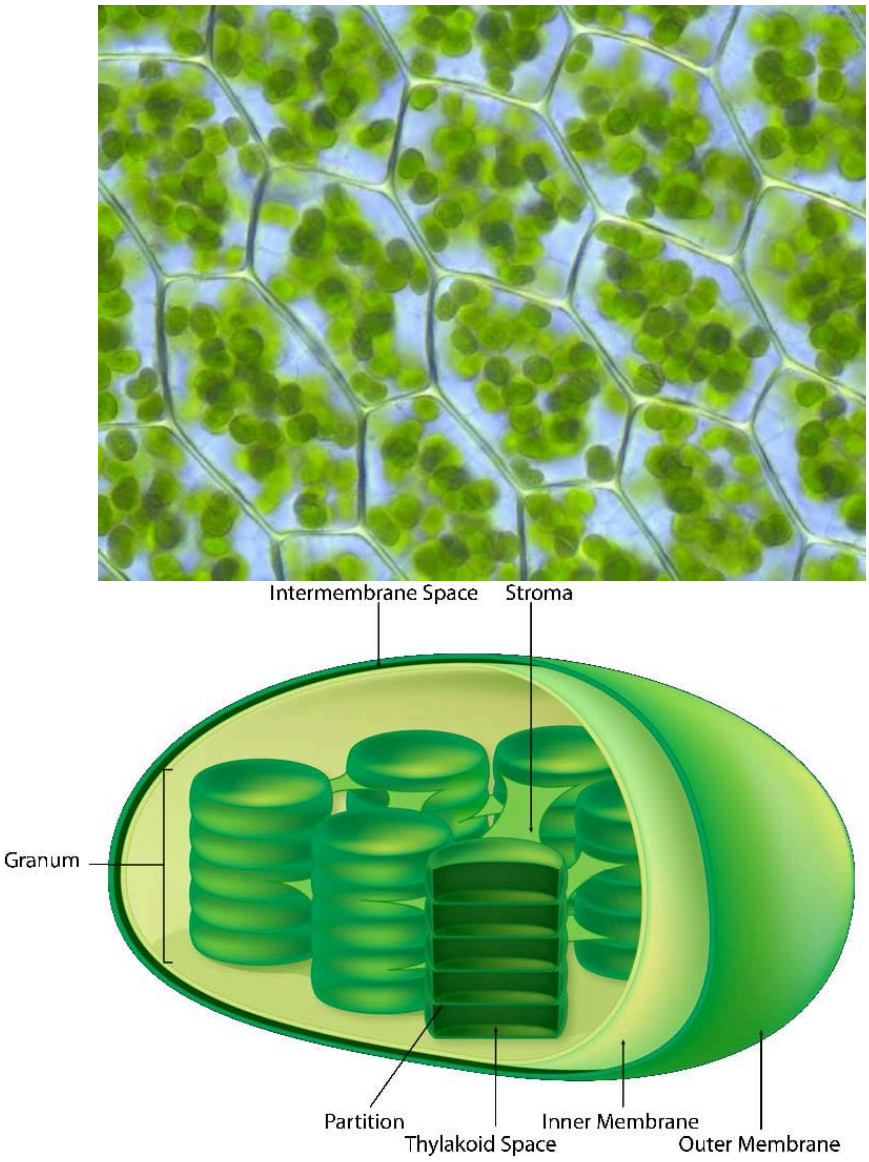


Figure 5.3.2: *The chloroplast. Image by Aleia Kim*

Electron transport: chloroplasts vs mitochondria

In some ways, the movement of electrons in chloroplasts during photosynthesis is opposite that of electron transport in mitochondria. In photosynthesis, water is the source of electrons and their final destination is NADP+ to make NADPH. In mitochondria, NADH/FADH2 are electron sources and H2O is their final destination. How do biological systems get electrons to go both ways? It would seem to be the equivalent of going to and from a particular place while always going downhill, since electrons will move according to potential.

Solar power

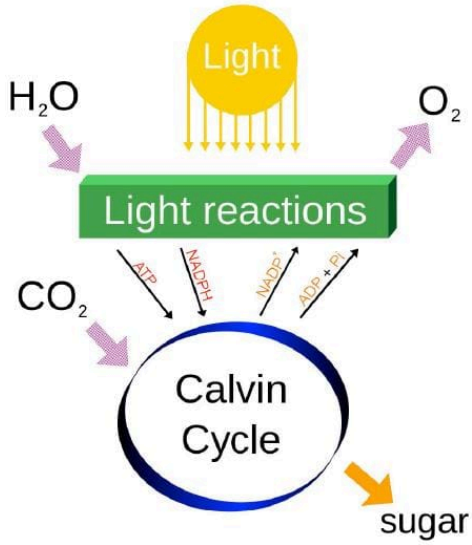


Figure 5.3.3: *Overview of photosynthesis. Wikipedia*

The answer is the captured energy of the photons from the sun (Figure 5.59), which elevates electrons to an energy where they move “downhill” to their NADPH destination in a Z-shaped scheme. The movement of electrons through this scheme in plants requires energy from photons in two places to “lift” the energy of the electrons sufficiently.

Last, it should be noted that photosynthesis actually has two phases, referred to as the light cycle (described above) and the dark cycle, which is a set of chemical reactions that captures CO2 from the atmosphere and “fixes” it, ultimately into glucose. The dark cycle is also referred to as the Calvin Cycle and is discussed HERE.

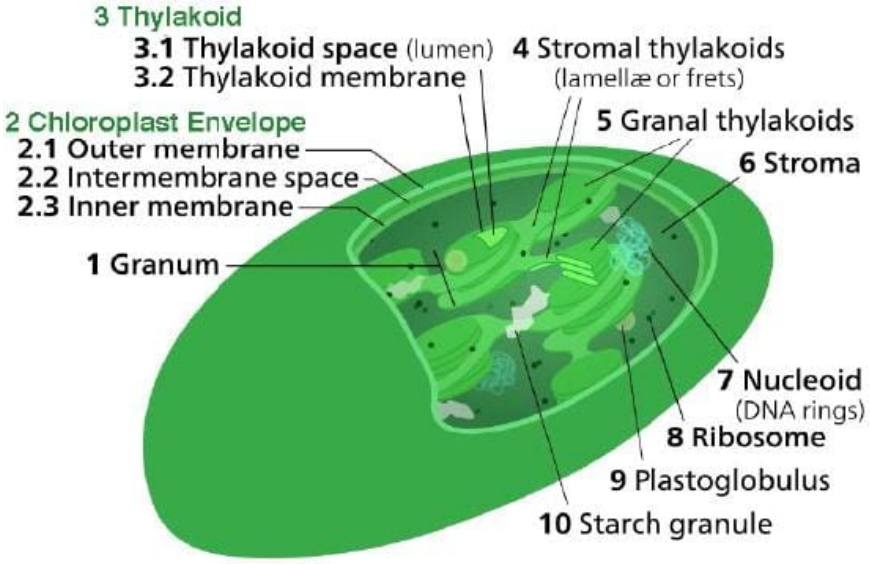


Figure 5.3.4: *Chloroplast anatomy. Wikipedia*

Photosynthesis

Photosynthesis is an energy capture process found in plants and other organisms to harvest light energy and convert it into chemical energy. This photochemical energy is stored ultimately in carbohydrates which are made using ATP (from the energy harvesting), carbon dioxide and water. In most cases, a byproduct of the process is oxygen, which is released from water in the capture process. Photosynthesis is responsible for most of the oxygen in the atmosphere and it supplies the organic materials

and most of the energy used by life on Earth.

Steps

The steps in the photosynthesis process varies slightly between organisms. In a broad overview, it always starts with energy capture from light by protein complexes, containing chlorophyll pigments, called reaction centers. Plants sequester these proteins in chloroplasts, but bacteria, which don't have organelles, embed them in their plasma membranes.

Energy from the light is used to strip electrons away from electron donors (usually water) and leave a byproduct (oxygen, if water was used). Electrons are donated to a carrier and ultimately are accepted by NADP+, to become NADPH. As electrons travel towards NADP+, they generate a proton gradient across the thylakoid membrane, which is used to drive synthesis of ATP. Thus NADPH, ATP, and oxygen are the products of the first phase of photosynthesis called the light reactions. Energy from ATP and electrons from NADPH are used to reduce CO2 and build sugars, which are the ultimate energy storage directly arising from photosynthesis.

Chloroplasts

Chloroplasts are found in almost all aboveground plant cells, but are primarily concentrated in leaves. The interior of a leaf, below the epidermis is made up of photosynthesis tissue called mesophyll, which can contain up to 800,000 chloroplasts per square millimeter.

The chloroplast's membrane has a phospholipid inner membrane, a phospholipid outer membrane, and a region between them called the intermembrane space (Figure 5.61). Within the inner chloroplast membrane is the stroma, in which the chloroplast DNA and the enzymes of the Calvin cycle are located. Also within the stroma are stacked, flattened disks known as thylakoids which are defined by their thylakoid membranes. The space within the thylakoid membranes are termed the thylakoid spaces or thylakoid lumen. The protein complexes containing the light-absorbing pigments, known as photosystems, are located on the thylakoid membrane. Besides chlorophylls, carotenes and xanthophylls are also present, allowing for absorption of light energy over a wider range. The same pigments are used by green algae and land plants.

Brown algae and diatoms add fucoxanthin (a xanthophyll) and red algae add phycoerythrin to the mix. In plants and algae, the pigments are held in a very organized fashion complexes called antenna proteins that help funnel energy, through resonance energy transfer, to the reaction center chlorophylls. A system so organized is called a light harvesting complex. The electron transport complexes of photosynthesis are also located on the thylakoid membranes.

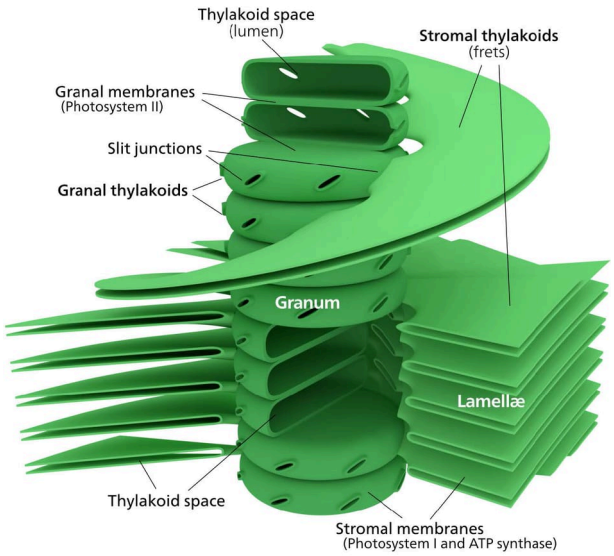


Figure 5.3.5: Side view of thylakoids. Wikipedia

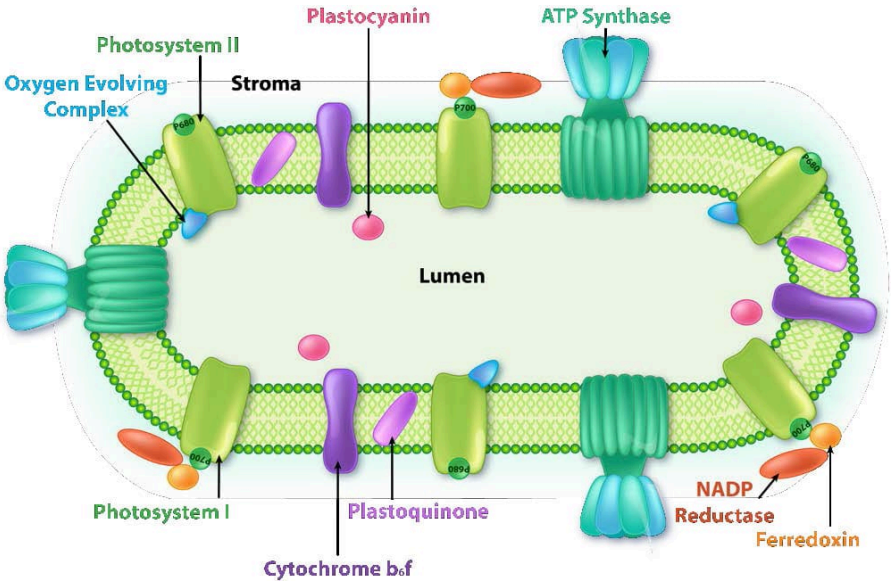


Figure 5.3.6***:*** *Complexes in the thylakoid membrane. Image by Aleia Kim*

Light reactions of photosynthesis

In chloroplasts, the light reactions of photosynthesis involving electron transfer occur in the thylakoid membranes (Figure 5.3.6). Separate biochemical reactions involving the assimilation of carbon dioxide to make glucose are referred to as the Calvin cycle, also sometimes referred to as the "dark reactions". This will be discussed elsewhere in the section on metabolism (HERE).

The chloroplasts are where the energy of light is captured, electrons are stripped from water, oxygen is liberated, electron transport occurs, NADPH is formed, and ATP is generated. The thylakoid membrane corresponds to the inner membrane of the mitochondrion for transport of electrons and proton pumping (Figure 5.3.4).

The thylakoid membrane does its magic using four major protein complexes. These include Photosystem II (PS II), Cytochrome b6f complex (Cb6f), Photosystem I (PS I), and ATP synthase. The roles of these complexes, respectively, are to capture light energy, create a proton gradient from electron movement, capture light energy (again), and use proton gradient energy from the overall process to synthesize ATP.

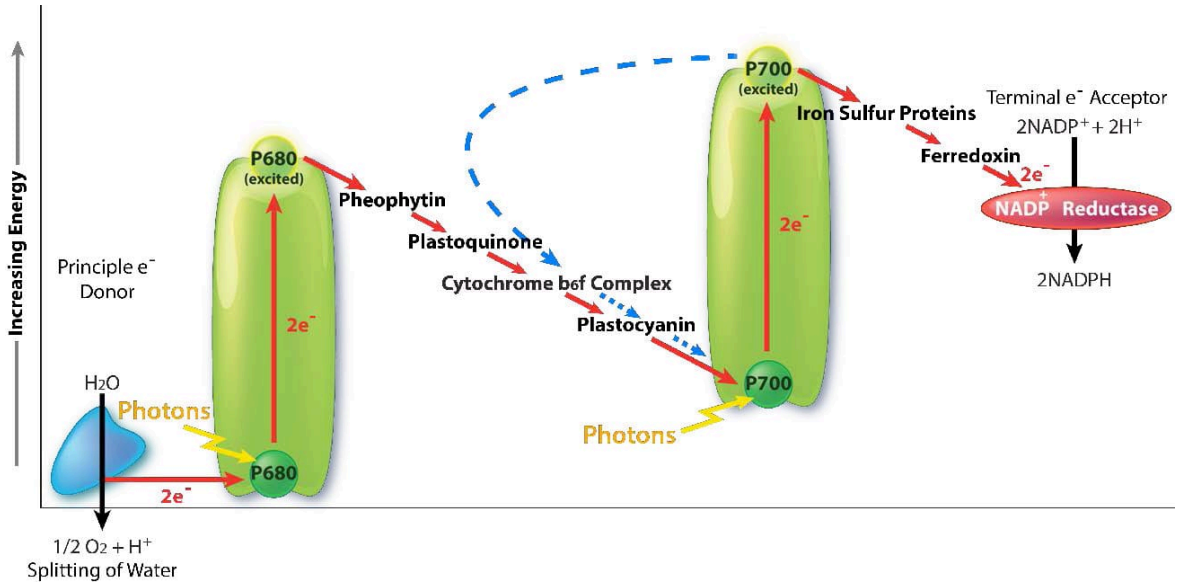


Figure 5.3.7: Movement of electrons through photosystems. Cyclic photophosphorylation shown by blue dashed line. Image by Aleia Kim and Pehr Jacobson

Light harvesting

Harvesting the energy of light begins in PS II with the absorption of a photon of light at a reaction center. PS II performs this duty best with light at a wavelength of 680 nm and it readily loses an electron to excitation when this occurs, leaving PS II with a positive charge. This electron must be replaced. The ultimate replacement source of electrons is water, but water must lose four electrons and PS II can only accept one at a time.

Manganese centers

An intermediate Oxygen Evolving Complex (OEC) contains four manganese centers that provide the immediate replacement electron that PSII requires. After four electrons have been donated by the OEC to PS II, the OEC extracts four electrons from two water molecules, liberating oxygen and dumping four protons into the thylakoid space, thus contributing to the proton gradient. The excited electron from PS II must be passed to another carrier very quickly, lest it decay back to its original state. It does this, giving its electron within picoseconds to pheophytin (Figure 5.3.8).

Pheophytin passes the electron on to protein-bound plastoquinones . The first is known as PQA. PQA hands the electron off to a second plastoquinone (PQB), which waits for a second electron and collects two protons to become PQH2, also known as plastoquinol (Figure 5.3.9). PQH2 passes these to the Cytochrome b6f complex (Cb6f) which uses passage of electrons through it to pump protons into the thylakoid space. ATP synthase makes ATP from the proton gradient created in this way. Cb6f drops the electron off at plastocyanin, which holds it until the next excitation process begins with absorption of another photon of light at 700 nm by PS I.

Absorption of light at PS I

With absorption of a photon of light by PS I, a process begins, that is similar to the process in PS II. PS I gains a positive charge as a result of the loss of an excited electron and pulls the electron in plastocyanin away from it. Meanwhile, the excited electron from PS I passes through an iron-sulfur protein, which gives the electron to ferredoxin (another iron sulfur protein). Ferredoxin then passes the electron off to the last protein in the system known as Ferredoxin:NADP+ oxidoreductase, which gives the electron and a proton to NADP+, creating NADPH.

Note that reduction of NADP+ to NADPH requires two electrons and one proton, so the four electrons and two protons from oxidation of water will result in production of two molecules of NADPH. At this point, the light cycle is complete - water has been oxidized, ATP has been created, and NADPH has been made. The electrons have made their way from water to NADPH via carriers in the thylakoid membrane and their movement has released sufficient energy to make ATP. Energy for the entire process came from four photons of light.

The two photosystems performing all of this magic are protein complexes that are similar in structure and means of operation. They absorb photons with high efficiency so that whenever a pigment in the photosynthetic reaction center absorbs a photon, an electron from the pigment is excited and transferred to another molecule almost instantaneously. This reaction is called photo-induced charge separation and it is a unique means of transforming light energy into chemical forms.

Cyclic photophosphorylation

Besides the path described above for movement of electrons through PS I, plants have an alternative route that electrons can take. Instead of electrons going through ferredoxin to form NADPH, they instead take a backwards path through the the proton-pumping b6f complex. This system, called cyclic photophosphorylation (Figure 5.3.8) which generates more ATP and no NADPH, is similar to a system found in green sulfur bacteria. The ability of plants to switch between non-cyclic and cyclic photosystems allows them to make the proper ratio of ATP and NADPH they need for assimilation of carbon in the dark phase of photosynthesis. This ratio turns out to be 3 ATPs to 2 NADPHs.

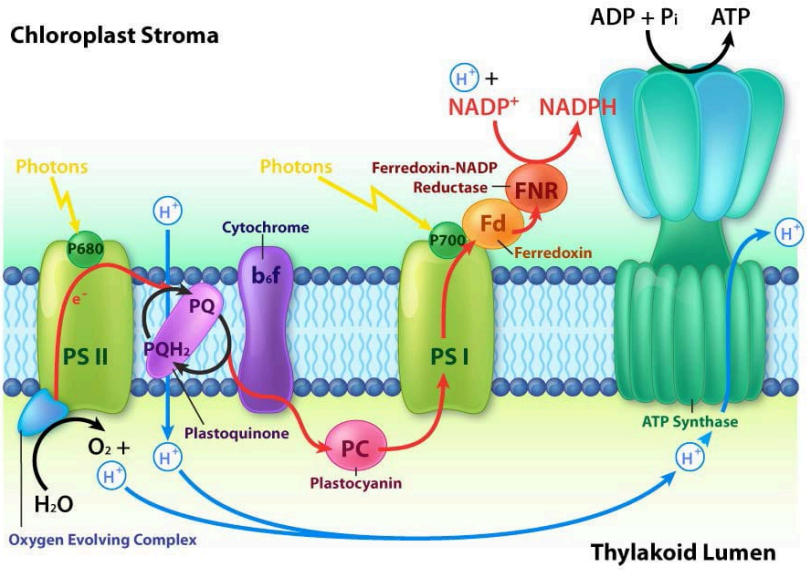


Figure 5.3.8 - Movement of electrons and protons through the thylakoid membrane. Image by Aleia Kim

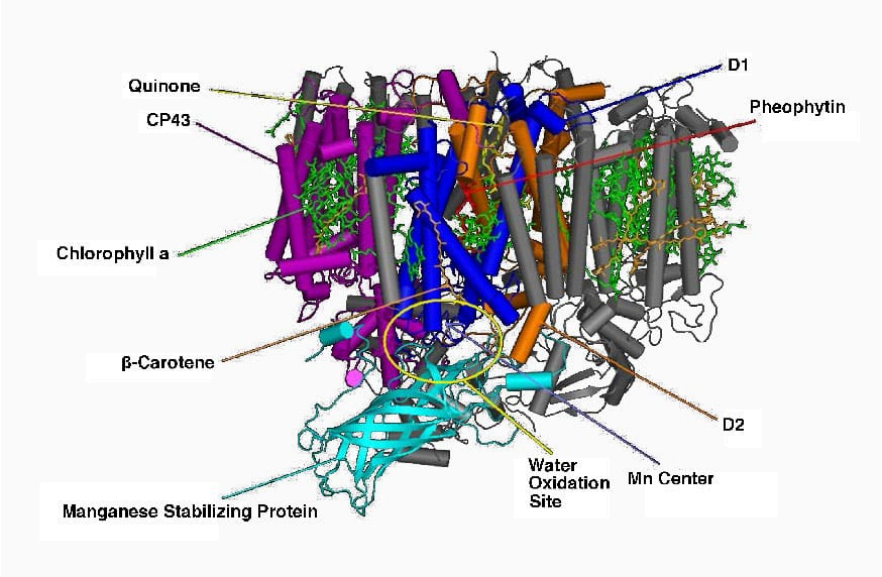


Figure 5.3.9 - Photosystem II of cyanobacteria. Wikipedia

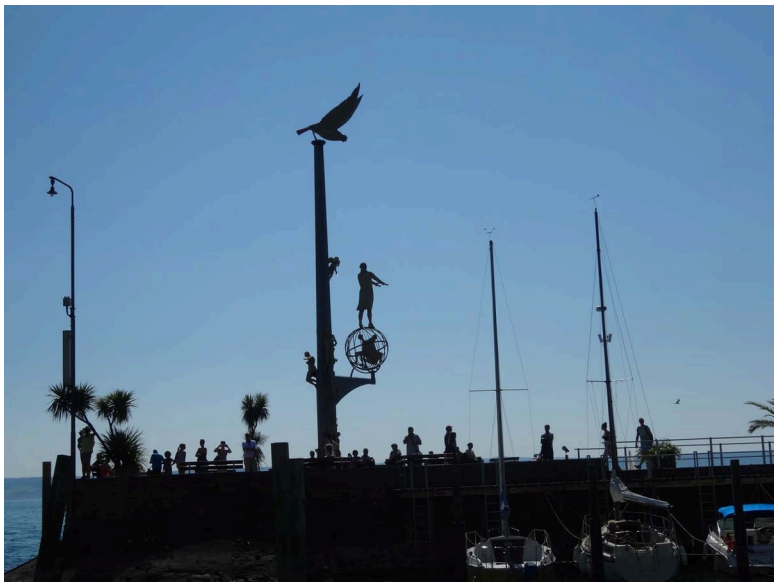
Photosynthetic energy

The output of the photophosphorylation part of photosynthesis (O2, NADPH, and ATP), of course, is not the end of the process of photosynthesis. For the growing plant, the NADPH and ATP are used to capture carbon dioxide from the atmosphere and convert it (ultimately) into glucose and other important carbon compounds. This, as noted previously, occurs in the Calvin Cycle (see HERE) in what is called the dark phase of the process. The oxygen liberated in the process is a necessary for respiration of all aerobic life forms on Earth. Indeed, it is believed that essentially all of the oxygen in the atmosphere today is the result the splitting of water in photosynthesis over the many eons that the process has existed.

This page titled 5.3: Energy - Photophosphorylation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

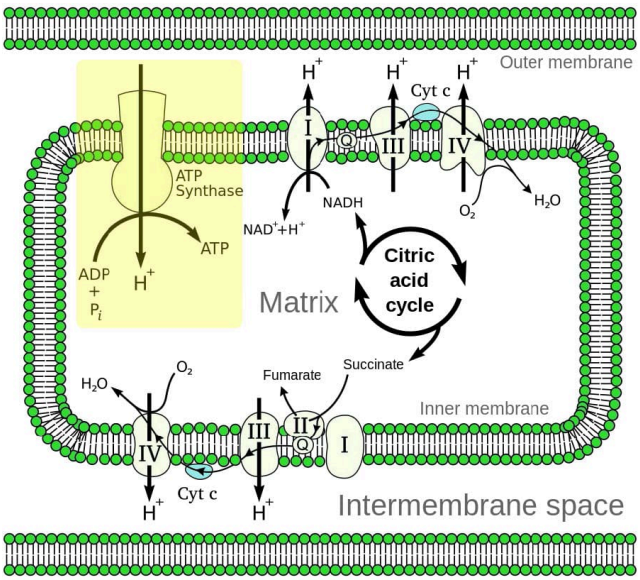
5.2: Electron Transport and Oxidative Phosphorylation

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

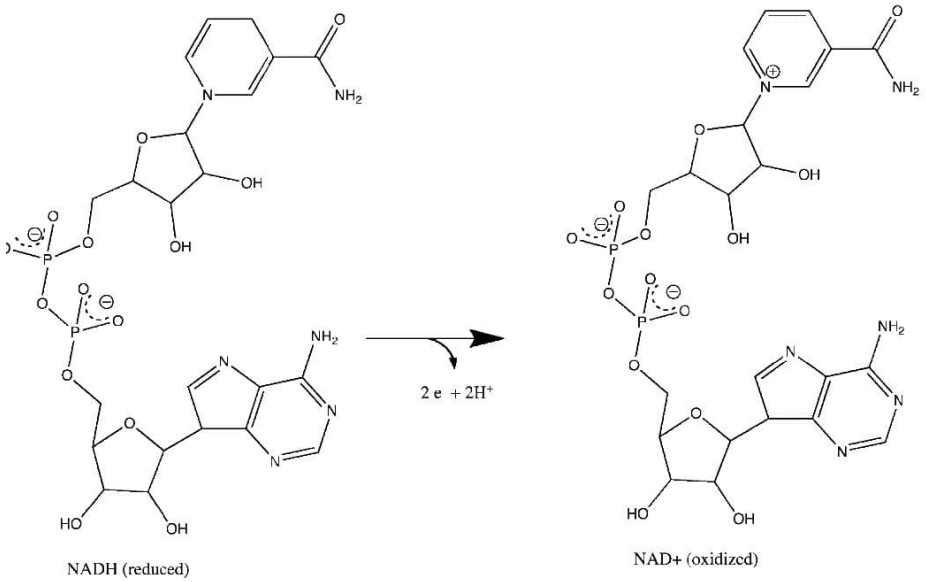
In oxidative phosphorylation, ATP synthesis is accomplished as a result of protons re-entering the mitochondrial matrix via the transmembrane ATP synthase complex, which combines ADP with inorganic phosphate to make ATP. Central to the proper functioning of mitochondria through this process is the presence 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

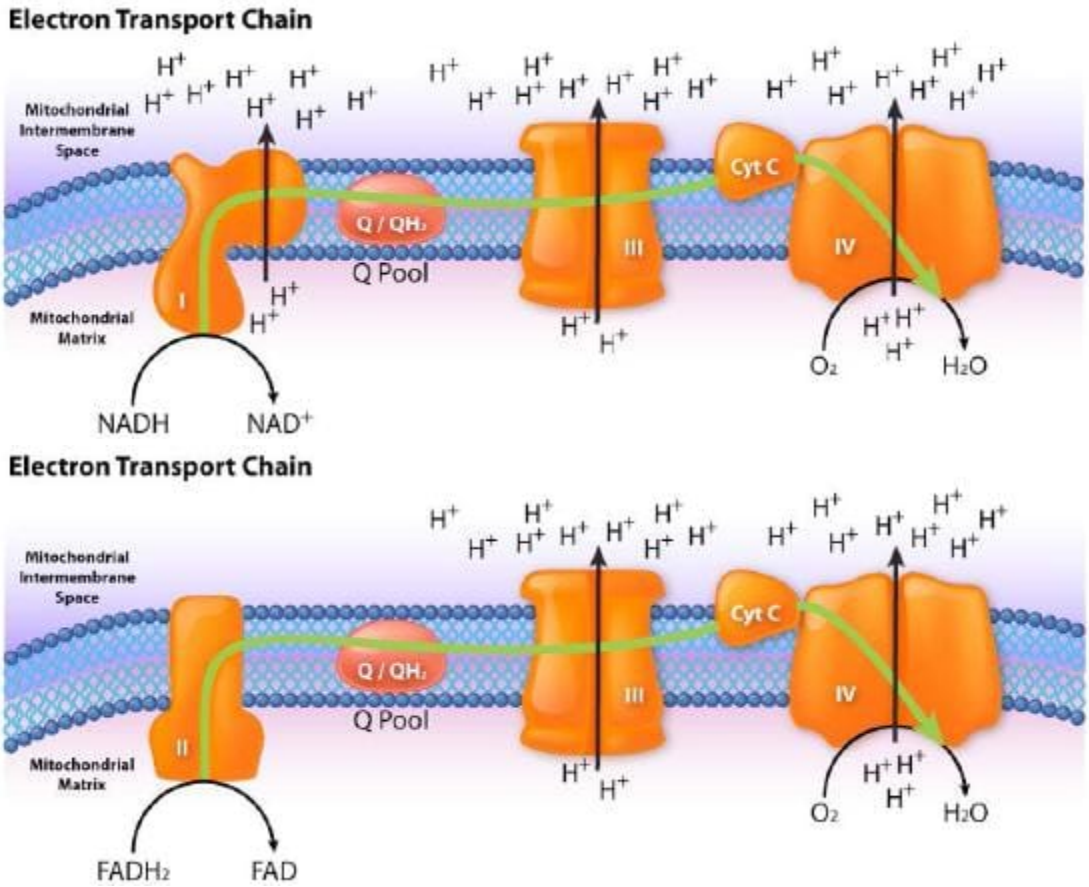
Mitochondria are called the power plants of the cell because most of a cell's ATP is produced there in the process of oxidative phosphorylation. The mechanism by which ATP is made in oxidative phosphorylation is one of the most interesting in all of biology.



NADH (reduced) NAD+ (oxidized)

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.



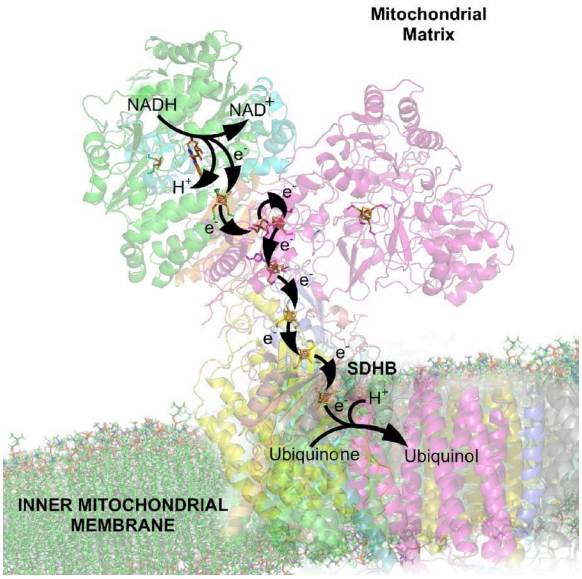
One can think of this occurring as a process where carriers “take” electrons away from complexes with lower reduction potential, much the way a bully takes lunch money from a smaller child. In this scheme, the biggest “bully” is oxygen in Complex IV. Electrons gained by a carrier cause it to be reduced, 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).

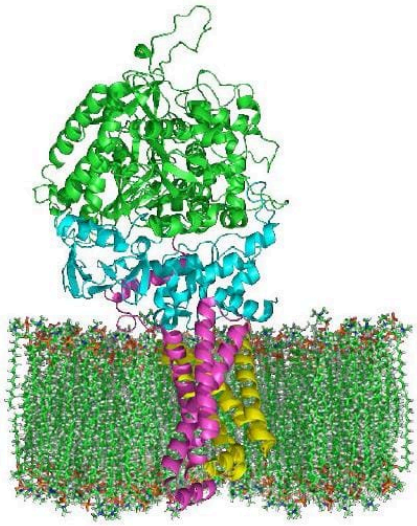
Complex III acts to ferry electrons from CoQ to cytochrome c. Cytochrome c takes one electron from Complex III and passes it to Complex IV (cytochrome oxidase). Complex IV is the final protein recipient of the electrons. It passes them to molecular oxygen (O2) to make two molecules of water. Making two 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.



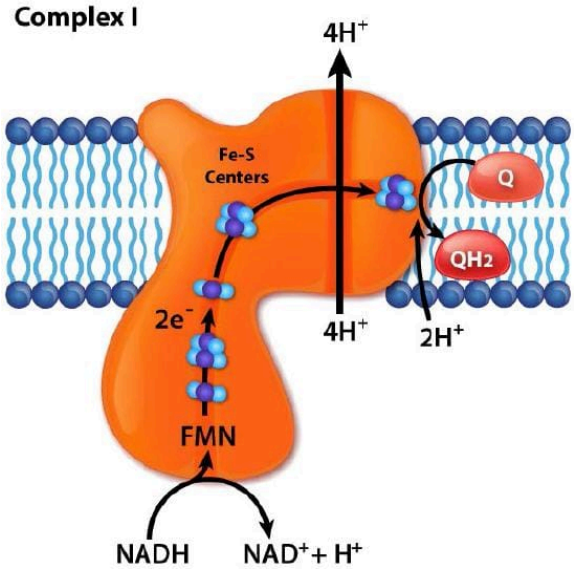
Having understood the overall picture of the synthesis of ATP linked to the movement of electrons through the ETS, we will take a closer look at the individual components of the ETS.

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.

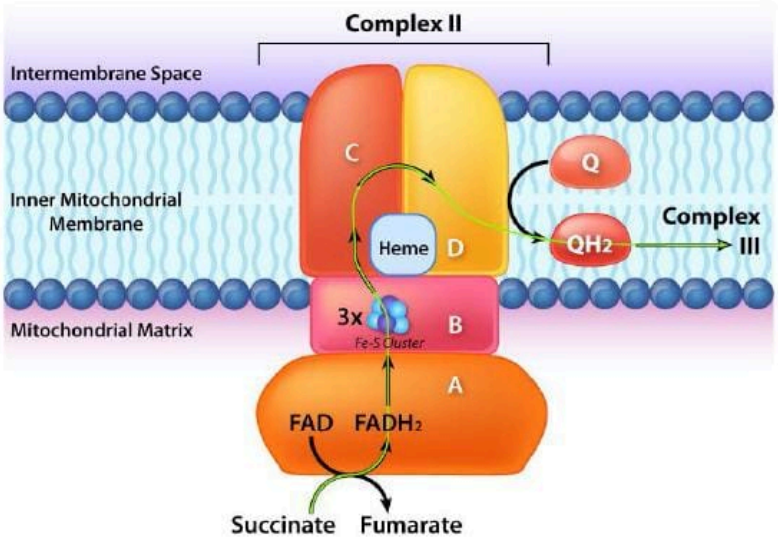
Complex I contains 44 individual polypeptide chains, numerous iron-sulfur centers, a molecule of flavin mononucleotide (FMN) and has an L shape with about 60 transmembrane domains. In the process of electron transport through it, four protons are pumped across the inner membrane into the intermembrane 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.

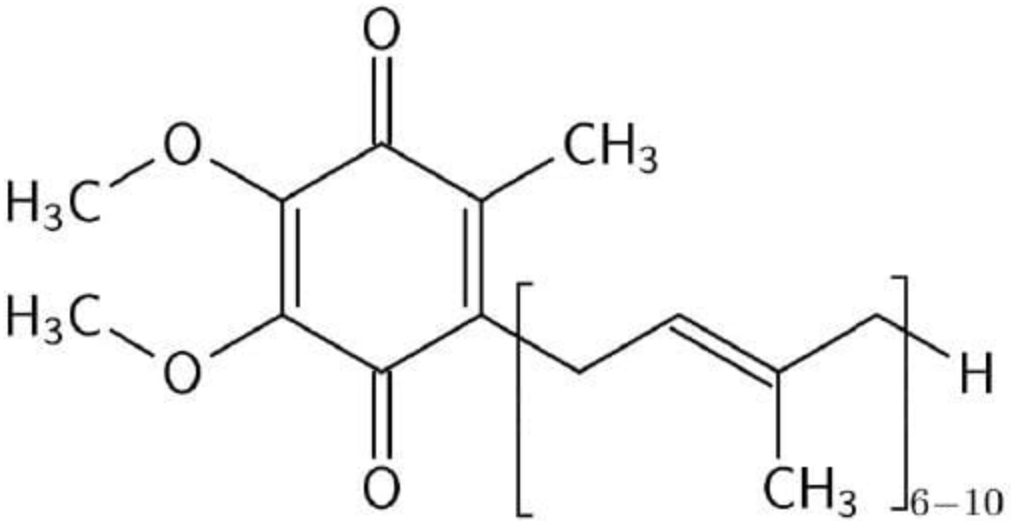


Coenzyme Q

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

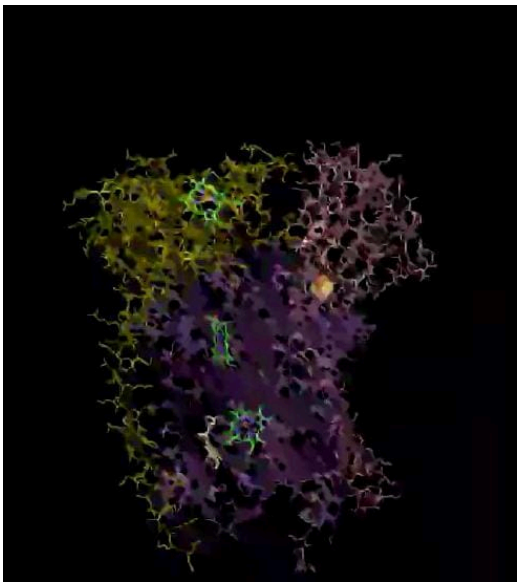
CoQ is useful because of its ability to carry and donate electrons and particularly because it can exist in forms with two extra electrons (fully reduced - ubiquinol), one extra electron (semi-reduced - ubisemiquinone), or no extra electrons (fully oxidized - ubiquinone). This ability allows CoQ to provide 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.

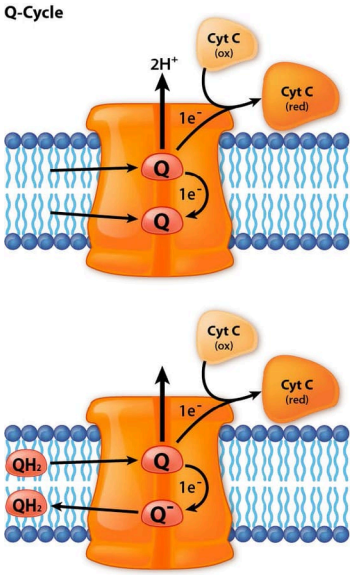
It accepts electrons from coenzyme Q in electron transport and passes them off to cytochrome c. In this cycle, known as the Q cycle, electrons arrive from CoQ in pairs, but get passed to cytochrome c individually. In the overall process, two protons are consumed from the matrix and four protons are pumped 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.

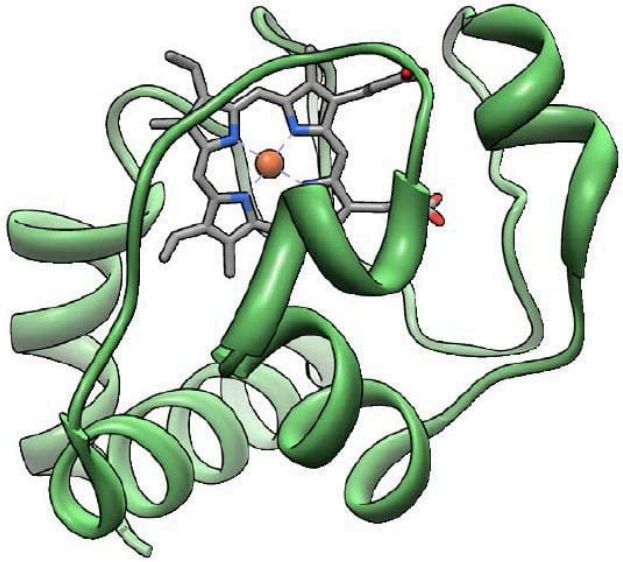
Step two

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.

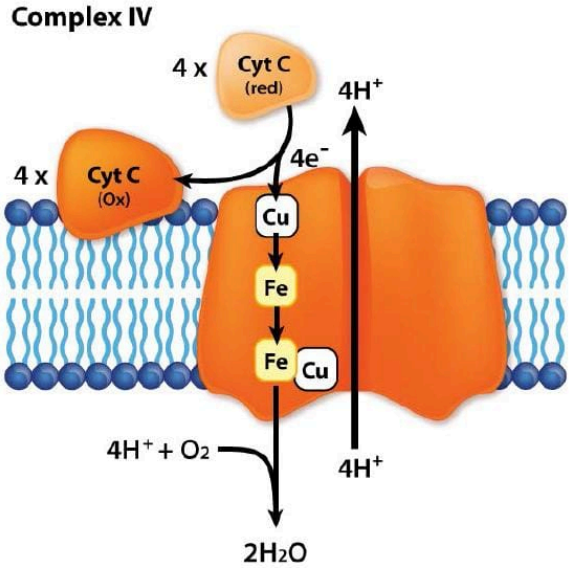
Cytochrome c

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.

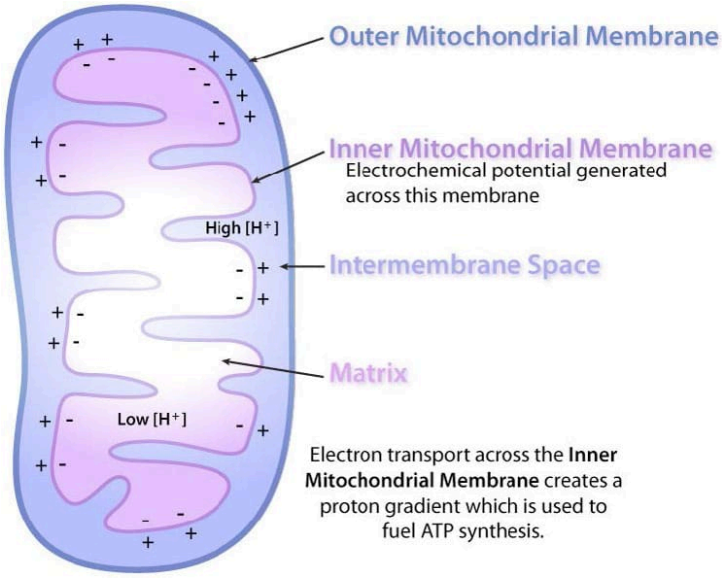


Respirasome

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).

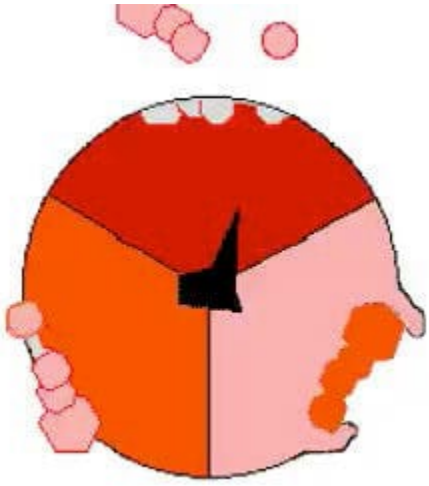
Another important consideration is that when ATP is made in oxidative phosphorylation, it is released into the mitochondrial matrix, but must be transported into the cytosol to meet the energy needs of the rest of the cell. This is accomplished by action of the adenine nucleotide translocase, an antiport 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



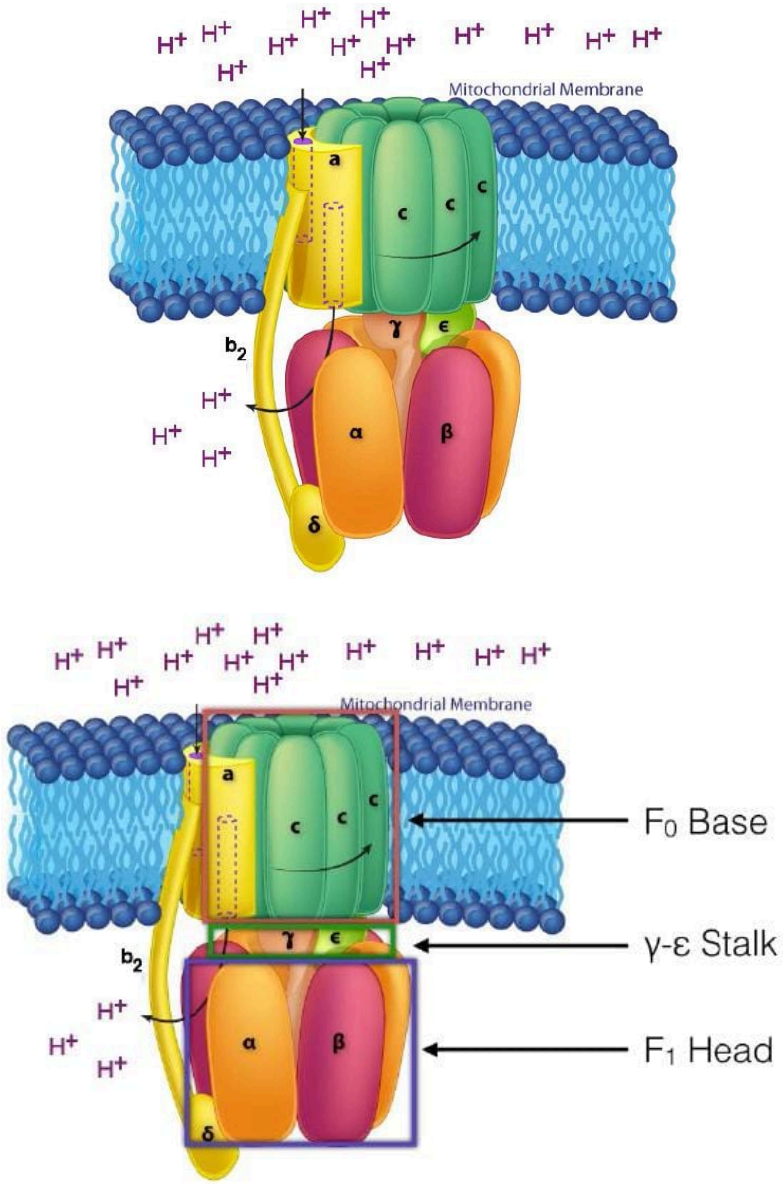


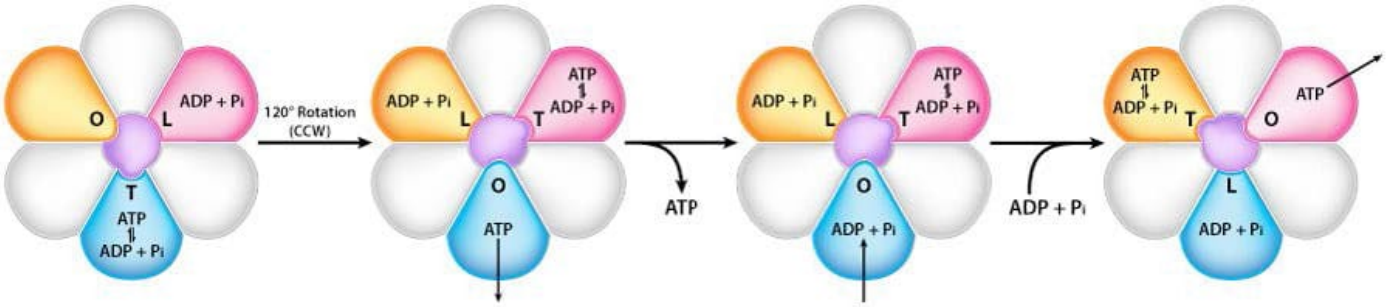
Figure 5.32 illustrate the multi-subunit nature of this membrane protein, which acts like a turbine at a hydroelectric dam. The movement of protons through the ATP Synthase c-ring causes it and the γ-ε stalk attached to it to turn. It is this action that is necessary for making ATP.

In ATP Synthase, the spinning components, or rotor, are the membrane portion (c ring) of the FO 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

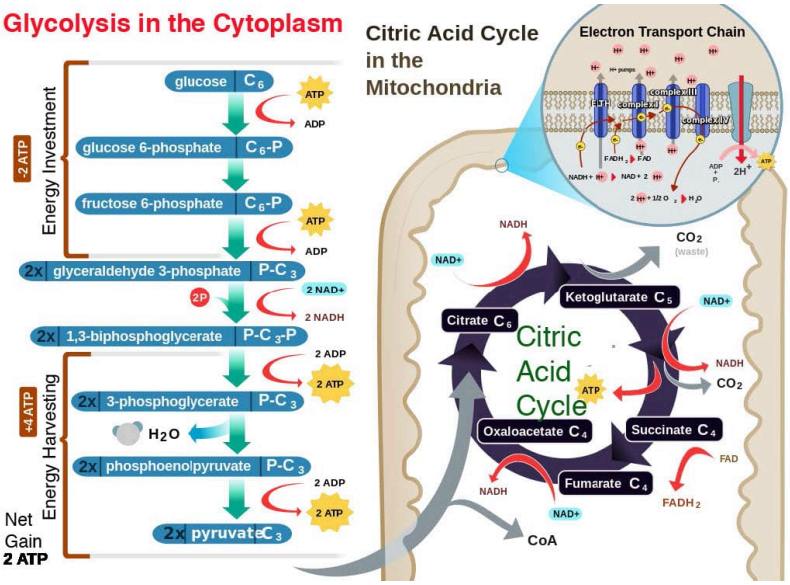
When running starts, muscular contraction, which uses energy, causes ATP to be converted to ADP. Increasing ADP in muscle cells favors oxidative phosphorylation to attempt to make up for the ATP being burned. ATP synthase begins working and protons begin to come back into the mitochondrial matrix. The 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.

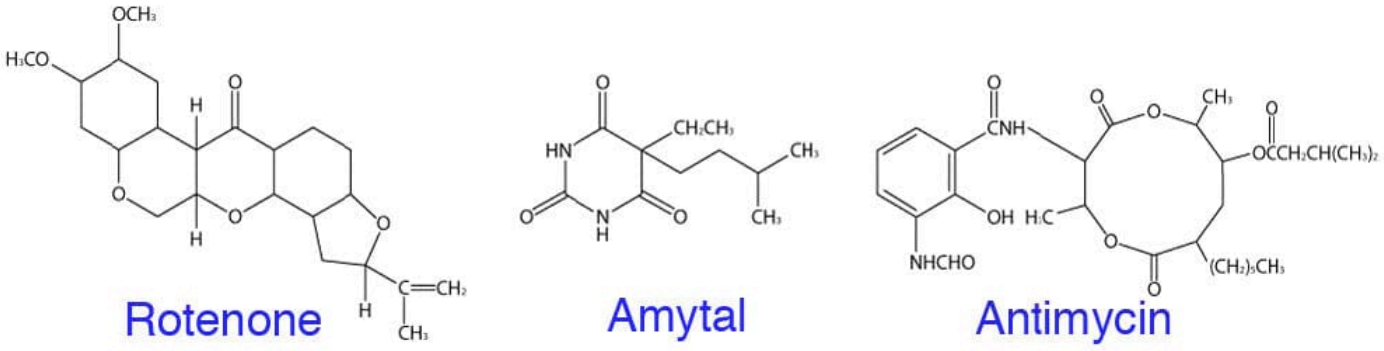
All of the chemicals described here are laboratory tools and should never be used by people. The first group for discussion are the inhibitors. In tightly coupled mitochondria, inhibiting either electron transport or oxidative phosphorylation has the effect of inhibiting the other one as well.

Electron transport inhibitors

Common inhibitors of electron transport include rotenone and amytal, which stop movement of electrons past Complex I, malonate, malate, and oxaloacetate, 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

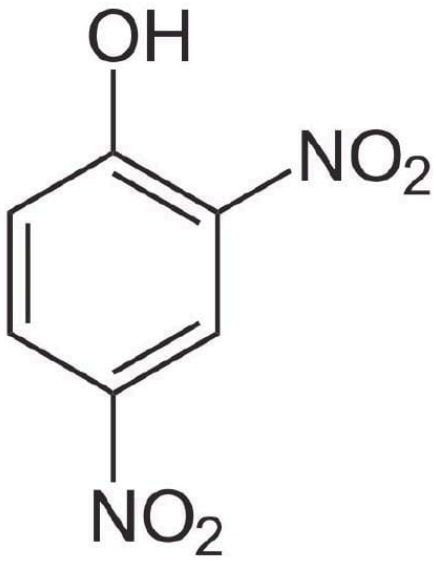


Figure 5.35). Treatment of mitochondria with 2,4 DNP makes the mitochondrial inner membrane “leaky” to protons. This has the effect of providing an alternate route for protons to reenter the matrix besides going through ATP synthase, and uncouples oxidative phosphorylation from electron transport.

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

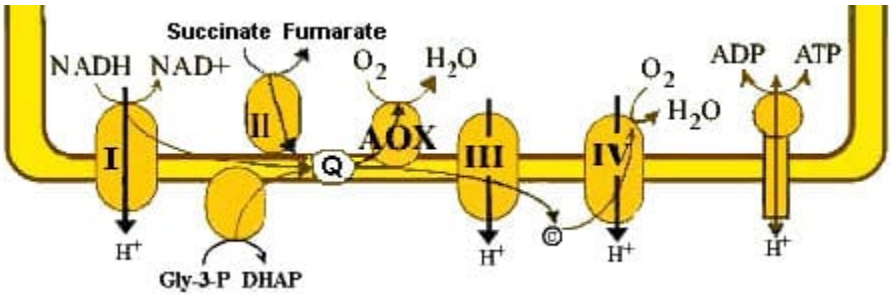
In uncoupling, whether through the action of an endogenous uncoupling protein or DNP, the energy that would have normally been captured in ATP is lost as heat. In the case of uncoupling by thermogenin, this serves the important purpose of keeping

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

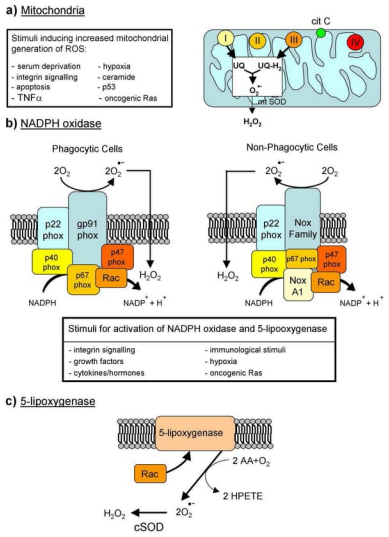


Figure 5.37) are oxygen containing molecules, such as peroxides, hydroxyl radical, superoxide, peroxynitrite, and others that are very chemically reactive. Though some ROS, such as peroxide and nitric oxide have important biological functions in signaling, increases in reactive oxygen species in times 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

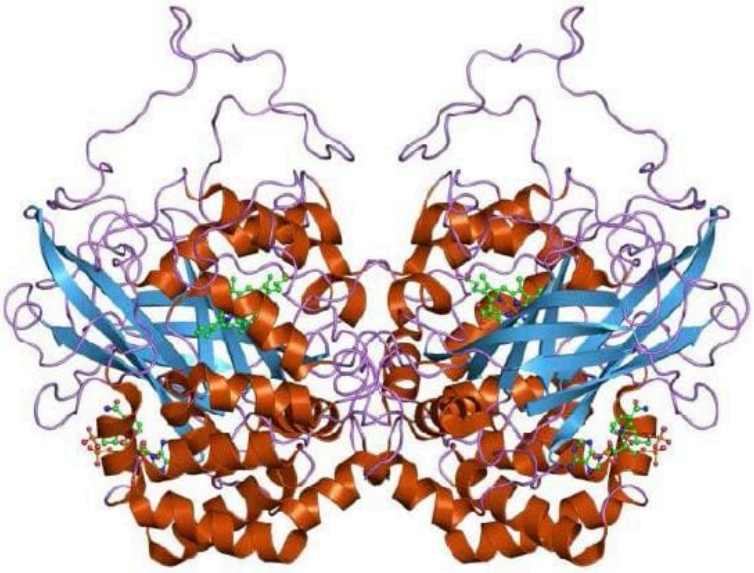


Figure 5.40) is an important enzyme for cells of all types that live in an oxygen environment. A first line of defense against reactive oxygen species, catalase catalyzes the breakdown of hydrogen peroxide into water and oxygen.

2 H2O2 <=> 2 H2O + O2

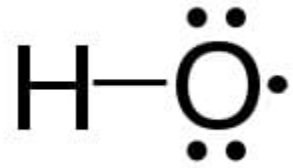
The enzyme, which employs four heme groups in its catalysis, works extremely rapidly, converting up to 40,000,000 molecules of hydrogen peroxide to water and oxygen per enzyme per second. It is abundantly found in peroxisomes.

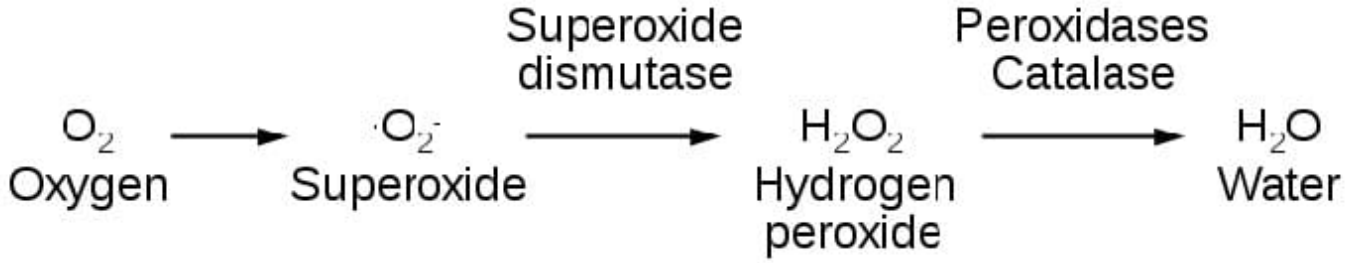
In addition to catalase's ability to break down hydrogen peroxide, the enzyme can also use hydrogen peroxide to oxidize a wide variety organic compounds, including phenols, formic acid, formaldehyde, acetaldehyde, and alcohols, but with much lower efficiency.

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



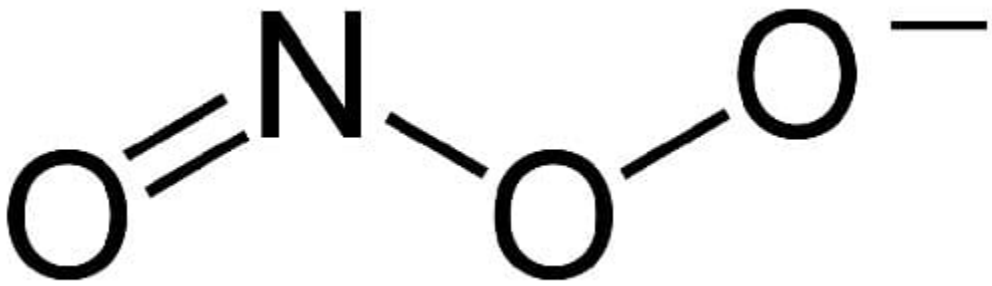




Another important enzyme for protection against reactive oxygen species is superoxide dismutase (SOD), which is found, like catalase, in virtually all organisms living in an oxygen environment. Superoxide dismutase, also like catalase, has a very high Kcat value and, in fact, has the highest Kcat/Km 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, as shown in Figure 5.45.



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 - Figure 5.45), mitochondria (SOD2 - Figure 5.46), and extracellular areas (SOD3 - Figure 5.47). 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 (Figure 5.47) 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.

Ferredoxin

Ferredoxins are iron-sulfur containing proteins performing electron transfer in a wide variety of biological systems and processes. They include roles in photosynthesis in chloroplasts. Ferredoxins are classified structurally by the iron-sulfur clustered centers they contain. Fe2S2 clusters (Figure 5.50) are found in chloroplast membranes and can donate electrons to glutamate synthase, nitrate reductase, and sulfite reductase and serve as electron carriers between reductase flavoproteins and bacterial dioxygenase systems. Adrenodoxin is a soluble human Fe2S2 ferredoxin (also called ferredoxin 1) serving as an electron carrier (to cytochrome P450) in mitochondrial monooxygenase systems. Fe4S4 ferredoxins are subdivided as low and high potential ferredoxins, with the latter ones functioning in anaerobic electron transport chains.

Ferritin

Ferritin is an intracellular iron-storage protein found in almost all living organisms, from bacteria to higher plants and animals. It is a globular protein complex with 24 subunits and is the primary intracellular iron-storage protein in eukaryotes and prokaryotes. Ferritin functions to keep iron in a soluble and non-toxic form. Its ability to safely store iron and release it in a controlled fashion allow it to act like the prime iron buffer and solubilizer in cells - keeping the concentration of free iron from going to high or falling too low. Ferritin is located in the cytoplasm in most tissues, but it is also found in the serum acting as an iron carrier. Ferritin that doesn't contain any iron is known as apoferritin.

Monoamine oxidases

Monoamine oxidases are enzymes that catalyze the oxidative deamination of monoamines, such as serotonin, epinephrine, and dopamine. Removal of the amine with oxygen results in the production of an aldehyde and ammonia. The enzymes are found inside and outside of the central nervous system.

There are two types of monoamine oxidase enzymes - MAO-A and MAO-B. MAO-A is particularly important for oxidizing monoamines consumed in the diet. Both MAO-A and MAO-B play important roles in inactivating monoaminergic neurotransmitters. Both enzymes act on dopamine, tyramine (Figure 5.50), and tryptamine. MAO-A is the primary enzyme for metabolizing melatonin, serotonin, norepinephrine, and epinephrine, while MAO-B is the primary enzyme for phenethylamine (Figure 5.51) and benzylamine. MAO-B levels have been reported to be considerably reduced with tobacco usage.

Actions of monoamine oxidases thus affects levels of neurotransmitters and consequently are thought to play roles in neurological and/or psychiatric disorders. Aberrant levels of MAOs have been linked to numerous psychological problems, including depression, attention deficit disorder (ADD), migraines, schizophrenia, and substance abuse. Medications targeting MAOs are sometimes used to treat depression as a last resort - due to potential side effects. Excess levels of catecholamines, such as epinephrine, norepinephrine, and dopamine, can result in dangerous hypertension events.

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.

Oxidative damage

DNA damage can occur in several ways. Oxidation can damage nucleotides and alter their base-pairing tendencies. Oxidation of guanine by reactive oxygen species, for example, can produce 8-oxo-guanine (Figures 5.52 and 5.53). This oxidized nucleobase commonly produced lesion in DNA arising from action of reactive oxygen species like superoxides. 8-oxoguanine is capable of forming a stable base pairing interaction within a DNA duplex with adenine, potentially giving rise to a mutation when DNA replication proceeds. 8-oxoguanine can be repaired if recognized in time by a DNA glycosylase, which acts to clip out the damaged base and it can then be replaced by the proper one. Polycyclic aromatic hydrocarbons from cigarette smoke, diesel exhaust, or overcooked meat can covalently bind to DNA and, if unrepaired, lead to mutation. Chemical damage to DNA can result in broken or cross-linked DNAs.

Diseases of DNA repair

The importance of DNA repair in the aging process is made clear by diseases affecting DNA repair that lead to premature aging. These include Werner syndrome, for whom the life expectancy is 47 years. It arises as a result of loss of two enzymes in base excision repair. People suffering from Cockayne syndrome have a life expectancy of 13 years due to mutations that alter transcription-coupled nucleotide excision repair, which is an important system for fixing oxidative damage.

Further, the life expectancies of 13 species of mammalian organisms correlates with the level of expression of the PARP DNA repair-inducing protein. Interestingly, people who lived past the age of 100 had a higher level of PARP than younger people in the population.

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.

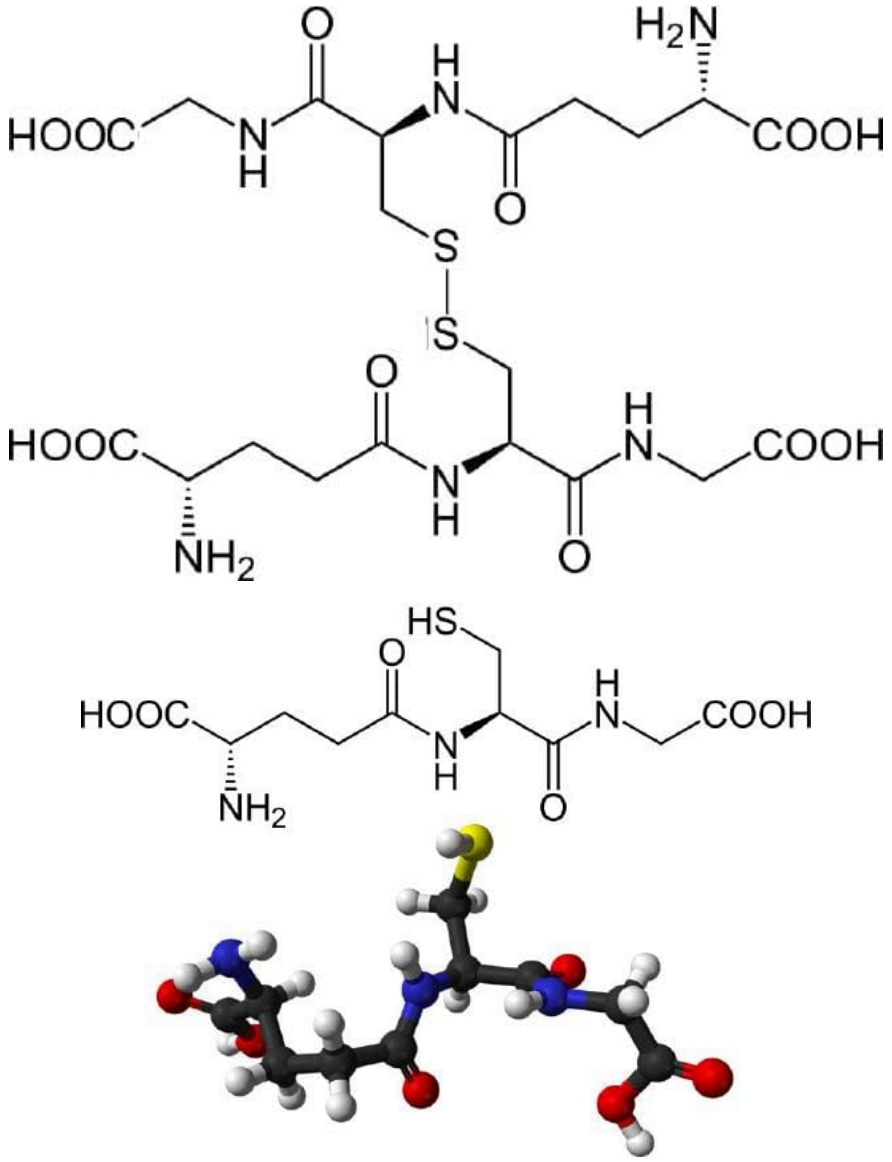
Biologically, cells have several lines of antioxidant defense. They include molecules, such as vitamins C, A, and E, glutathione, and enzymes that destroy ROS such as superoxide dismutase, catalase, and peroxidases.



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



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.

Non-ribosomal synthesis

Glutathione is not made by ribosomes. Rather, two enzymes catalyze its synthesis. The enzyme γ-glutamylcysteine synthetase catalyzes the joining of the glutamate to the cysteine and then glutathione synthetase catalyzes the peptide bond formation between the cysteine and the glycine. Each step requires energy from ATP.

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.

Regulation of the nitric oxide cycle

Resveratrol

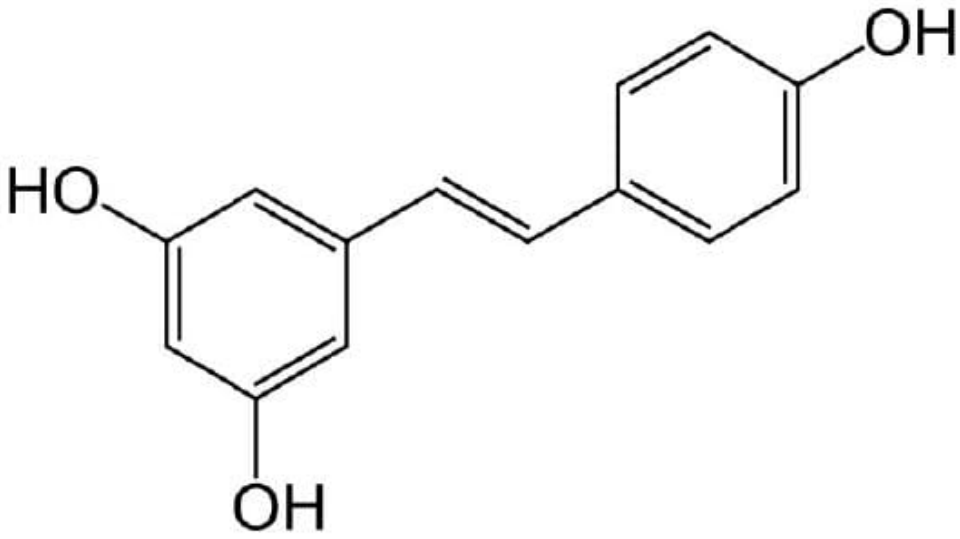


Figure 5.57) is a phenolic compound produced in the skin of plants such as grapes, raspberries, and blueberries, in response to injury or when they are being attacked by pathogens. Numerous health benefits are claimed for the compound, though evidence of such benefits is in short supply. Resveratrol 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

This page titled 5.2: Electron Transport and Oxidative Phosphorylation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6: METABOLISM

6.1: Metabolism - Sugars
6.2: Citric Acid Cycle & Related Pathways
6.3: Fats and Fatty Acids
6.4: Other Lipids
6.5: Amino Acids and the Urea Cycle
6.6: Nucleotides

This page titled 6: Metabolism is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6.1: Metabolism - Sugars

Glycolysis

Carbohydrates, whether synthesized by photosynthetic organisms, stored in cells as glycogen, or ingested by heterotrophs, must be broken down to obtain energy for the cell's activities as well as to synthesize other molecules required by the cell. Starch and glycogen, polymers of glucose, are the main energy storage forms of carbohydrates in plants and animals, respectively. To use these sources of energy, cells must first break down the polymers to yield glucose. The glucose is then taken up by cells through transporters in cell membranes. The metabolism of glucose, as well as other six carbon sugars (hexoses) begins with the catabolic pathway called glycolysis. In this pathway, sugars are oxidized and broken down into pyruvate molecules. The corresponding anabolic pathway by which glucose is synthesized is termed gluconeogenesis. Neither glycolysis nor gluconeogenesis is a major oxidative/reductive process, with one step in each one involving loss/gain of electrons, but the product of glycolysis, pyruvate, can be completely oxidized to carbon dioxide (Figure 6.2). Indeed, without production of pyruvate from glucose in glycolysis, a major energy source for the cell would not be available.

Glucose is the most abundant hexose in nature and is traditionally used to illustrate the reactions of glycolysis, but fructose (in the form of fructose-6- phosphate) is also readily metabolized, while galactose can easily be converted into glucose for catabolism in the pathway as well. The end metabolic products of glycolysis are two molecules of ATP, two molecules of NADH and two molecules of pyruvate (Figure 6.3), which, in turn, can be oxidized further in the citric acid cycle.

Entry points for glycolysis

Glucose and fructose are the sugar 'funnels' serving as entry points to the glycolytic pathway. Other sugars must be converted to either of these forms to be metabolized in glycolysis. Some pathways, including the Calvin

Figure 6.2 - Metabolic fates of glucose Image by Aleia Kim Your cells may have a mounting crisis Should they not go through glyco-lye-sis No glucose energy releases Until it's fractured into pieces

Figure 6.3 - Glycolysis and its Regulators Image by Ben Carson Cycle and the Pentose Phosphate Pathway (PPP) contain intermediates in common with glycolysis, so in that sense, almost any cellular sugar can be metabolized here.

Other pathways

Intermediates of glycolysis and gluconeogenesis that are common to other pathways include glucose-6-phosphate (PPP, glycogen metabolism), Fructose-6-phosphate (Calvin Cycle, PPP), Glyceraldehyde-3- phosphate (Calvin Cycle, PPP), dihydroxyacetone phosphate (PPP, glycerol metabolism, Calvin Cycle), 3- phosphoglycerate (Calvin Cycle, PPP), phosphoenolpyruvate (C4 plant metabolism, Calvin Cycle), and pyruvate (fermentation, acetyl-CoA genesis, amino acid metabolism). It is worth noting that glycerol from the breakdown of fat can readily be metabolized to dihydroxyacetone phosphate (DHAP) and thus enter the glycolysis pathway. It is the only part of a fat that is used in these pathways.

Reaction 1

Glucose gets a phosphate from ATP to make glucose-6-phosphate (G6P) in a reaction catalyzed by the enzyme hexokinase, a transferase enzyme.

Glucose + ATP ⇄ G6P + ADP + H+

Hexokinase is one of three regulated enzymes in glycolysis and is inhibited by one of the products of its action - G6P. Hexokinase has flexibility in its substrate binding and is able to phosphorylate a variety of hexoses, including fructose, mannose, and galactose.

Why phosphorylate glucose?

Phosphorylation of glucose serves two important purposes. First, the addition of a phosphate group to glucose effectively traps it in the cell, as G6P cannot diffuse across the lipid bilayer. Second, the reaction decreases the concentration of free glucose, favoring additional import of the molecule. G6P is a substrate for the pentose phosphate pathway and can also be converted to

glucose-1-phosphate (G1P) for use in glycogen synthesis and galactose metabolism (Figure 6.5).

It is worth noting that the liver has an enzyme like hexokinase called glucokinase, which Figure 6.4 - Reaction #1 - Phosphorylation of glucose - catalyzed by hexokinase has a much higher Km (lower affinity) for glucose. This is important, because the liver is a site of glucose synthesis (gluconeogenesis) where cellular concentrations of glucose can be relatively high. With a lower affinity glucose phosphorylating enzyme, glucose is not converted to G6P unless glucose concentrations get high, so the liver is able to release the glucose it makes into the bloodstream for the rest of the body to use.

Reaction 2

Next, G6P is converted to fructose-6-phosphate (F6P), in a reaction catalyzed by the enzyme

$$\mathrm{G_6P} \rightleftarrows \mathrm{F_6P} \tag{6.1.1}$$

The reaction has a low ΔG°' , so it is readily favorable in either direction with Figure 6.6 - Mechanism of conversion of G6P to F6P in reaction #2 Figure 6.5 - The centrality of glucose-6-phosphate in metabolism Image by Aleia Kim only slight changes in concentration of reactants.

Reaction 3

$$ceF6P + ATP \rightleftarrows F1,6BP + ADP + H+ \tag{6.1.2}$$

The second input of energy occurs when F6P gets another phosphate from ATP in a reaction catalyzed by the enzyme phosphofructokinase-1 (PFK-1 - another transferase) to make fructose-1,6- bisphosphate (F1,6BP). PFK-1 is a very important enzyme regulating glycolysis, with several allosteric activators and inhibitors (see HERE).

Like the hexokinase reaction the energy from ATP is needed to make the reaction energetically favorable. PFK-1 is the most important regulatory enzyme in the pathway and this reaction is the ratelimiting step. It is also one of three essentially irreversible reactions in glycolysis.

A variant enzyme found in plants and some bacterial uses pyrophosphate rather than ATP as the energy source and due to the lower energy input from hydrolysis of the pyrophosphate, that reaction is reversible.

Reaction 4

$$\mathrm{F_1,6\,BP} \rightleftarrows \mathrm{D{-}GLYAL_3P + DHAP} \tag{6.1.3}$$

With the glycolysis pump thus primed, the pathway proceeds to split the F1,6BP into two 3-carbon intermediates. This reaction catalyzed by the lyase known as aldolase is energetically a “hump” to overcome in the glycolysis direction (ΔG°' = +24 kJ/mol Figure 6.7 - Reaction #3 - Conversion of F6P to F1,6BP by PFK Wikipedia Figure 6.8 - Reaction #4 - Breakdown of F1,6BP into GLYAL3P (left) and DHAP (right) by aldolase °K) so to get over the energy hump, cells must increase the concentration the reactant (F1,6BP) and decrease the concentration of the products, which are D-glyceraldehyde- 3-phosphate (D-GLYAL3P) and dihydroxyacetone phosphate (DHAP).

A novel scheme facilitates decreasing concentration of the products (see below). Aldolases cut the ketose ring by two different mechanisms and these enzymes are grouped as Class I (in animals and plants) and Class II (in fungi and bacteria).

Reaction 5

$$\mathrm{DHAP} \rightleftarrows \mathrm{D{-}GLYAL_3P} \tag{6.1.4}$$

In the next step, DHAP is converted to DGLYAL3P in a reaction catalyzed by the enzyme triosephosphate isomerase. At this point, the six carbon glucose molecule has been broken down to two units of three carbons each - D-GLYAL3P. From this point forward each reaction of glycolysis contains two of each molecule. Reaction #5 is fairly readily reversible in cells.

The enzyme is of note because it is one example of a “perfect enzyme.” Enzymes in this category have very high ratios of Kcat/Km that approach a theoretical maximum limited only by the diffusion of substrate into the active site of the enzyme. The apparent reason for the enzyme evolving in this way is that the mechanism of the reaction produces an unstable, toxic intermediate (Figure 6.9). With the reaction proceeding as rapidly as it does, there is less chance of the intermediate escaping and causing damage in the cell.

Reaction 6

$$\mathrm{D{-}GLYAL_3P + NAD^+ + Pi\ D\text{-}1{,}3\,BPG + NADH + H^+} \tag{6.1.5}$$

Figure 6.9 - Reaction #5 - Triose phosphate isomerase with unstable, toxic intermediate (methyl glyoxal) Image by Ben Carson

In this reaction, D-GLYAL3P is oxidized in the only oxidation step of glycolysis catalyzed by the enzyme glyceraldehyde-3-phosphate dehydrogenase, an oxidoreductase. The aldehyde in this reaction is oxidized, then linked to a phosphate to make an ester - D-1,3-bisphospho-glycerate (D- 1,3BPG). Electrons from the oxidation are donated to NAD+, creating NADH.

NAD+ is a critical constituent in this reaction and is the reason that cells need a fermentation option at the end of the pathway (see below).

Note here that ATP energy was not required to put the phosphate onto the oxidized D-GLYAL3P. The reason for this is because the energy provided by the oxidation reaction is sufficient for adding the phosphate.

Reaction 7

$$\mathrm{D\text{-}1{,}3\,BPG + ADP \rightleftarrows\ 3\,PG + ATP} \tag{6.1.6}$$

The two phosphates in the tiny 1,3BPG molecule repel each other and give the molecule high potential energy. This energy is utilized by the enzyme phosphoglycerate kinase (another transferase) to phosphorylate ADP and make ATP, as well as the product, 3-phosphoglycerate (3-PG). This is an example of a substrate-level phosphorylation. Such mechanisms for making ATP require an intermediate with a high enough energy to phosphorylate ADP to make ATP. Figure 6.10 - Reaction #6 - Oxidation of GLYAL3P, catalyzed by glyceraldehyde-3-phosphate dehydrogenase Figure 6.11 - Reaction #7 - Substrate-level Phosphorylation by 1,3-BPG

Though there are a few substrate level phosphorylations in cells (including another one at the end of glycolysis), the vast major of ATP is made by oxidative phosphorylation in the mitochondria (in animals). In addition to oxidative phosphorylation, plants also make ATP by photophosphorylation in their chloroplasts. Since there are two 1,3 BPGs produced for every glucose, the two ATPs produced in this reaction replenish the two ATPs used to start the cycle and the net ATP count at this point of the pathway is zero.

Reaction 8

$$\mathrm{3\text{-}PG \rightleftarrows\ 2\text{-}PG} \tag{6.1.7}$$

Conversion of the 3-PG intermediate to 2-PG (2- phosphoglycerate) occurs by an important mechanism. An intermediate in this readily reversible reaction (catalyzed by phosphoglycerate mutase - a mutase enzyme) is 2,3-BPG. This intermediate, which is stable, is released with low frequency by the enzyme instead of being con- Figure 6.13 - Two routes to formation of 2,3-BPG Figure 6.14 - 2,3- Bisphosphoglycerate (2,3-BPG) Figure 6.12 - Reaction #8 - Conversion of 3-PG to 2-PG verted to 2-PG. 2,3BPG is important because it binds to hemoglobin and stimulates release of oxygen. The molecule can also be made from 1,3-BPG as a product of a reaction catalyzed by bisphophglycerate mutase (Figure 6.13). Cells which are metabolizing glucose rapidly release more 2,3-BPG and, as a result, get more oxygen, supporting their needs. Notably, cells which are metabolizing rapidly are using oxygen more rapidly and are more likely to be deficient in it.

Reaction 9

$$\mathrm{2\text{-}PG \rightleftarrows\ PEP + H_2O} \tag{6.1.8}$$

2-PG is converted by enolase (a lyase) to phosphoenolpyruvate (PEP) by removal of water, creating a very high energy intermediate. The reaction is readily reversible, but with PEP, the cell has one of its highest energy molecules and that is important for the next reaction.

Reaction 10

$$\mathrm{PEP + ADP + H^+ \rightleftarrows\ PYR + ATP} \tag{6.1.9}$$

Conversion of PEP to pyruvate by pyruvate kinase is the second substrate level phosphorylation of glycolysis, creating ATP. This reaction is what some refer to as the “Big Bang” of glycolysis because there is almost enough energy in PEP to stimulate production of a second ATP ($\Delta G^{\circ\prime}$ = 31.6 kJ/ mol), but it is not used. Consequently, this energy is lost as heat. If you wonder why you get hot when you exercise, the heat produced in the breakdown of glucose is a prime contributor and the pyruvate kinase reaction is a major source. Figure 6.16 - Reaction #10 - The big bang - PEP phosphorylates ADP with a lot of energy to spare

Wikipedia Figure 6.15 - Reaction #9 - Enolase-catalyzed removal of water Wikipedia

Pyruvate kinase is the third and last enzyme of glycolysis that is regulated (see below). The primary reason this is the case is to be able to prevent this reaction from occurring when cells are making PEP while going through gluconeogenesis (see more HERE).

Catabolism of other sugars

Though glycolysis is a pathway focused on the metabolism of glucose and fructose, the fact that other sugars can be readily metabolized into glucose means that glycolysis can be used for extracting energy from them as well. Galactose is a good example. It is commonly produced in the produced in the body as a result of hydrolysis of lactose, catalyzed by the enzyme known as lactase (Figure 6.17). Deficiency of lactase is the cause of lactose intolerance.

Galactose begins preparation for entry into glycolysis by being converted to galactose-1- phosphate (catalyzed by galactokinase - Figure 6.18). Galactose-1-phosphate swaps with glucose-1-phosphate from UDP-glucose to make UDP-galactose (Figure 6.19). An epimerase converts UDPgalactose back to UDP-glucose and the cycle is complete. Each turn of the cycle thus takes in one galactose-1-phosphate and releases one glucose-1-phosphate.

Deficiency of galactose conversion enzymes results in accumulation of galactose (from breakdown of lactose). Excess galactose is converted to galactitol, a sugar alcohol. Galactitol in the human eye lens causes it to absorb water and this may be a factor in formation of cataracts.Figure 6.17 - Breakdown of lactose to glucose and galactose by lactase Image by Pehr Jacobson Figure 6.18 - Galactokinase Reaction Image by Penelope Irving Free fructose can also enter glycolysis by two mechanisms. First, it can be phosphorylated to fructose-6-phosphate by hexokinase. A more interesting alternate entry point is that shown in Figure 6.20. Phosphorylation of fructose by fructokinase produces fructose-1-phosphate and cleavage of that by fructose-1- phosphate aldolase yields DHAP and glyceraldehyde.

Phosphorylation of glyceraldehyde by triose kinase yields GLYAL3P. This alternative entry means for fructose may have important implications because DHAP and GLYAL3P are introduced into the glycolysis pathway while bypassing PFK-1 regulation. Some have proposed this may be important when considering metabolism of high fructose corn syrup, since it forces production of pyruvate, a precursor of acetyl-CoA, which is itself a precursor of fatty acids when ATP levels are high.

Mannose metabolism

Mannose can also be metabolized in glycolysis. In this case, it enters via fructose by the following two-step process - 1) phosphoryla- Figure 6.19 - Conversion of galactose-1-phosphate into glucose-6-phosphate Image by Aleia Kim tion by hexokinase to make mannose-6- phosphate followed by its conversion to fructose-6-phosphate, catalyzed by phosphomannoisomerase (Figure 6.21).

Glycerol metabolism

Glycerol is an important molecule for the synthesis of fats, glycerophospholipids, and other membrane lipids. Most commonly it is made into glycerol-3- phosphate (Figure 6.22) and the glycolysis/gluconeogenesis pathways are important both for producing the compound and for metabolizing it. The relevant intermediate in these pathways both for producing and for using glycerol-3-phosphate is DHAP. The enzyme glycerol-3-phosphate dehydrogenase reversibly converts glycerol-3- phosphate into DHAP (Figure 6.22).

This reaction, which is an oxidation, transfers electrons to NAD+ to produce NADH. In the reverse reaction, production of glycerol-3- phosphate from DHAP, of course, requires electrons from NADH for the reduction. Both glycolysis and gluconeogenesis are sources DHAP, meaning when the cell needs glycerol- 3-phosphate that it can use sugars (glucose, fructose, mannose, or galactose) as sources in glycolysis. For gluconeogenesis, sources include pyruvate, alanine and Figure 6.20 - Entry of fructose into glycolysis, bypassing PFK-1 Image by Penelope Irving Figure 6.21 - Entry of other sugars into glycolysis Image by Penelope Irving lactate (both can easily be made into pyruvate), oxaloacetate, aspartic acid (which can be made into oxaloacetate by transamination), and others. All of the intermediates of the citric acid cycle (and glyoxylate cycle) can be converted ultimately to oxaloacetate, which is a gluconeogenesis intermediate, as well.

It is worth noting that animals are unable to use fatty acids as materials for gluconeogenesis in net amounts, but they can, in fact, use glycerol in both glycolysis and gluconeogenesis. It is the only part of the fat molecule that can be so used.

Pyruvate metabolism

As noted, pyruvate produced in glycolysis can be oxidized to acetyl-CoA, which is itself oxidized in the citric acid cycle to carbon

dioxide. That is not the only metabolic fate of pyruvate, though (Figure 6.23).

Pyruvate is a “starting” point for gluconeogenesis, being converted to oxaloacetate in the mitochondrion in the first step. Pyruvate in animals can also be reduced to lactate by adding electrons from NADH (Figure 6.24). This reaction produces NAD+ and is critical for generating the latter molecule to keep the glyceraldehyde-3-phosphate dehydrogenase reaction of glycolysis (reaction #6) going under conditions when there is no oxygen.

This is because oxygen is necessary for the electron transport system (ETS) to operate and it performs the important function of converting NADH back to NAD+. When the ETS is running, NADH donates electrons to Complex I and is oxidized to NAD+ in the process, generating the intermediate needed for oxidizing GLYAL-3P. In the absence of oxygen, however, NADH cannot be converted to Figure 6.22 - Reactions in glycerol metabolism Image by Penelope Irving NAD+ by the ETS, so an alternative means of making NAD+ is necessary for keeping glycolysis running under low oxygen conditions (fermentation).

Bacteria and yeast generate NAD+ under oxygen deprived conditions by doing fermentation in a different way (Figure 6.25). They use NADH-requiring reactions that regenerate NAD+ while producing ethanol from pyruvate instead of making lactate. Thus, fermentation of pyruvate is essential to keep glycolysis operating when oxygen is limiting. It is also for these reasons that brewing of beer (using yeast) involves depletion of oxygen and muscles low in oxygen produce lactic acid (animals).

Pyruvate is a precursor of alanine which can be easily synthesized by transfer of a nitrogen from an amine donor, such as glutamic acid. Pyruvate can also be converted into oxaloacetate by carboxylation in the process of gluconeogenesis (see below).

The enzymes involved in pyruvate metabolism include pyruvate dehydrogenase (makes acetyl-CoA), lactate dehydrogenase (makes lactate), transaminases (make alanine), pyruvate carboxylase (makes ox- Figure 6.23 - Pyruvate metabolism. When oxygen is absent, pyruvate is converted to lactate (animals) or ethanol (bacteria and yeast). When oxygen is present, pyruvate is converted to acetyl-CoA. Not shown - Pyruvate transamination to alanine or carboxylation to form oxaloacetate. aloacetate), and pyruvate decarboxylase (a part of pyruvate dehydrogenase that makes acetaldehyde in bacteria and yeast).

Catalytic action and regulation of the pyruvate dehydrogenase complex is discussed in the section on the citric acid cycle (HERE).

Gluconeogenesis

The anabolic counterpart to glycolysis is gluconeogenesis (Figure 6.26), which occurs mostly in the cells of the liver and kidney and virtually no other cells in the body. In seven of the eleven reactions of gluconeogenesis (starting from pyruvate), the same enzymes are used as in glycolysis, but the reaction directions are reversed. Notably, the ΔG values of these reactions in the cell are typically near zero, meaning their direction can be readily controlled by changing substrate and product concentrations by small amounts.

The three regulated enzymes of glycolysis all catalyze reactions whose cellular ΔG values are not close to zero, making manipulation of reaction direction for their reac- Figure 6.24 - Formation of lactate in animal fermentation produces NAD+ for G3PDH Image by Ben Carson Figure 6.25 - Formation of ethanol in microbial fermentation produces NAD+ for G3PDH Image by Ben Carson tions non-trivial. Consequently, cells employ “work-around” reactions catalyzed by four different enzymes to favor gluconeogenesis, when appropriate.

Bypassing pyruvate kinase

Two of the enzymes (pyruvate carboxylase and PEP carboxykinase - PEPCK) catalyze reactions that bypass pyruvate kinase. F1,6BPase bypasses PFK-1 and G6Pase bypasses hexokinase. Notably, pyruvate carboxylase and G6Pase are found in the mitochondria and endoplasmic reticulum, respectively, whereas the other two are found in the cytoplasm along with all of the enzymes of glycolysis.

Biotin An important coenzyme used by pyruvate carboxylase is biotin (Figure 6.27). Biotin is commonly used by carboxylases to carry CO2 to incorporate into the substrate.

Also known as vitamin H, biotin is a water soluble B vitamin (B7) needed for many metabolic processes, including fatty acid synthesis, gluconeogenesis, and amino acid metabolism. Deficiency of the vitamin is rare, since it is readily produced by gut Gluconeogenesis and glycolysis. Only the enzymes differing in gluconeogenesis are shown Image by Aleia Kim teria. There are many claims of advantages of taking biotin supplements, but there is no strong indication of benefits in most cases. Deficiencies are associated with inborn genetic errors, alcoholism, burn patients, and people who have had a gastrectomy. Some pregnant and

lactating women may have reduced levels due to increased biotin catabolism.

Reciprocal regulation

All of the enzymes of glycolysis and nine of the eleven enzymes of gluconeogenesis are all in the cytoplasm, necessitating a coordinated means of controlling them. Cells generally need to minimize the extent to which paired anabolic and catabolic pathways are occurring simultaneously, lest they produce a futile cycle, resulting in wasted energy with no tangible product except heat. The mechanisms of controlling these pathways have opposite effects on catabolic and anabolic processes. This method of control is called reciprocal regulation (see above).

Reciprocal regulation is a coordinated means of simultaneously controlling metabolic pathways that do opposite things. In reciprocal regulation, a single molecule (allosteric regulation) or a single covalent modification (phosphorylation/ dephosphorylation,

Allosteric Regulation of Glycolysis & Gluconeogenesis

Reciprocal Regulation

AMP - Activates PFK-1, Inhibits F1,6BPase

F2,6BP - Activates PFK-1, Inhibits F1,6BPase

Citrate - Activates PFK-1, Inhibits F1,6BPase

Glycolysis Only

ATP - Inhibits PFK-1 and Pyruvate Kinase

Alanine - Inhibits Pyruvate Kinase

Gluconeogenesis Only

ADP - Inhibits Pyruvate Carboxylase and PEPCK

Acetyl-CoA - Activates Pyruvate Carboxylase

Figure 6.27 - Biotin carrying carbon dioxide (red) Wikipedia for example) has opposite effects on the different pathways.

Reciprocal allosteric effects For example, in glycolysis, the enzyme known as phosphofructokinase (PFK-1) is allosterically activated by AMP and a molecule known as F2,6BP (Figure 6.28). The corresponding enzyme from gluconeogenesis catalyzing a reversal of the glycolysis reaction is known as F1,6BPase. F1,6BPase is inhibited by both AMP and F2,6BP.

Reciprocal covalent effects

In glycogen metabolism, the enzymes phosphorylase kinase and glycogen phosphorylase catalyze reactions important for the breakdown of glycogen. The enzyme glycogen synthase catalyzes the synthesis of glyco- Directional velocity Inverts with reciprocity If glycolysis is flowing Glucose synthesis awaits But when the latter is a-going Sugar breakdown then abates Figure 6.28 - Regulation of glycolysis (orange path) and gluconeogenesis (black path) Image by Aleia Kim gen. Each of these enzymes is, at least partly, regulated by attachment and removal of phosphate.

Phosphorylation of phosphorylase kinase and glycogen phosphorylase has the effect of making them more active, whereas phosphorylation of glycogen synthase makes it less active. Conversely, dephosphorylation has the reverse effects on these enzymes - phosphorylase kinase and glycogen phosphorylase become less active and glycogen synthase becomes more active.

Simple and efficient

The advantage of reciprocal regulation schemes is that they are very efficient. It doesn't take separate molecules or separate treatments to control two pathways simultaneously. Further, its simplicity ensures that when one pathway is turned on, the other is turned off.

This is especially important with catabolic/ anabolic regulation, because having both pathways going on simultaneously in a cell is not very productive, leading only to production of heat in a futile cycle. A simple futile cycle is shown on Figure 6.29. If unregulated, the cyclic pathway in the figure (shown in black) will make ATP in creating pyruvate from PEP and will use ATP to make oxaloacetate from pyruvate.

It will also use GTP to make PEP from oxaloacetate. Thus, each turn of the cycle will make one ATP, use one ATP and use one GTP for a net loss of energy. The process will start with pyruvate and end with pyruvate, so there is no net production of molecules. (see HERE for one physiological use of a futile cycle).

Specific gluconeogenesis controls

Besides reciprocal regulation, other mechanisms help control gluconeogenesis. First, PEPCK is controlled largely at the level of synthesis. Overexpression of PEPCK (stimulated by glucagon, glucocorticoid hormones, and cAMP and inhibited by insulin) produces symptoms of diabetes.

Pyruvate carboxylase is sequestered in the mitochondrion (one means of regulation) Figure 6.29 - A simple futile cycle - follow the black lines Image by Aleia Kim Interactive Learning Module HERE and is sensitive to acetyl-CoA, which is an allosteric activator. Acetyl-CoA concentrations increase as the citric acid cycle activity decreases. Glucose-6- phosphatase is present in low concentrations in many tissues, but is found most abundantly and importantly in the major gluconeogenic organs – the liver and kidney cortex.

Specific glycolysis controls

Control of glycolysis and gluconeogenesis is unusual for metabolic pathways, in that regulation occurs at multiple points. For glycolysis, this involves three enzymes:

1. Hexokinase (Glucose ⇄ G6P)

2. Phosphofructokinase-1 (F6P ⇄ F1,6BP)

3. Pyruvate kinase (PEP ⇄ Pyruvate).

Regulation of hexokinase is the simplest of these. The enzyme is unusual in being inhibited by its product, glucose-6-phosphate. This ensures when glycolysis is slowing down hexokinase is also slowing down to reduce feeding the pathway.

Pyruvate kinase

It might also seem odd that pyruvate kinase, the last enzyme in the pathway, is regulated (Figure 6.30), but the reason 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 the gluconeogenesis pathway. In other words, it takes two enzymes, two reactions, and two triphosphates (ATP and GTP) to go from one pyruvate back to one PEP in gluconeogenesis. When cells are needing to make glu- igure 6.30 - Regulation of pyruvate kinase For cells a glucose cycling's cost Is energy in reams Four ATPs each time is lost From breaking/making schemes So use for metabolic heat To make it warm inside your feet Else it's of no utility To practice such futility cose, 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 must be inhibited during gluconeogenesis or a futile cycle will occur and no glucose will be made.

Another interesting control mechanism called feedforward activation involves pyruvate kinase. Pyruvate kinase is activated allosterically by the glycolysis intermediate, F1,6BP. This molecule is a product of the PFK-1 reaction and a substrate for the aldolase reaction.

Reactions pulled

As noted above, the aldolase reaction is energetically unfavorable (high positive $\Delta G^{\circ\prime}$), thus allowing F1,6BP to accumulate. When this happens, some of the excess F1,6BP binds to pyruvate kinase, which activates and jump- Figure 6.31 - Regulation of Synthesis and Breakdown of F2,6BP Image by Penelope Irving 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 GLYAL3P and DHAP fall, helping to pull the aldolase reaction forward.

PFK-1 regulation

PFK-1 has a complex regulation scheme. First, it is reciprocally regulated (relative to F1,6BPase) by three molecules. F2,6BP activates PFK-1 and inhibits F1,6BPase. PFK-1 is also allosterically activated by AMP, whereas F1,6BPase is inhibited. On the other hand, citrate inhibits PFK-1, but activates F1,6BPase.

PFK-1 is also inhibited by ATP and is exquisitely sensitive to proton concentration, easily losing activity when the pH drops only slightly. PFK- 1's inhibition by ATP is noteworthy and odd at first glance because ATP is also a substrate whose increasing concentration should favor the reaction instead of inhibit it. The root of this conundrum is that PFK-1 has two ATP binding sites - one at an allosteric site that binds ATP relatively inefficiently and one that the active site that binds ATP with high affinity. Thus, only when ATP concentration is high is binding at the allosteric site favored and only then can ATP turn off the enzyme.

F2,6BP regulation

Regulation of PFK-1 by F2,6BP is simple at the PFK-1 level, but more complicated at the level of synthesis of F2,6BP. Despite having a name sounding like a glycolysis/ gluconeogenesis intermediate (F1,6BP), F2,6BP is not an intermediate in either pathway. Instead, it is made from fructose-6-phosphate and ATP by the enzyme known as phosphofructokinase-2 (PFK- 2 - Figure 6.31).

Cori cycle

With respect to energy, the liver and muscles act complementarily. The liver is the major or- Figure 6.32 - The Cori cycle Image by Aleia Kim gan in the body for the synthesis of glucose. Muscles are major users of glucose to make ATP. Actively exercising muscles use oxygen faster than the blood can deliver it. As a consequence, the muscles go anaerobic and produce lactate. This lactate is of no use to muscle cells, so they dump it into the blood. Lactate travels in the blood to the liver, which takes it up and reoxidizes it back to pyruvate, catalyzed by the enzyme lactate dehydrogenase (Figure 6.32).

Pyruvate in the liver is then converted to glucose by gluconeogenesis. The glucose thus made by the liver is dumped into the bloodstream where it is taken up by muscles and used for energy, completing the important intercellular pathway known as the Cori cycle.

Glucose alanine cycle

The glucose alanine cycle (also known as the Cahill Cycle), has been described as the amine equivalent of the Cori cycle (Figure 6.33). The Cori cycle, of course, exports lac- Figure 6.33 - Overlap between the Cori cycle and the glucose alanine cycle tate from muscles (when oxygen is limiting) to the liver via the bloodstream. The liver, in turn, converts lactate to glucose, which it ships back to the muscles via the bloodstream. The Cori Cycle is an essential source of glucose energy for muscles during periods of exercise when oxygen is used faster than it can be delivered.

In the glucose-alanine cycle, cells are generating toxic amines and must export them. This is accomplished by transaminating pyruvate (the product of glycolysis) to produce the amino acid alanine.

The glucose-alanine process requires the enzyme alanine aminotransferase, which is found in muscles, liver, and intestines. Alanine is exported in the process to the blood and picked up by the liver, which deaminates it to release the amine for synthesis of urea and excretion. The pyruvate left over after the transamination is a substrate for gluconeogenesis. Glucose produced in the liver is then exported to the blood for use by cells, thus completing the cycle.

Polysaccharide metabolism

Sugars are metabolized rapidly in the body and that is one of the primary reasons they are used. Managing levels of glucose in the body is very important - too much leads to complications related to diabetes and too little gives rise to hypoglycemia (low blood sugar). Sugars in the body are maintained by three processes - 1) diet; 2) synthesis (gluconeogenesis); and 3) storage. The storage forms of sugars are, of course, the polysaccharides and their metabolism is our next topic of discussion.

Amylose and amylopectin

The energy needs of a plant are much less dynamic than those of animals. Muscular contraction, nervous systems, and information processing in the brain require large amounts of quick energy. Because of this, the polysaccharides stored in plants are somewhat less complicated than those of animals. Plants store glucose for energy in the form of amylose (Figure 6.34 and see HERE) and amylopectin and for structural integrity in the form of cellulose (see HERE). These structures differ in that cellulose contains glucose units solely joined by β-1,4 bonds, whereas amylose has only α-1,4 bonds and amylopectin has α-1,4 and α-1,6 bonds. Figure 6.34 Amylose, a polymer of glucose in plants

Glycogen

Animals store glucose primarily in liver and muscle in the form of a compound related to amylopectin known as glycogen. The structural differences between glycogen and amylopectin are solely due to the frequency of the α-1,6 branches of glucoses. In glycogen they occur about every 10 residues instead of every 30-50, as in amylopectin (Figure 6.35).

Glycogen provides an additional source of glucose besides that produced via gluconeogenesis. Because glycogen contains so many glucoses, it acts like a battery backup for the body, providing a quick source of glucose when needed and providing a place to store excess glucose when glucose concentrations in the blood rise.

The branching of glycogen is an important feature of the molecule metabolically as well. Since glycogen is broken down from the "ends" of the molecule, more branches translate to more ends, and more glucose that can be released at once.

Just as in gluconeogenesis, the cell has a separate mechanism for glycogen synthesis that is distinct from glycogen breakdown. As noted previously, this allows the cell to separately control the reactions, avoiding futile cycles, and enabling a process to occur efficiently (synthesis of glycogen) that would not occur if Figure 6.35 - Glycogen Structure - α-1,4 links with α-1,6 branches every 7-10 residues it were simply the reversal of glycogen breakdown.

Glycogen breakdown

Breakdown of glycogen involves 1) release of glucose-1-phosphate (G1P), 2) rearranging the remaining glycogen (as necessary) to permit continued breakdown, and 3) conversion of G1P to G6P for further metabolism. G6P can be 1) used in glycolysis, 2) converted to glucose by gluconeogenesis, or 3) oxidized in the pentose phosphate pathway.

Glycogen phosphorylase (sometimes simply called phosphorylase) catalyzes breakdown of glycogen into glucose-1- Phosphate (G1P - Figure 6.36). The reaction that produces G1P from glycogen is a phosphorolysis, not a hydrolysis reaction. The distinction is that hydrolysis reactions use water to cleave bigger molecules into smaller ones, but phosphorolysis reactions use phosphate instead for the same purpose. Note that the phosphate is just that - it does NOT come from ATP. Since ATP is not used to put phosphate on G1P, the reaction saves the cell energy.

Glycogen debranching enzyme

Glycogen phosphorylase will only act on nonreducing ends of a glycogen chain that are at least 5 glucoses away from a branch point. A second enzyme, Glycogen Debranching Enzyme (GDE) (also called debranching enzyme), is therefore needed to convert α (1-6) branches to α (1-4) branches. GDE acts on glycogen branches that have reached their limit of phosphorylysis with glycogen phosphorylase. Figure 6.36 - Breaking of α-1,4 bonds of glycogen by glycogen phosphorylase Image by Aleia Kim Interactive Learning Module HERE

GDE acts to transfer a trisaccharide from an α-1,6 branch onto an adjacent α-1,4 branch, leaving a single glucose at the 1,6 branch. Note that the enzyme also catalyzes the hydrolysis of the remaining glucose at the 1,6 branch point (Figure 6.37). Thus, the breakdown products from glycogen are G1P and glucose (mostly G1P). Glucose can, of course, be converted to Glucose-6-Phosphate (G6P) as the first step in glycolysis by either hexokinase or glucokinase.

G1P can be converted to G6P by action of an enzyme called phosphoglucomutase. This reaction is readily reversible, allowing G6P and G1P to be interconverted as the concentration of one or the other increases. This is important, because phosphoglucomutase is needed to form G1P for glycogen synthesis.

Regulation of glycogen metabolism

Regulation of glycogen metabolism is complex, occurring both allosterically and via hormone-receptor controlled events that result in protein phosphorylation or dephosphorylation. In order to avoid a futile cycle of glycogen synthesis and breakdown simultaneously, cells have evolved an elaborate set of controls that ensure only one pathway is primarily active at a time.

Regulation of glycogen metabolism is managed by the enzymes glycogen phosphorylase and glycogen synthase. Glycogen phosphorylase is regulated by both allosteric factors (ATP, G6P, AMP, and glucose) and by covalent modification (phosphorylation / dephosphorylation). Its regulation is consistent with the energy needs of the cell. High energy molecules (ATP, G6P, glucose) al- Figure 6.37 - Catalytic activity of debranching enzyme losterically inhibit glycogen phosphorylase, while the low energy molecule AMP allosterically activates it.

GPa/GPb allosteric regulation

Glycogen phosphorylase exists in two different covalent forms – one form with phosphate (called GPa here) and one form lacking phosphate (GPb here). GPb is converted to GPa by phosphorylation by an enzyme known as phosphorylase kinase. GPa and GPb can each exist in an 'R' state and a 'T' state (Figure 6.38). For both GPa and GPb, the R state is the more active form of the enzyme. GPa's negative allosteric effector (glucose) is usually not abundant in cells, so GPa does not flip into the T state often. There is no positive allosteric effector of GPa. When glucose is absent, GPa automatically flips into the R (more active) state (Figure 6.39). It is for this reason that people tend to think of GPa as being the more active covalent form of the enzyme.

GPb can convert from the GPb T state to the GPb R state by binding AMP. Unless a cell is low in energy, AMP concentration is low. Thus GPb is not converted Figure 6.38 - Glycogen phosphorylase regulation - covalent (horizontal) and allosteric (vertical) Image by Aleia Kim to the R state very often. This is why people think of the GPb form as less active than GPa. On the other hand, ATP and/or G6P are usually present at high enough concentration in cells that GPb is readily flipped into the T state (Figure 6.40).

GPa/GPb covalent regulation

The relative amounts of GPa and GPb largely govern the overall process of glycogen breakdown, since GPa tends to be active more often than GPb. It is i

Phosphorylase kinase itself has two covalent forms – phosphorylated (active) and dephosphorylated (inactive). It is phosphorylated by the enzyme Protein Kinase A (PKA -). Another way to activate the enzyme is allosterically with calcium (Figure 6.41). Phosphory- Figure 6.39 - Allosteric regulation of GPa Image by Aleia Kim Figure 6.40 - Allosteric regulation of GPb Image by Aleia Kim lase kinase is dephosphorylated by phosphoprotein phosphatase, the same enzyme that removes phosphate from GPa.

PKA and cAMPcAMP

PKA is activated by cAMP, which is, in turn, produced by adenylate cyclase after activation by a G-protein (See HERE for overview). G-proteins are activated ultimately by binding of ligands to specific membrane receptors called 7-TM receptors, also known as Gprotein coupled receptors. These are discussed in greater detail HERE. Common ligands for 7-TM receptors include epinephrine (binds β- adrenergic receptor) and glucagon (binds glucagon receptor). Epinephrine exerts its greatest effects on muscle and glucagon works preferentially on the liver. Thus, epinephrine and glucagon can activate glycogen breakdown by stimulating synthesis of cAMP followed by the cascade of events described above.

Turning off glycogen breakdown

Turning off signals is as important, if not more so, than turning them on. Glycogen is a precious resource. If its breakdown is not controlled, a lot of energy used in its synthesis is wasted. The steps in the glycogen breakdown regulatory pathway can be reversed at every level. First, the ligand (epinephrine or glucagon) can leave the receptor, turning off the stimulus. Second, the G-proteins have an inherent GTPase activity. GTP, of course, is what activates Gproteins, so a GTPase activity converts the GTP it is carrying to GDP and the G-protein becomes inactive. Thus, G-proteins turn off Figure 6.41 - Activation of phosphorylase kinase Image by Aleia Kim their own activity. Interfering with their ability to convert GTP to GDP can have dire consequences, including cancer in some cases.

Third, cells have phosphodiesterase enzymes (inhibited by caffeine) for breaking down cAMP. cAMP is needed to activate PKA, so breaking it down stops PKA from activating phosphorylase kinase. Fourth, the enzyme known as phosphoprotein phosphatase (also called PP1) plays a major role. It can remove phosphates from phosphorylase kinase (inactivating it) and form GPa, converting it to the less likely to be active GPb. Regulation of phosphoprotein phosphatase activity occurs at several levels. Two of these are shown in Figures 6.42 & 6.43.

In Figure 6.42, phosphoprotein phosphatase is shown being inactivated by phosphorylation of an inhibitor (called PI-1 - see below). This happens as a result of cascading actions from binding of epinephrine (or glucagon) to a cell's β-adrenergic receptor. Reversal of these actions occurs when insulin binds to the cell's insulin receptor, resulting in activation of phosphoprotein phosphatase.

PI-1

The inhibitor PI-1 can block activity of phosphpoprotein phosphatase only if it (PI-1) is phosphorylated. When PI-1 gets dephosphorylated, it no longer functions as an inhibitor, so phosphoprotein phosphatase be- Figure 6.42 - Inactivation of phosphoprotein phosphatase by protein kinase A via phosphorylation of PI-1 (Inhibitor) and the GM binding protein Image by Pehr Jacobson Interactive Learning Module HERE comes active. Now, here is the clincher - PI-1 gets phosphorylated by PKA

(thus, when epinephrine or glucagon binds to a cell) and gets dephosphorylated when insulin binds to a cell.

Another regulatory mechanism

Another way to regulate phosphoprotein phosphatase in the liver involves GPa directly (Figure 6.43). In liver cells, phosphoprotein phosphatase is bound to a protein called GL. GL can also bind to GPa. As shown in the figure, if the three proteins are complexed together (top of figure), then PP1 (phosphoprotein phosphatase) is inactive. When glucose is present (such as when the liver has made too much glucose), then the free glucose binds to the GPa and causes GPa to be released from the GL.

This has the effect of activating phosphoprotein phosphatase, which begins dephosphorylating enzymes. As shown in the figure, two such enzymes are GPa (making GPb) and glycogen synthase b, making glycogen synthase a. These dephosphorylations have opposite effects on the two enzymes, making GPb, which is less active and glycogen synthase a, which is much more active.

Glycogen synthesis

The anabolic pathway opposing glycogen breakdown is that of glycogen synthesis. Just Figure 6.43 - Regulation of phosphoprotein phosphatase (PP-1) activity by GPa Image by Penelope Irving as cells reciprocally regulate glycolysis and gluconeogenesis to prevent a futile cycle between these pathways, so too do cells use reciprocal schemes to regulate glycogen breakdown and synthesis.

Synthesis of glycogen starts with G1P, which is converted to an 'activated' intermediate, UDPglucose. This activated intermediate is what 'adds' the glucose to the growing glycogen chain in a reaction catalyzed by the enzyme known as glycogen synthase (Figure 6.44). Once the glucose is added to glycogen, the glycogen molecule may need to have branches inserted in it by the enzyme known as branching enzyme (Figure 6.45).

Steps

Let us first consider the steps in glycogen synthesis. 1) Glycogen synthesis from glucose involves phosphorylation to form G6P, and isomerization to form G1P (using phospho- igure 6.45 - Branch formation in glycogen by branching enzyme Image by Penelope Irving Figure 6.44 - Catalytic activity of glycogen synthase Image by Penelope Irving glucomutase, common to glycogen breakdown). G1P is reacted with UTP to form UDP-glucose in a reaction catalyzed by UDP-glucose pyrophosphorylase. Glycogen synthase catalyzes synthesis of glycogen by joining carbon #1 of the UDP-derived glucose onto the carbon #4 of the non-reducing end of a glycogen chain, to form the familiar α(1,4) glycogen links. Another product of the reaction is UDP.

“Primer” requirements

It is also worth noting, in passing, that glycogen synthase will only add glucose units from UDP-Glucose onto a preexisting glycogen chain that has at least four glucose residues. Linkage of the first few glucose units to form the minimal "primer" needed for glycogen synthase recognition is catalyzed by a protein called glycogenin, which attaches to the first glucose and catalyzes linkage of the first eight glucoses by α(1,4) bonds. 3) The characteristic α(1,6) branches of glycogen are the products of the enzyme known as branching enzyme. Branching enzyme breaks α(1,4) chains and carries the broken chain to the carbon #6 and forms an α(1,6) linkage (Figure 6.45).

Regulation of glycogen synthesis

The regulation of glycogen biosynthesis is reciprocal to that of glycogen breakdown. It also has a cascading covalent modification system similar to the glycogen breakdown system described above. In fact, part of the system is identical to glycogen breakdown. Epinephrine or glucagon signaling stimulates adenylate cyclase to make cAMP, which activates PKA. Figure 6.46 - Reciprocal regulation by the phosphorylation cascade - glycogen breakdown activated / glycogen synthesis inhibited Image by Penelope Irving

Effect of phosphorylation

In glycogen synthesis, protein kinase A phosphorylates the active form of glycogen synthase (GSa), and converts it into the usually inactive b form (called GSb).

Note the conventions for glycogen synthase and glycogen phosphorylase. For both enzymes, the more active forms are called the 'a' forms (GPa and GSa) and the less active forms are called the 'b' forms (GPb and GSb). The major difference, however, is that GPa has a phosphate, but GSa does not and GPb has no phosphate, but GSb does.

Thus phosphorylation and dephosphorylation have opposite effects on the enzymes of glycogen metabolism (Figure 6.46). This is the hallmark of reciprocal regulation. It is of note that the less active glycogen synthase form, GSb, can be activated by G6P. Recall that G6P had the exactly opposite effect on GPb.

Glycogen synthase, glycogen phosphorylase (and phosphorylase kinase) can all be dephosphorylated by the same enzyme - phosphoprotein phosphatase - and it is activated when insulin binds to its receptor in the cell membrane.

Big picture

In the big picture, binding of epinephrine or glucagon to appropriate cell receptors stimulates a phosphorylation cascade which simultaneously activates breakdown of glycogen by glycogen phosphorylase and inhibits synthesis of glycogen by glycogen synthase. Epinephrine, is also known as adrenalin, and the properties that adrenalin gives arise from a large temporary increase of blood glucose, which powers muscles.

On the other hand, insulin stimulates dephosphorylation by activating phosphoprotein phosphatase. Dephosphorylation reduces action of glycogen phosphorylase (less glycogen breakdown) and activates glycogen synthase (starts glycogen synthesis). Our bodies make glycogen when blood glucose levels rise. Since high blood glucose levels are harmful, insulin stimulates cells to take up glucose. In the liver and in muscle cells, the uptaken glucose is made into glycogen. Figure 6.47 - Cotton - the purest natural form of cellulose Wikipedia Interactive Learning Module HERE

Cellulose synthesis

Cellulose is synthesized as a result of catalysis by cellulose synthase. Like glycogen synthesis it requires an activated intermediate to add glucose residues and there are two possible ones - GDP-glucose and UDPglucose, depending on which cellulose synthase is involved. In plants, cellulose provides support to cell walls.

The reaction catalyzed is shown next where Cellulosen = a polymer of [(1→4)-β-Dglucosyl] n units long.

The GDP-glucose reaction is the same except with substitution of GDP-glucose for UDP-Figure 6.48 - The Pentose Phosphate Pathway - Enzymes - 1 = G6P dehydrogenase / 2 = 6-Phosphogluconolactonase / 3 = 6-PG dehydrogenase / 4a = Ribose 5- phosphate isomerase / 4b = Ribulose 5-phosphate 3-epimerase / 5,7 = Transketolase / 6 = Transaldolase UDP-glucose + Cellulosen UDP + Cellulosen+1 glucose. UDP-glucose for the reaction is obtained by catalysis of sucrose synthase. The enzyme is named for the reverse reaction.

Pentose phosphate pathway

The pentose phosphate pathway (PPP - also called the hexose monophosphate shunt) is an oxidative pathway involving sugars that is sometimes described as a parallel pathway to glycolysis. It is, in fact, a pathway with multiple inputs and outputs (Figure 6.48). PPP is also a major source of NADPH for biosynthetic reactions and can provide ribose-5-phosphate for nucleotide synthesis.

Though when drawn out, the pathway's "starting point" is often shown as glucose-6-phosphate (G6P), in fact there are multiple entry points including other glycolysis intermediates, such as fructose-6-phosphate (F6P) and glyceraldehyde-3-phosphate (GLYAL-3-P), as well as less common sugar compounds with 4,5, and 7 carbons.

The multiple entry points and multiple outputs gives the cell tremendous flexibility to meet its needs by allowing it to use a variety of materials to make any of these products.

Oxidation #1

Beginning with G6P, PPP proceeds through its oxidative phase as follows:

The enzyme catalyzing the reaction is G6P dehydrogenase. It is the rate limiting step of the pathway and the enzyme is inhibited both by NADPH and acetyl-CoA. NADPH is important for anabolic pathways, such as fatty acid synthesis and also for maintaining glutathione in a reduced state. The latter is important in protection against damage from reactive oxygen species.

Deficiency of the G6P dehydrogenase enzyme is not rare, leading to acute hemolytic anemia, due to reduced NADPH concentration, and a reduced ability of the cell to disarm reactive oxygen species with glutathione. Reduced activity of the enzyme appears to have a protective effect against malarial infection, likely due to the increased fragility of the red blood cell membrane, which is then unable to sustain an infection by the parasite. Hydrolysis Reaction #2 is a hydrolysis and it is catalyzed by

Hydrolysis

Reaction #2 is a hydrolysis and it is catalyzed by 6-phosphogluconolactonase. The reac- Sucrose + UDP UDP-glucose + Fructose G6P + NADP+ 6-Phosphoglucono-δ-lactone + NADPH tion converts the circular 6-phosphoglucono- δ-lactone into the linear 6-phosphogluconate (6-PG) in preparation for oxidation in the next step.

Decarboxylation

Reaction #3 is the only decarboxylation in the PPP and the last oxidative step. It is catalyzed by 6-phosphogluconate dehydrogenase.

Mutations disabling the protein made from this gene negatively impact red blood cells. At this point, the oxidative phase of PPP is complete and the remaining reactions involve molecular rearrangements. Ru5P has two possible fates and these are each described below.

Isomerization

Reaction 4a: The enzyme catalyzing this reversible reaction is Ru5P isomerase (top of next column). It is important because this is the way cells make R-5-P for nucleotide synthesis. The R-5-P can also be used in other PPP reactions shown elsewhere.

Epimerization

Reaction 4b (catalyzed by Ru-5-P epimerase) is another source of a pentose sugars and provides an important substrate for subsequent reactions.

Transketolase reactions

The other reactions don't really have an order to them and whether they occur or not depends on cellular needs. The first enzyme, transketolase, is flexible in terms of its substrate/product combinations and is used not only in PPP, but also in the Calvin cycle of plants. It catalyzes the next two reactions

In the first reaction (above), two phosphorylated sugars of 5 carbons each are converted into one phosphorylated sugar of 3 carbons and one of 7 carbons. In the second (next page), a five carbon sugar phosphate and aRu-5-P Xylulose-5-phosphate (Xu-5-P) Xu-5-P + R-5-P GLYAL-3-P + Sedoheptulose-7-phosphate (S-7-P) 6-PG + NADP+ Ribulose-5-phosphate (Ru-5-P) + NADPH + CO2 6-Phosphoglucono-δ-lactone + H2O 6-phosphogluconate (6-PG) + H+ Ru-5-P Ribose-5-phosphate (R-5-P) four carbon sugar phosphate are rearranged into sugar phosphates with 3 and 6 carbons.

Glycolysis intermediates

In the reversible reactions of the pentose phosphate pathway, one can see how glycolysis intermediates can easily be rearranged and made into other sugars. Thus, GLYAL-3-P and F6P can be readily made into Ribose-5- phosphate for nucleotide synthesis.

Involvement of F6P in the pathway permits cells to continue making nucleotides (by making R-5-P) or tryptophan (by making E-4-P) even if the oxidative reactions of PPP are inhibited.

The last reaction is catalyzed by the enzyme known as transaldolase.

TPP co-factor

Transketolase uses thiamine pyrophosphate (TPP) to catalyze reactions. TPP's thiaFigure 6.49 - Intermediates of the pentose phosphate pathway Xu-5-P + Erythrose-4-phosphate (E-4-P) GLYAL-3-P + F6P GLYAL-3-P + S-7-P E-4-P + F6P zole ring's nitrogen and sulfur atoms on either side of a carbon, allow it to donate a proton and act as an acid, thus forming a carbanion, which gets stabilized by the adjacent tetravalent nitrogen (Figures 6.50 & 6.51)).

The stabilized carbanion plays important roles in the reaction mechanism of enzymes, such as transketolase that use TPP as a cofactor. Commonly, the carbanion acts as a nucleophile that attacks the carbonyl carbon of the substrate. Such is the case with transketolase. Attack by the carbanion breaks the carbonyl bond on the substrate and covalently links it to the ionized carbon of TPP, thus allowing it to "carry" the carbonyl group to the other substrate for attachment. In this way, two carbons are moved from Xu- 5-P to E-4-P to make F6P (from E-4-P) and GLYAL-3-P (from Xu-5-P). Similarly, S-7-P and GLYAL-3-P are made from R-5-P and Xu-5-P, respectively.

Thiamines

Thiamines are a class of compounds involved in catalysis of important respiration-related The Pentose Phosphate Pathway by Kevin Ahern I need erythrose phosphate And don't know what to do My cells are full of G-6-P And NADP too But I just hit upon a plan As simple as can be I'll run reactions through the path That's known as PPP In just two oxidations There's ribulose-5P Which morphs to other pentoses Each one attached to P The next step it is simple Deserving of some praise The pentose carbons mix and match Thanks to transketolase Glyceraldehyde's a product Sedoheptulose is too Each with a trailing phosphate But we are not quite through Now three plus seven is the same As adding six and four By swapping carbons back and forth There's erythrose-P and more At last I've got the thing I need From carbons trading places I'm happy that my cells are full Of some transaldolases Figure 6.50 - Thiamine pyrophosphate reactions in the citric acid cycle, pyruvate metabolism, the pentose phosphate pathway, and the Calvin cycle. Thiamine was the first water-soluble vitamin (B1) to be discovered via association with the peripheral nervous system disease known as Beriberi. Thiamine pyrophosphate (TPP) is an enzyme cofactor found in all living systems derived from thiamine by action of the enzyme thiamine diphosphokinase. TPP facilitates catalysis of several biochemical reactions essential for tissue respiration.

Deficiency of the vitamin is rare today, though people suffering from Crohn's disease, anorexia, alcoholism or undergoing kidney dialysis may develop deficiencies. TPP is required for the oxidative decarboxylation of pyruvate to form acetyl-CoA and similar reactions. Transketolase, an important enzyme in the pentose phosphate pathway, also uses it as a coenzyme. Besides these reactions, TPP is also required for oxidative decarboxylation of α-keto acids like α-ketoglutarate and branched-chain α-keto acids arising from metabolism of valine, isoleucine, and leucine. Figure 6.51 - Mechanism of action of thiamine pyrophosphate (TPP) - 1) Carbanion formation; 2) Nucleophilic attack; 3) Covalent attachment of carbonyl; 4) Transfer to second group; 5) Release of product and regeneration of TPP

TPP acts in the pyruvate dehydrogenase complex to assist in decarboxylation of pyruvate and "carrying" the activated acetaldehyde molecule to its attachment (and subsequent oxidation) to lipoamide. Central to TPP's function is the thiazolium ring, which stabilizes carbanion intermediates (through resonance) by acting as an electron sink (Figure 6.51). Such action facilitates breaking of carbon-carbon bonds such as occurs during decarboxylation of pyruvate to produce the activated acetaldehyde.

Thiamine deficiency

Thiamine is integral to respiration and is needed in every cell. Acute deficiency of thiamine leads to numerous problems - the best known condition is beriberi, whose symptoms include weight loss, weakness, swelling, neurological issues, and irregular heart rhythms. Figure 6.52 - The Calvin cycle - The resynthesis phase has multiple steps and is described below. Image by Aleia Kim

Causes of deficiency include poor nutrition, significant intake of foods containing the enzyme known as thiaminase, foods with compounds that counter thiamine action (tea, coffee), and chronic diseases, including diabetes, gastrointestinal diseases, persistent vomiting. People with severe alcoholism often are deficient in thiamine.

Calvin cycle

The Calvin cycle (Figure 6.52) is a metabolic pathway occurring exclusively in photosynthetic organisms. Commonly referred to as the "Dark Cycle" or the Light-Independent Cycle, the Calvin cycle does not actually occur in the dark. The cell/chloroplast simply is not directly using light energy to drive it.

Assimilation

It is in the Calvin cycle of photosynthesis that carbon dioxide is taken from the atmosphere and ultimately built into glucose (or other sugars). Reactions of the Calvin cycle take place in regions of the chloroplast known as the stroma, the fluid areas outside of the thylakoid membranes. The cycle can be broken into three phases

1) assimilation of CO2

2) reduction reactions

3) regeneration of the starting material, ribulose 1,5 bisphosphate (Ru1,5BP).

Though reduction of carbon dioxide to glucose ultimately requires electrons from twelve molecules of NADPH (and 18 ATPs), it is confusing because one reduction occurs 12 times (1,3 BPG to GLYAL-3P) to input the overall reduction necessary to make one glucose.

Carbon dioxide

Another reason students find the pathway confusing is because the carbon dioxide molecules are absorbed one at a time into six different molecules of Ru1,5BP. At no point are the six carbons ever together in the same molecule to make a single glucose.

Instead, six molecules of Ru1,5BP (30 carbons) gain six more carbons via carbon dioxide and then split into 12 molecules of 3-phosphoglycerate (36 carbons). The gain of six carbons allows two three carbon molecules to be produced in excess for each turn of the cycle. These two molecules molecules are then converted into glucose using the enzymes of gluconeogenesis. The other ten molecules of 3-PG are used to regenerate the six molecules of Ru1,5BP. Figure 6.53 - Rubisco, the most abundant enzyme on Earth

Cyclic pathway

Like the citric acid cycle, the Calvin cycle doesn't really have a starting or ending point, but can we think of the first reaction as the fixation of carbon dioxide to Ru1,5BP. This reaction is catalyzed by the enzyme known as ribulose-1,5 bisphosphate carboxylase (RUBISCO - Figure 6.53). The resulting six carbon intermediate is unstable and is rapidly converted to two molecules of 3-phosphoglycerate.

As noted, if one starts with 6 molecules of Ru1,5BP and makes 12 molecules of 3-PG, the extra 6 carbons that are a part of the cycle can be shunted off as two three-carbon molecules of glyceraldehyde-3-phosphate (GLYAL3P) to gluconeogenesis, leaving behind 10 molecules to be reconverted into 6 moleFigure 6.54 - Resynthesis phase of the Calvin cycle - All paths lead to regenerating Ru1,5BP, which is the aim of the resynthesis phase. Glycolysis/gluconeogenesis intermediates shown in blue. Enzyme numbers explained in text. cules of Ru1,5BP. This occurs in what is called the resynthesis phase.

Resynthesis phase

The resynthesis phase (Figure 6.54) requires multiple steps, but only utilizes two enzymes unique to plants - sedoheptulose-1,7 bisphosphatase and phosphoribulokinase. RUBISCO is the third (and only other) enzyme of the pathway that is unique to plants.

All of the other enzymes of the pathway are common to plants and animals and include some found in the pentose phosphate pathway and gluconeogenesis. Enzymes shown as numbers in Figure 6.54 are as follows (enzymes unique to plants in green):

1 - Phosphoglycerate kinase

2 - G3PDH

3 - Triosephosphate Isomerase

4 - Aldolase

5 - Fructose 1,6 bisphosphatase

6 - Transketolase

7 - Phosphopentose Epimerase

8 - Phosphoribulokinase

9 - Sedoheptulose 1,7 bisphosphatase

10 - Phosphopentose Isomerase

Reactions

The resynthesis phase begins with conversion of the 3-PG molecules into GLYAL3P (there are actually 10 GLYAL3P molecules involved in resynthesis, as noted above, but we are omitting numbers to try to help students to see the bigger picture. Suffice it to say that there are sufficient quantities of all of the molecules to complete the reactions described). Some GLYAL3P is converted to DHAP by triose phosphate isomerase. Some DHAPs are converted (via gluconeogenesis) to F6P (one phosphate is lost for each F6P).

Two carbons from F6P are given to GLYAL3P to create E-4P and Xu-5P (reversal of PPP reaction). E- 4P combines with DHAP to form sedoheptulose-1,7 bisphosphate (S1,7BP). The phosphate at position #1 is Figure 6.55 - Use of CO2 (Calvin cycle) vs. O2 (photorespiration) by RUBISCO. Image by Pehr Jacobson cleaved by sedoheptulose-1,7 bisphosphatase to yield S-7-P. Transketolase (another PPP enzyme) catalyzes transfer to two carbons from S-7-P to GLYAL3P to yield Xu-5P and R5P.

Phosphopentose isomerase catalyzes conversion of R5P to Ru5P and phosphopentose epimerase similarly converts Xu-5P to Ru5P. Finally, phosphoribulokinase transfers a phosphate to Ru5P (from ATP) to yield Ru1,5BP.

Photorespiration

In the Calvin cycle of photosynthesis, the enzyme ribulose-1,5-bisphosphate carboxylase (RUBISCO) catalyzes the addition of carbon dioxide to ribulose-1,5- bisphosphate (Ru1,5BP) to create two molecules of 3-phosphoglycerate. Molecular oxygen (O2), however, competes with CO2 for this enzyme, so about 25% of the time, the molecule that gets added is not CO2, but rather O2 (Figure 6.55). When this happens, the following reaction occurs

This is the first step in the process known as photorespiration. The process of photorespiration is inefficient relative to the carboxylation of Ru1,5BP. Phosphoglycolate is converted to glyoxylate in the glyoxysome and then transamination of that yields glycine. Two glycines can combine in a complicated coupled set of reactions in the mitochondrion shown next. Figure 6.56 - Maize - a C4 plant Ru1,5BP + O2 Phosphoglycolate + 3-phosphoglycerate + 2H+ 2 Glycine + NAD+ + H2O Serine + CO2 + NH3 + NADH + H+

Deamination and reduction of serine yields pyruvate, which can be then be converted back to 3-phosphoglycerate. The end point of oxygenation of Ru1,5BP is the same as the carboxylation of Ru1,5BP reactions, but there are significant energy costs associated with it, making the process less efficient.

C4 plants

The Calvin cycle is the means by which plants assimilate carbon dioxide from the atmosphere, ultimately into glucose. Plants use two general strategies for doing so. The first is employed by plants called C3 plants (most plants) and it simply involves the pathway described above. They are called C3 plants because the first stable intermediate after absorbing carbon dioxide contains three carbons - 3-phosphoglycerate. Another class of plants, called C4 plants (Figure 6.56) employ a novel strategy for concentrating the Figure 6.57 - Assimilation of CO2 by C4 plants Image by Aleia Kim CO2 prior to assimilation. C4 plants are generally found in hot, dry environments where conditions would otherwise favor the wasteful photorespiration reactions of RUBISCO and loss of water.

Capture by PEP

In C4 plants, carbon dioxide is captured in special mesophyll cells first by phosphoenolpyruvate (PEP) to make oxaloacetate (contains four carbons and gives the C4 plants their name - Figure 6.57). The oxaloacetate is converted to malate and transported into bundle sheath cells where the carbon dioxide is released and captured by Ru1,5BP, as in C3 plants. The Calvin cycle proceeds from there. The advantage of the C4 plant scheme is that it allows concentration of carbon dioxide while minimizing loss of water and photorespiration.

Peptidoglycan synthesis

Bacterial cell walls contain a layer of protection known as the peptidoglycan layer. Assembly of the layer begins in the cytoplasm.

Steps in the process follow

1. Donation of an amine from glutamine to fructose-6- phosphate and isomerization to make glucosamine-6- phosphate.

2. Donation of an acetyl group from acetyl-CoA to make N-acetylglucosamine-6- phosphate

3. Isomerization of N-acetylglucosamine-6- phosphate makes N-acetylglucosamine-1- phosphate Figure 6.58 - Peptidoglycan layer in a bacterial outer cell wall Wikipedia

4. UTP combines with N-acetylglucosamine-1-phosphate to make UDP-N-acetyl-glucosamine-1- phosphate

5. Addition of PEP and electrons from NADPH yields UDP-Nacetylmuramic acid

6. A pentapeptide or tetrapeptide chain is attached to the UDP-Nacetylmuramic acid. The sequence varies a bit between species, but commonly is L-Ala - D-Glu - L-Lys - DAla - D-Ala

7. Dolichol phosphate replaces UMP on the UDP-N-acetylmuramic acid-pentapeptide.

8. UDP-N-acetyl-glucosamine donates a glucose to the Nacetylmuramic acid part of the Dolichol-PP-N-acetylmuramic acidpentapeptide

9. A pentapeptide chain of glycines (pentaglycine) is linked to lysine of the pentapeptide chain to create a Dolichol-PP-Nacetylmuramic acid-N-acetylglucosaminedecapeptide. The pentaglycine serves as cross links in the overall structure.

10. Dolichol-PP is removed to yield Nacetylmuramic acid-N-acetylglucosaminedecapeptide Figure 6.60 - Catalytic activity of DDtranspeptidase Wikipedia Figure 6.59 - Penicillin

11. This last group is added to the growing peptidoglycan network by joining the pentaglycine of one chain to the tetrapeptide/pentapeptide of another.

The enzyme catalyzing the addition of the N-acetylmuramic acid-N-acetylglucosaminedecapeptide to the network in the last step is DD-transpeptidase. This is the cellular enzyme targeted by penicillin and its derivatives. One reason penicillin is so effective is because synthesis of a peptidoglycan cell wall for a single bacterium requires millions of cycles of reactions above. Even slowing down the process can have a major effect on bacterial growth. On the flip side, resistance to penicillin and derivatives arises as a result of mutations in one enzyme - the transpeptidase.

Metabolons

At this point, it is appropriate to bring up the concept of metabolons. Metabolons are cellular complexes containing multiple enzymes of a metabolic pathway that appear to be arranged so that the product of one enzymatic reaction is passed directly as substrate to the enzyme that catalyzes the next reaction in the metabolic pathway. The structural complexes are temporary and are held together by non-covalent forces.

Metabolons appear to offer advantages of reducing the amount of water needed to hydrate enzymes. Activity of enzymes in the complex is increased. Most of the basic metabolic pathways are thought to use metabolons. They include glycolysis, the citric acid cycle, nucleotide metabolism, glycogen synthesis, steroid synthesis, DNA synthesis, RNA synthesis, the urea cycle, and the process of electron transport.

Hypoxia

Hypoxia occurs when the body or a region of it has an insufficient oxygen supply. Varia- Figure 6.61 - Hypoxia inducible factor tions in arterial oxygen concentrations in normal physiology may lead to hypoxia, for example, during hypoventilation training or strenuous physical exercise. Generalized hypoxia may appear in healthy people when at high altitudes. Cancer cells, which may be undergoing faster respiration than surrounding tissues may also tend to be hypoxic. Hypoxia is an important consideration for sugar metabolism due to the ability of cells to change their sugar metabolism (fermentation) when these conditions exist.

The body's response to hypoxia is to produce Hypoxia-Inducible Factors (HIFs), which are transcription factors that induce expression of genes to help cells adapt to the hypoxic conditions. Many of the genes activated by HIFs are enzymes of glycolysis and GLUTs (glucose transport proteins). The combination of these gene products allows cells to 1) import more glucose and 2) metabolize it more rapidly when it arrives. This is to be expected because anaerobic sugar metabolism is only about 1/15th as efficient as aerobic metabolism. Consequently, it requires much more sugar metabolism to keep the cancer cells alive. A recently discovered protein called cytoglobin is believed to help assist in hypoxia by facilitating transfer of oxygen from arteries to the brain.

Covalent modification

HIFs are regulated partly by an interesting covalent modification. When oxygen concentration is high, the enzyme prolyl hydroxylase will hydroxylate proline residues in HIFs. This stimulates the protein degradation system (proteasome) to degrade them. When oxygen concentration is low, the hydroxylation occurs to a much lower extent (or does not occur at all), reducing/stopping degradation of HIFs and allowing them to activate genes. In this way, the concentration of HIFs is kept high under low oxygen concentration (to activate HIF genes) and low under high oxygen concentrations (to stop synthesis of HIF genes).

This page titled 6.1: Metabolism - Sugars is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6.2: Citric Acid Cycle & Related Pathways

Source: BiochemFFA_6_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Citric acid cycle

The primary catabolic pathway in the body is the citric acid cycle because it is here that oxidation to carbon dioxide occurs for breakdown products of the cell's major building blocks - sugars, fatty acids, and amino acids. The pathway is cyclic (Figure 6.63) and thus, doesn't really have a starting or ending point. All of the reactions occur in mitochondria, though one enzyme is embedded in the organelle's inner 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 (see HERE).

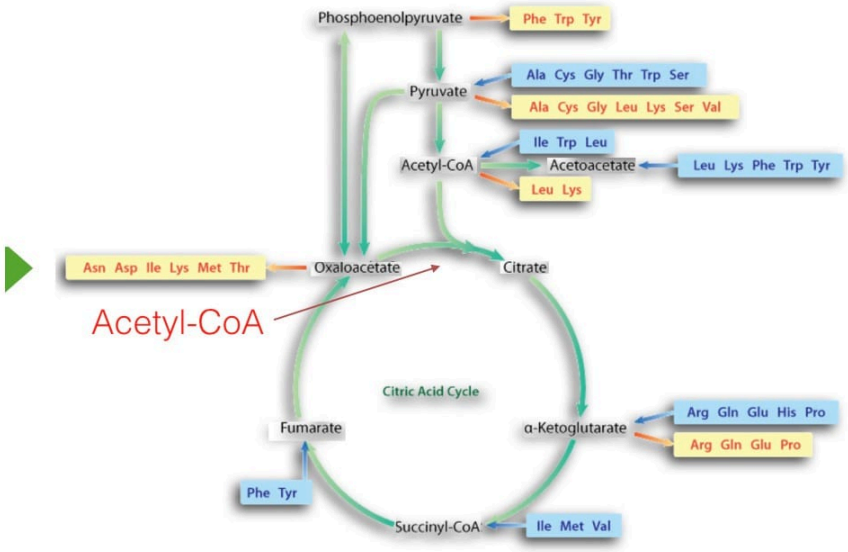


Figure 6.63 - Amino acid metabolism and the citric acid cycle. Amino acids boxed in yellow are made from the indicated intermediate. Amino acids in blue are made into the intermediate in catabolism. Image by Aleia Kim

Acetyl-CoA

The molecule "feeding" the citric acid cycle is acetyl-CoA and it can be obtained from pyruvate (from glycolysis), from fatty acid β-oxidation, from ketone bodies, and from amino acid metabolism. Molecules from other pathways feeding into the citric acid cycle for catabolism make the citric acid cycle 'cataplerotic'. It is worth noting that acetyl-CoA has very different fates, depending on the cell's energy status/needs (see HERE). The description below describes oxidation (catabolism) in citric acid cycle.

Anabolically, acetyl-CoA is also very important for providing building blocks for synthesis of fatty acids, ketone bodies, amino acids and cholesterol. Other citric acid cycle intermediates are also important in amino acid metabolism (Figure 6.63), heme synthesis, electron shuttling, and shuttling of acetyl-CoA across the mitochondrial inner membrane. The ability of the citric acid cycle to supply intermediates to pathways gives rise to the term 'anaplerotic.' It means 'to fill up.' Before discussing the citric acid cycle, it is important to first describe one important enzyme complex that is a major source of acetyl-CoA for the cycle.

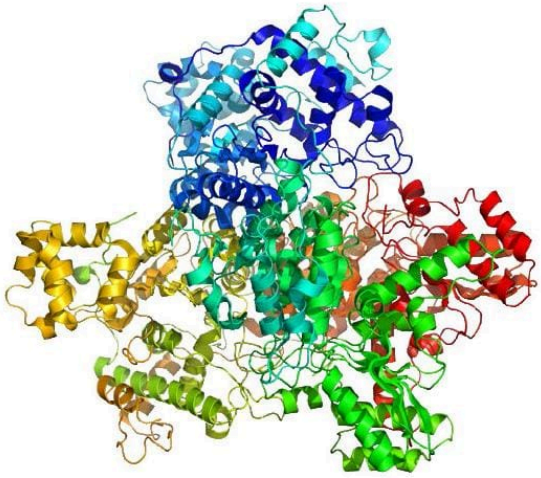


Figure 6.64 - E1 Subunit of Pyruvate Dehydrogenase. Wikipedia

Pyruvate decarboxylation

The pyruvate dehydrogenase enzyme is a complex of multiple copies of three subunits that catalyze the decarboxylation of pyruvate to form acetyl-CoA. The reaction mechanism requires use of five coenzymes. Pyruvate dehydrogenase is an enormous complex in mammals with a size five times greater than ribosomes.

Subunits

The three subunits are designated by E1, E2, and E3. E2 is also referred to as dihydrolipoamide acetyltransferase and E3 is more precisely called dihydrolipoyl dehydrogenase. Confusion arises with the name for E1. Some call it pyruvate dehydrogenase and others give it the name pyruvate decarboxylase. We will use pyruvate decarboxylase solely to refer to E1 and pyruvate dehydrogenase only to refer to the complex of E1, E2, and E3.

The catalytic actions of pyruvate dehydrogenase can be broken down into three steps, each taking place on one of the subunits. The steps, sequentially occurring on E1, E2, and E3, are 1) decarboxylation of pyruvate; 2) oxidation of the decarboxylated product; and 3) transfer of electrons to ultimately form NADH (Figure 6.65).

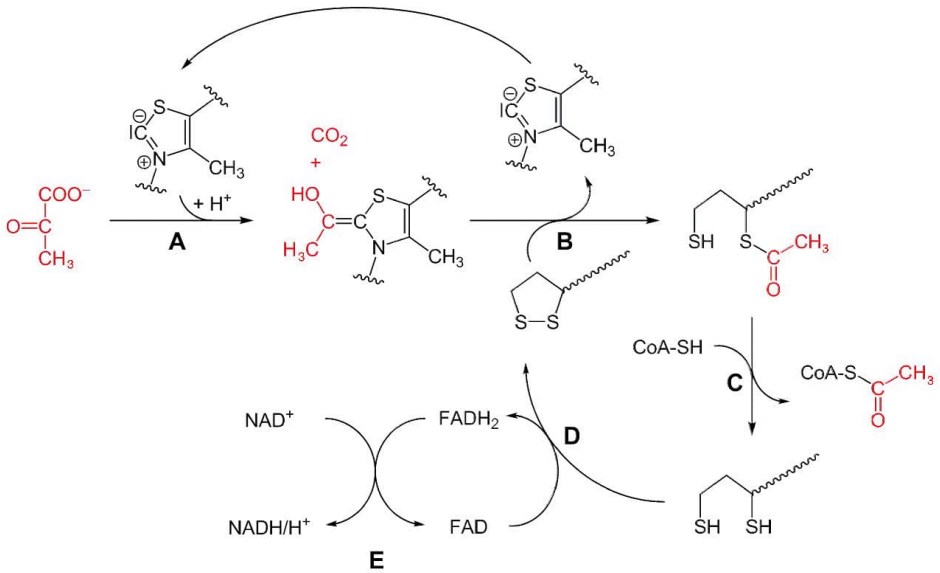


Figure 6.65 - Mechanism of action of pyruvate decarboxylation and oxidation by pyruvate dehydrogenase.

Catalysis

The catalytic process begins after binding of the pyruvate substrate with activation of the thiamine pyrophosphate coenzyme through formation of an ylide intermediate. The nucleophilic carbanion of the ylide attacks the electrophilic ketone carbon on the pyruvate, releasing carbon dioxide and creating an enol that loses a proton on the carbon to become a 1,3 dipole that includes the positively charged nitrogen of the thiamine. The reaction (step A in Figure 6.65) is a non-oxidative decarboxylation. Oxidation of the two carbon hydroxyethyl unit occurs in the transfer to the lipoamide.

Reductive acetylation

Reductive acetylation occurs next (Step B) as the 2-carbon hydroxyethyl unit is transferred to lipoamide on E2. (Lipoamide is the name for a molecule of lipoic acid covalently attached to a lysine side chain in the E2 subunit). In prokaryotes in the absence of oxygen, the hydroxyethyl group is not passed to lipoamide, but instead is released as free acetaldehyde , which can accept electrons from NADH (catalyzed by alcohol dehydrogenase) and become ethanol in the process of fermentation. In the presence of oxygen in almost all aerobic organisms, the process continues with transfer of the hydroxyethyl unit to E2 and continuation of the cycle below.

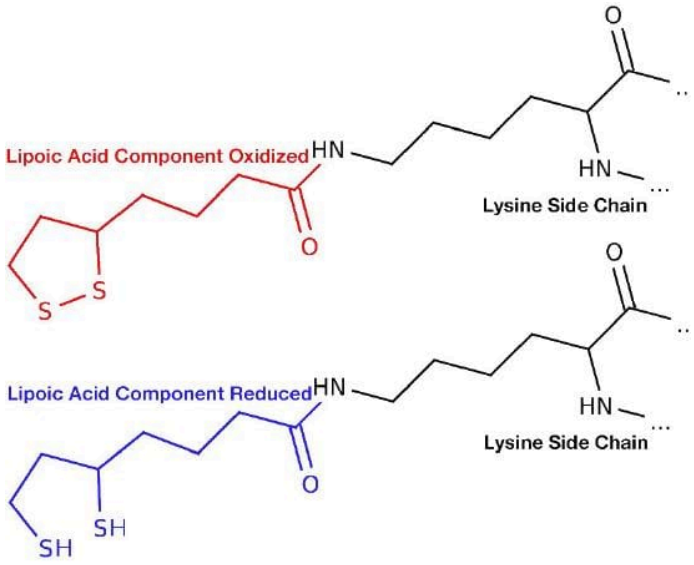


Figure 6.66 - Oxidized and reduced structures of lipoamide (lipoic acid linked to lysine)

Oxidation step

Transfer of the hydroxyethyl group from E1 to the lipoamide coenzyme in E2 is an oxidation, with transfer of electrons from the hydroxyethyl group to lipoamide's disulfide (reducing it) and formation on the lipoamide of an acetyl-thioester (oxidizing it).

The acetyl group is then transferred from lipoamide to coenzyme A in E2 (Step C in Figure 6.65), forming acetyl-CoA, which is released and leaving reduced sulfhydryls on the lipoamide. In order for the enzyme to return to its original state, the disulfide bond on lipoamide must be re-formed. This occurs with transfer of electrons from reduced lipoamide to an FAD covalently bound to E3 (Step D). This reduces FAD to FADH2.

Formation of NADH

In the last step in the process, electrons from FADH2 are transferred to external NAD+, forming NADH (Step E) and completing the overall cycle. Then enzyme can then begin another catalytic round by binding to a pyruvate.

Pyruvate dehydrogenase regulation

Pyruvate deyhdrogenase is regulated both allosterically and by covalent modification - phosphorylation / dephosphorylation. Regulation of pyruvate dehydrogenase, whether by allosteric or covalent mechanisms has the same strategy. Indicators of high energy shut down the enzyme, whereas indicators of low energy stimulate it. For allosteric regulation, the high energy indicators affecting the enzyme are ATP, acetyl-CoA, NADH, and fatty acids, which inhibit it. AMP, Coenzyme A, NAD+, and calcium, on the other hand, stimulate it (Figure 6.67).

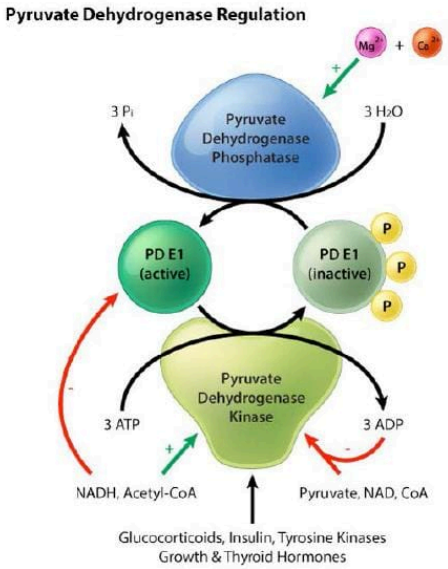


Figure 6.67 - Regulation scheme for pyruvate dehydrogenase (PD). Image by Aleia Kim

Covalent modification

Covalent modification regulation of pyruvate dehydrogenase is a bit more complicated. It occurs as a result of phosphorylation by pyruvate dehydrogenase kinase (PDK - Figure 6.67) or dephosphorylation by pyruvate dehydrogenase phosphatase (PDP).

PDK puts phosphate on any one of three serine residues on the E1 subunit, which causes pyruvate kinase to not be able to perform its first step of catalysis - the decarboxylation of pyruvate. PDP can remove those phosphates. PDK is allosterically activated in the mitochondrial matrix when NADH and acetyl-CoA concentrations rise.

Product inhibition

Thus, the products of the pyruvate dehydrogenase reaction inhibit the production of more products by favoring its phosphorylation by PDK. Pyruvate, a substrate of pyruvate dehydrogenase, inhibits PDK, so increasing concentrations of substrate activate pyruvate dehydrogenase by reducing its phosphorylation by PDK. As concentrations of NADH and acetyl-CoA fall, PDP associates with pyruvate kinase and removes the phosphate on the serine on the E1 subunit.

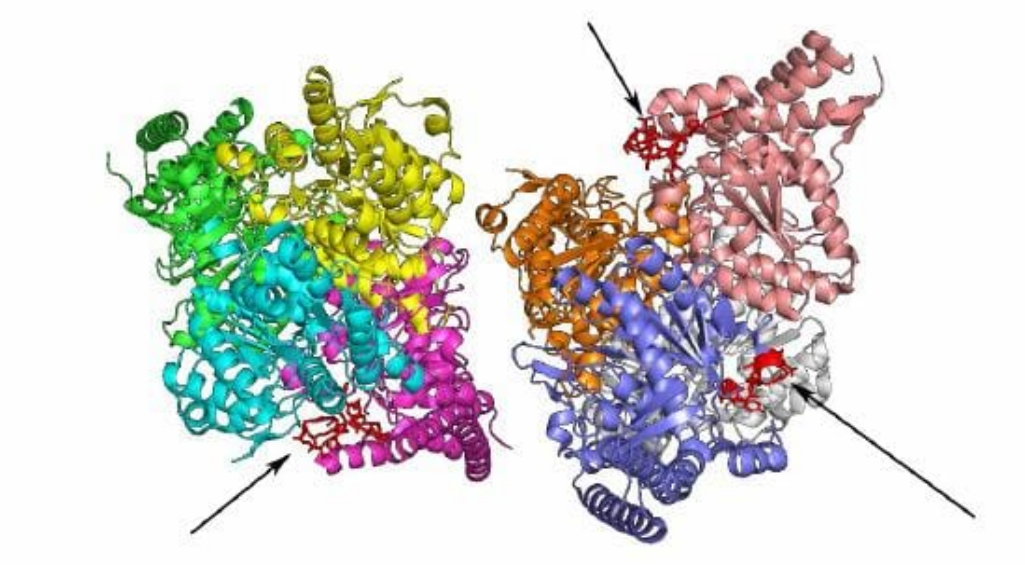


Figure 6.68 - Pyruvate dehydrogenase complex with three phosphorylation sites in red marked by arrows.Wikipedia

Low concentrations of NADH and acetyl-CoA are necessary for PDP to remain on the enzyme. When those concentrations rise, PDP dissociates and PDK gains access to the serine for phosphorylation. Insulin and calcium can also activate the PDP. This is very important in muscle tissue, since calcium is a signal for muscular contraction, which requires energy. Insulin also also activates pyruvate kinase and the glycolysis pathway to use internalized glucose. It should be noted that the cAMP phosphorylation cascade from the β-adrenergic receptor has no effect on pyruvate kinase, though the insulin cascade does, in fact, affect PDP and pyruvate kinase.

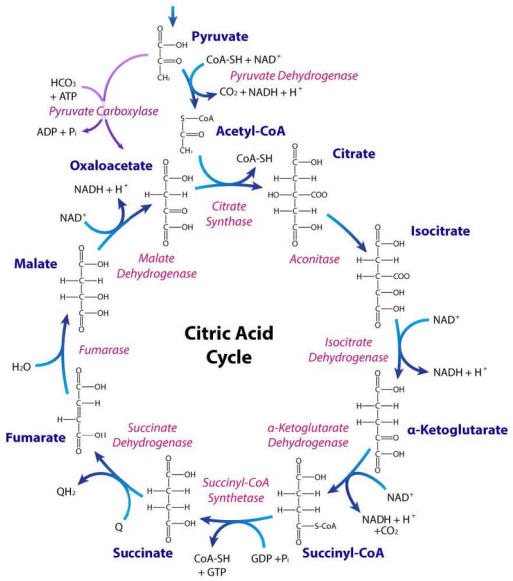


Figure 6.69 - The citric acid cycle. Image by Aleia Kim

Citric acid cycle reactions

Focusing on the pathway itself (Figure 6.69), the usual point to start discussion is addition of acetyl-CoA to oxaloacetate (OAA) to form citrate.

Acetyl-CoA for the pathway can come from a variety of sources. The reaction joining it to OAA is catalyzed by citrate synthase and the ΔG°' is fairly negative. This, in turn, helps to "pull" the malate dehydrogenase reaction preceding it in the cycle.

In the next reaction, citrate is isomerized to isocitrate by action of the enzyme called aconitase.

Isocitrate is a branch point in plants and bacteria for the glyoxylate cycle (see HERE). Oxidative decarboxylation of isocitrate by isocitrate dehydrogenase produces the first NADH and yields α-ketoglutarate.

This five carbon intermediate is a branch point for synthesis of the amino acid glutamate. In addition, glutamate can also be made easily into this intermediate in the reverse reaction. Decarboxylation of α-ketoglutarate produces succinyl-CoA and is catalyzed by α-ketoglutarate dehydrogenase.

The enzyme α-ketoglutarate dehydrogenase is structurally very similar to pyruvate dehydrogenase and employs the same five coenzymes – NAD+, FAD, CoA-SH, thiamine pyrophosphate, and lipoamide.

Regeneration of oxaloacetate

The remainder of the citric acid cycle involves conversion of the four carbon succinyl-CoA into oxaloacetate. Succinyl-CoA is a branch point for the synthesis of heme (see HERE). Succinyl-CoA is converted to succinate in a reaction catalyzed by succinyl-CoA synthetase (named for the reverse reaction) and a GTP is produced, as well – the only substrate level phosphorylation in the cycle.

The energy for the synthesis of the GTP comes from hydrolysis of the high energy thioester bond between succinate and the CoA-SH. Evidence for the high energy of a thioester bond is also evident in the citrate synthase reaction, which is also very energetically favorable. Succinate is also produced by metabolism of odd-chain fatty acids (see HERE).

Succinate Oxidation

Oxidation of succinate occurs in the next step, catalyzed by succinate dehydrogenase. This interesting enzyme both catalyzes this reaction and participates in the electron transport system, funneling electrons from the FADH2 it gains in the reaction to coenzyme Q. The product of the reaction, fumarate, gains a water across its trans double bond in the next reaction, catalyzed by fumarase to form malate.

Fumarate is also a byproduct of nucleotide metabolism and of the urea cycle. Malate is important also for transporting electrons across membranes in the malate-aspartate shuttle (see HERE) and in ferrying carbon dioxide from mesophyll cells to bundle sheath cells in C4 plants (see HERE).

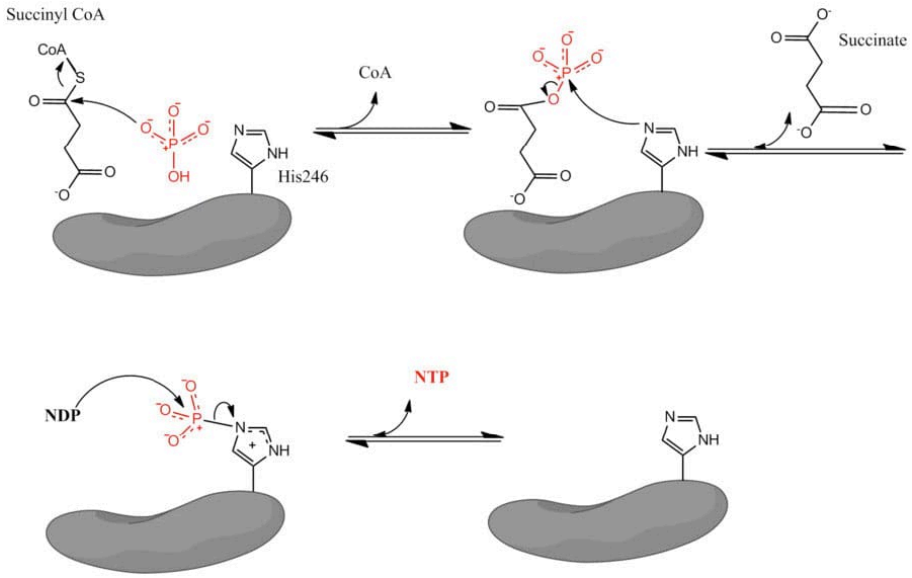


Figure 6.70 - Succinyl-CoA synthetase mechanism

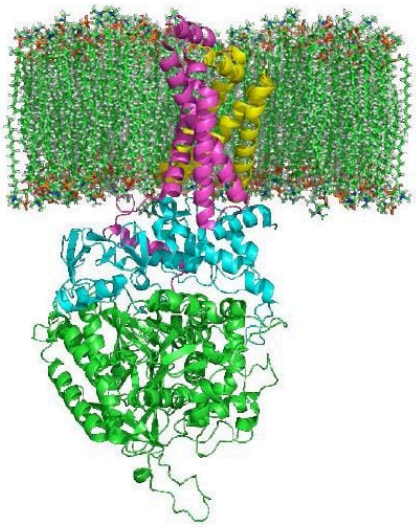


Figure 6.71 - Succinate dehydrogenase embedded in the mitochondrial inner membrane (top). Wikipedia

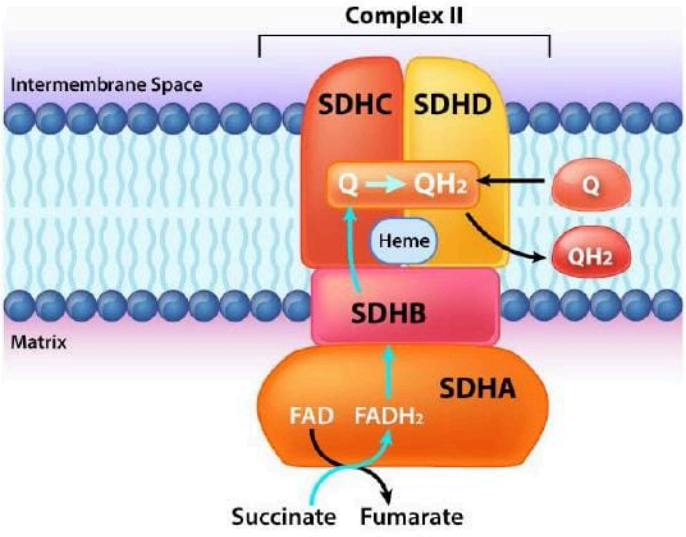


Figure 6.72 - Succinate dehydrogenase reaction. Image by Aleia Kim

Rare oxidation

Conversion of malate to oxaloacetate by malate dehydrogenase is a rare biological oxidation that has a $\Delta G^{\circ\prime}$ with a positive value (29.7 kJ/mol).

The reaction is 'pulled' by the energetically favorable conversion of oxaloacetate to citrate in the citrate synthase reaction described above. Oxaloacetate intersects other important pathways, including amino acid metabolism (readily converted to aspartic acid), transamination (nitrogen movement) and gluconeogenesis.

It is worth noting that reversal of the citric acid cycle theoretically provides a mechanism for assimilating CO2. In fact, this reversal has been noted in both anaerobic and microaerobic bacteria, where it is called the Arnon-Buchanan cycle (Figure 6.73).

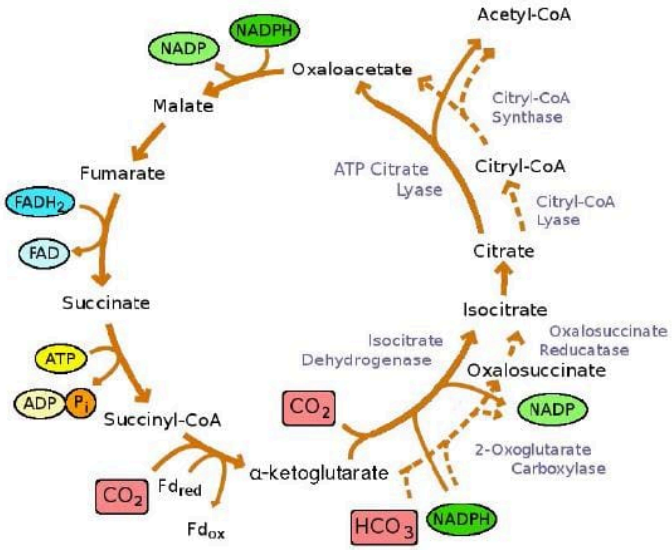


Figure 6.73 - Arnon-Buchanon cycle. Alternative enzymes shown on right in lavender. Fd = ferredoxin. Wikipedia

Regulation of the citric acid cycle

Allosteric regulation of the citric acid cycle is pretty straightforward. The molecules involved are all substrates/products of the pathway or molecules involved in energy transfer. Substrates/products that regulate or affect the pathway include acetyl-CoA and succinyl-CoA .

Inhibitors and activators

High energy molecular indicators, such as ATP and NADH will tend to inhibit the cycle and low energy indicators (NAD+, AMP, and ADP) will tend to activate the cycle. Pyruvate dehydrogenase, which catalyzes formation of acetyl-CoA for entry into the cycle is allosterically inhibited by its product (acetyl-CoA), as well as by NADH and ATP.

Regulated enzymes

Regulated enzymes in the cycle include citrate synthase (inhibited by NADH, ATP, and succinyl-CoA), isocitrate dehydrogenase (inhibited by ATP, activated by ADP and NAD+), and α-ketoglutarate dehydrogenase (inhibited by NADH and succinyl-CoA and activated by AMP).

Anaplerotic/cataplerotic pathway

The citric acid cycle is an important catabolic pathway oxidizing acetyl-CoA into CO2 and generating ATP, but it is also an important source of molecules needed by cells and a mechanism for extracting energy from amino acids in protein breakdown and other breakdown products. This ability of the citric acid cycle to supply molecules as needed and to absorb metabolic byproducts gives great flexibility to cells. When citric acid cycle intermediates are taken from the pathway to make other molecules, the term used to describe this is cataplerotic, whereas when molecules are added to the pathway, the process is described as anaplerotic.

Cataplerotic molecules

The citric acid cycle's primary cataplerotic molecules include α-ketoglutarate, succinyl-CoA, and oxaloacetate. Transamination of α-ketoglutarate and oxaloacetate produces the amino acids glutamate and aspartic acid, respectively. Oxaloacetate is important for the production of glucose in gluconeogenesis.

Glutamate plays a very important role in the movement of nitrogen through cells via glutamine and other molecules and is also needed for purine synthesis. Aspartate is a precursor of other amino acids and for production of pyrimidine nucleotides. Succinyl-CoA is necessary for the synthesis of porphyrins, such as the heme groups in hemoglobin, myoglobin and cytochromes.

Citrate is an important source of acetyl-CoA for making fatty acids. When the citrate concentration is high (as when the citric acid cycle is moving slowly or is stopped), it gets shuttled across the mitochondrial membrane into the cytoplasm and broken down by the enzyme citrate lyase to oxaloacetate and acetyl-CoA. The latter is a precursor for fatty acid synthesis in the cytoplasm.

Anaplerotic molecules

Anaplerotic molecules replenishing citric acid cycle intermediates include acetyl-CoA (made in many pathways, including fatty acid oxidation, pyruvate decarboxylation, amino acid catabolism, and breakdown of ketone bodies), α-ketoglutarate (from amino acid metabolism), succinyl-CoA (from propionic acid metabolism), fumarate (from the urea cycle and purine metabolism), malate (carboxylation of PEP in plants), and oxaloacetate (many sources, including amino acid catabolism and pyruvate carboxylase action on pyruvate in gluconeogenesis)

Glyoxylate cycle

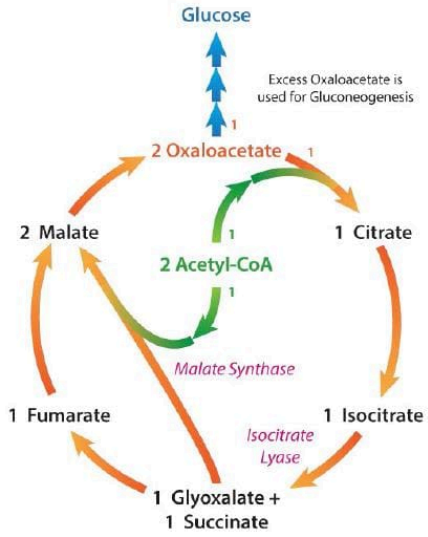


Figure 6.74 - Overview of the glyoxylate cycle. Image by Aleia Kim

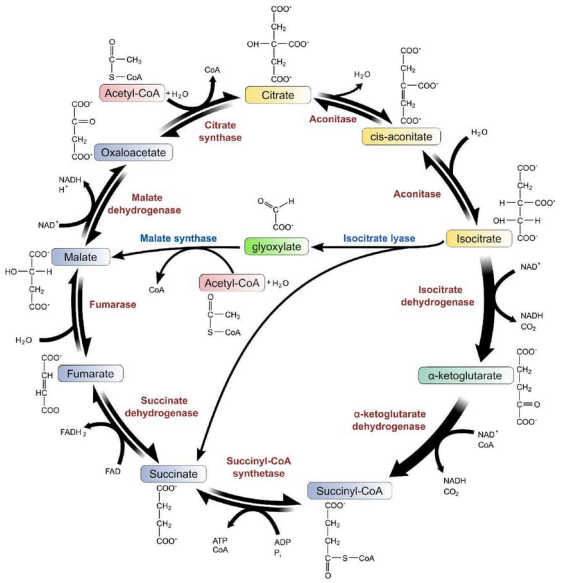


Figure 6.75 - Reactions of the glyoxylate cycle. Wikipedia

A pathway related to the citric acid cycle found only in plants and bacteria is the glyoxylate cycle (Figures 6.74 & 6.75). The glyoxylate cycle, which bypasses the decarboxylation reactions while using most of the non-decarboxylation reactions of the citric acid cycle, does not operate in animals, because they lack two enzymes necessary for it – isocitrate lyase and malate synthase. The cycle occurs in specialized plant peroxisomes called glyoxysomes. Isocitrate lyase catalyzes the conversion of isocitrate into succinate and glyoxylate. Because of this, all six carbons of the citric acid cycle survive each turn of the cycle and do not end up as carbon dioxide.

Succinate continues through the remaining reactions to produce oxaloacetate. Glyoxylate combines with another acetyl-CoA (one acetyl-CoA was used to start the cycle) to create malate (catalyzed by malate synthase). Malate can, in turn, be oxidized to oxaloacetate.

It is at this point that the glyoxylate pathway's contrast with the citric acid cycle is apparent. After one turn of the citric acid cycle, a single oxaloacetate is produced and it balances the single one used in the first reaction of the cycle. Thus, in the citric acid cycle, there is no net production of oxaloacetate in each turn of the cycle.

Net oxaloacetate production

On the other hand, thanks to assimilation of carbons from two acetyl-CoA molecules, each turn of the glyoxylate cycle results in two oxaloacetates being produced, after starting with one. The extra oxaloacetate of the glyoxylate cycle can be used to make other molecules, including glucose in gluconeogenesis. This is particularly important for plant seed germination (Figure 6.76), since the seedling is not exposed to sunlight. With the glyoxylate cycle, seeds can make glucose from stored lipids.

Because animals do not run the glyoxylate cycle, they cannot produce glucose from acetyl-CoA in net amounts, but plants and bacteria can. As a result, plants and bacteria can turn acetyl-CoA from fat into glucose, while animals can't. Bypassing the oxidative decarboxylations (and substrate level phosphorylation) has energy costs, but, there are also benefits. Each turn of the glyoxylate cycle produces one FADH2 and one NADH instead of the three NADHs, one FADH2, and one GTP made in each turn of the citric acid cycle.



Figure 6.76 - A gingko seed embryo. Wikipedia

Carbohydrate needs

Organisms that make cell walls, such as plants, fungi, and bacteria, need large quantities of carbohydrates as they grow to support the biosynthesis of the complex structural polysaccharides of the walls. These include cellulose, glucans, and chitin. Notably, each of the organisms can operate the glyoxylate cycle using acetyl-CoA from β-oxidation.

Coordination of the glyoxylate cycle and the citric acid cycle

The citric acid cycle is a major catabolic pathway producing a considerable amount of energy for cells, whereas the glyoxylate cycle's main function is anabolic - to allow production of glucose from fatty acids in plants and bacteria. The two pathways are physically separated from each other (glyoxylate cycle in glyoxysomes / citric acid cycle in mitochondria), but nonetheless a coordinated regulation of them is important.

The enzyme that appears to provide controls for the cycle is isocitrate dehydrogenase. In plants and bacteria, the enzyme can be inactivated by phosphorylation by a kinase found only in those cells. Inactivation causes isocitrate to accumulate in the mitochondrion and when this happens, it gets shunted to the glyoxysomes, favoring the glyoxylate cycle. Removal of the phosphate from isocitrate dehydrogenase is catalyzed by an isocitrate dehydrognease-specific phosphoprotein phosphatase and restores activity to the enzyme.

When this happens, isocitrate oxidation resumes in the mitochondrion along with the rest of the citric acid cycle reactions. In bacteria, where the enzymes for both cycles are present together in the cytoplasm, accumulation of citric acid cycle intermediates and glycolysis intermediates will tend to favor the citric acid cycle by activating the phosphatase, whereas high energy conditions will tend to favor the glyoxylate cycle by inhibiting it.

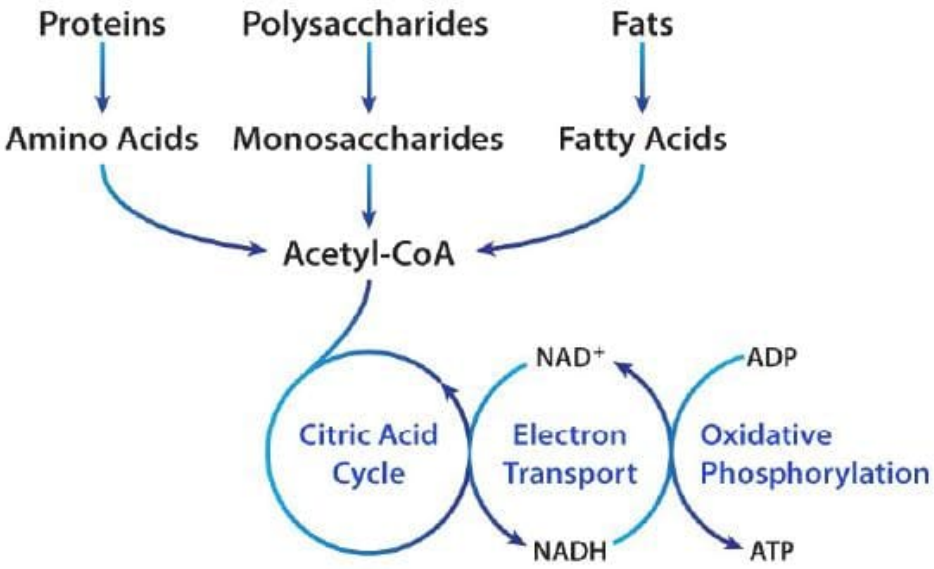


Figure 6.77 - Acetyl-CoA metabolism. Image by Aleia Kim

Acetyl-CoA metabolism

Acetyl-CoA is one of the most “connected” metabolites in biochemistry, appearing in fatty acid oxidation/synthesis, pyruvate oxidation, the citric acid cycle, amino acid anabolism/catabolism, ketone body metabolism, steroid/bile acid synthesis, and (by extension from fatty acid metabolism) prostaglandin synthesis . Most of these pathways will be dealt with separately. Here we will cover ketone body metabolism.

Ketone body metabolism

Ketone bodies are molecules made when the blood levels of glucose fall very low. Ketone bodies can be converted to acetyl-CoA by reversing the reaction of the pathway that makes them (Figure 6.78). Acetyl CoA, of course, can be used for ATP synthesis via the citric acid cycle. People who are very hypoglycemic (including some diabetics) will produce ketone bodies (Figure 6.79) and these are often first detected by the smell of acetone on their breath.

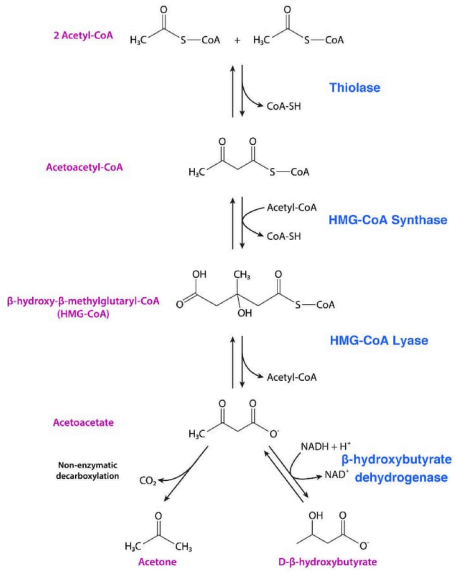


Figure 6.78 Ketone body metabolism. Image by Pehr Jacobson

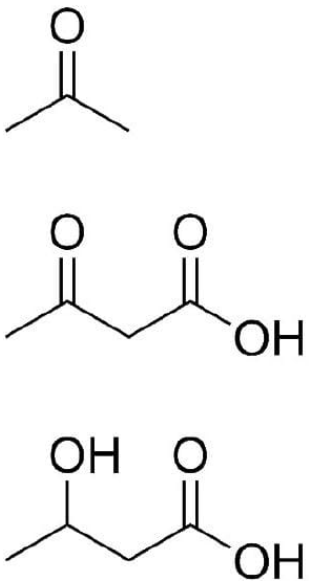


Figure 6.79 - Three ketone bodies - acetone (top), acetoacetic acid (middle), and β-hydroxybutyrate (bottom)

Overlapping pathways

The pathways for ketone body synthesis and cholesterol biosynthesis (Figure 6.80 and see HERE) overlap at the beginning. Each of these starts by combining two acetyl-CoAs together to make acetoacetyl-CoA. Not coincidentally, that is the next to last product of β-oxidation of fatty acids with even numbers of carbons (see HERE for fatty acid oxidation). In fact, the enzyme that catalyzes the joining is the same as the one that catalyzes its breakage in fatty acid oxidation – thiolase. Thus, these pathways start by reversing the last step of the last round of fatty acid oxidation.

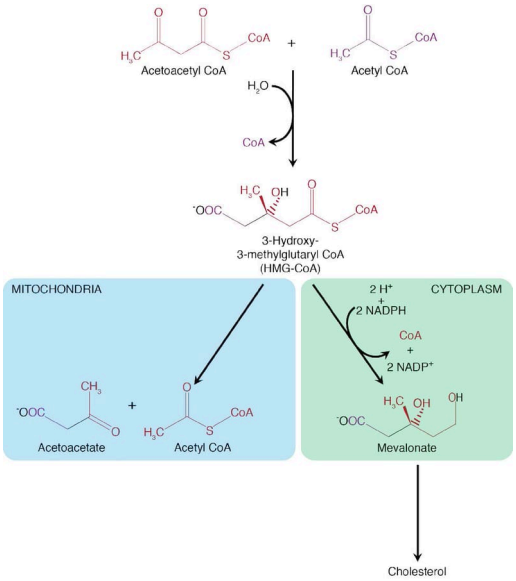


Figure 6.80 - Diverging biosynthetic pathways for ketone bodies (left) and cholesterol biosynthesis (right). Image by Penelope Irving

HMG-CoA formation

Both pathways also include addition of two more carbons to acetoacetyl-CoA from a third acetyl-CoA to form hydroxy-methyl-glutaryl-CoA, or HMG-CoA, as it is more commonly known. It is at this point that the two pathways diverge. HMG-CoA is a branch point between the two pathway and can either go on to become cholesterol or ketone bodies. In the latter pathway, HMG-CoA is broken down into acetyl-CoA and acetoacetate.

Acetoacetate is itself a ketone body and can be reduced to form another one, D-β-hydroxybutyrate (not actually a ketone, though). Alternatively, acetoacetate can be converted into acetone. This latter reaction can occur either spontaneously or via catalysis by acetoacetate decarboxylase. Acetone can be converted into pyruvate and pyruvate can be made into glucose.

D-β-hydroxybutyrate travels readily in the blood and crosses the blood-brain barrier. It can be oxidized back to acetoacetate, converted to acetoacetyl-CoA, and then broken down to two molecules of acetyl-CoA for oxidation in the citric acid cycle.

Ketosis

When a body is producing ketone bodies for its energy, this state in the body is known as ketosis. Formation of ketone bodies in the liver is critical. Normally glucose is the body's primary energy source. It comes from the diet, from the breakdown of storage carbohydrates, such as glycogen, or from glucose synthesis (gluconeogenesis). Since the primary stores of glycogen are in muscles and liver and since gluconeogenesis occurs only in liver, kidney, and gametes, when the supply of glucose is interrupted for any reason, the liver must supply an alternate energy source.

From fatty acid breakdown

In contrast to glucose, ketone bodies can be made in animals from the breakdown of fat/fatty acids. Most cells of the body can use ketone bodies as energy sources. Ketosis may arise from fasting, a very low carbohydrate diet or, in some cases, diabetes.

Acidosis

The term acidosis refers to conditions in the body where the pH of arterial blood drops below 7.35. It is the opposite of the condition of alkalosis, where the pH of the arterial blood rises above 7.45. Normally, the pH of the blood stays in this narrow pH range. pH values of the blood lower than 6.8 or higher than 7.8 can cause irreversible damage and may be fatal. Acidosis may have roots in metabolism (metabolic acidosis) or in respiration (respiratory acidosis).

There are several causes of acidosis. In metabolic acidosis, production of excess lactic acid or failure of the kidneys to excrete acid can cause blood pH to drop. Lactic acid is produced in the body when oxygen is limiting, so anything that interferes with oxygen delivery may create conditions favoring production of excess lactic acid. These may include restrictions in the movement of blood to target tissues, resulting in hypoxia (low oxygen conditions) or decreases in blood volume. Issues with blood movement can result from heart problems, low blood pressure, or hemorrhaging.

Strenuous exercise can also result in production of lactic acid due to the inability of the blood supply to deliver oxygen as fast as tissues require it (hypovolemic shock). At the end of the exercise, though, the oxygen supply via the blood system quickly catches up.

Respiratory acidosis arises from accumulation of carbon dioxide in the blood. Causes include hypoventilation, pulmonary problems, emphysema, asthma, and severe pneumonia.

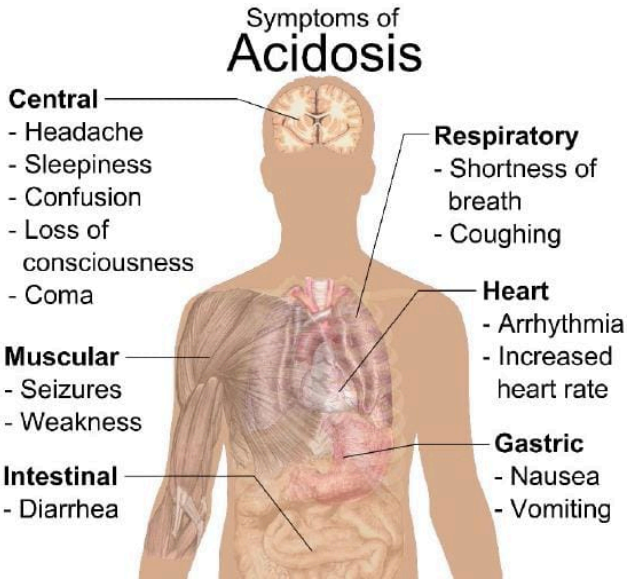


Figure 6.81 - Symptoms of acidosis

This page titled 6.2: Citric Acid Cycle & Related Pathways is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6.3: Fats and Fatty Acids

Source: BiochemFFA_6_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

In the modern Western world, which is fat and getting fatter, there is a tremendous amount of interest in the metabolism of fat and fatty acids. Fat is the most important energy storage form of animals, storing considerably more energy per carbon than carbohydrates, but its insolubility in water requires the body to package it specially for transport. Surprisingly, fat/fatty acid metabolism is not nearly as tightly regulated as that of carbohydrates. Neither are the metabolic pathways of breakdown and synthesis particularly complicated, either.

Movement of dietary fat

Before we discuss the breakdown and synthesis of fat, let us first discuss the movement of dietary fat and oil (triglycerides - Figure 6.82) in the body. Upon consumption of triglycerides in the diet, they first are solubilized in the digestive system by the churning action of the stomach and the emulsifying properties of the bile acids.

Upon passing into the lumen of the intestine, the triglycerides are acted on first by enzymes known as lipases that use water twice on each triglyceride to release two fatty acids, leaving behind a monoacylglyceride. As shown in Figure 6.83, the fatty acids and the monoacylglyceride are moved across the intestinal wall into the lymph system where they are reassembled back into a triglyceride. In the lymph system triglycerides and other insoluble lipids are packaged into lipoprotein complexes called chylomicrons that enter the blood stream and travel to target cells. The journey of lipids in the body after leaving the digestive system is long and is discussed in more depth HERE.

In the body, fat is stored in specialized cells known as adipocytes. When these cells receive appropriate signals, they begin the breakdown of fat into glycerol and fatty acids.

Breakdown of fat

Breakdown of fat in adipocytes requires catalytic action of three enzymes. The first of these is controlled by binding of hormones to the cell membrane (Figure 6.84). It is the only regulated enzyme of fat breakdown and is known as hormone sensitive triacylglycerol lipase. It removes the first fatty acid from the fat. Diacylglyceride lipase removes the second one and monoacylglyceride lipase removes the third. As noted, only the first one is regulated and it appears to be the rate limiting reaction when active.

Epinephrine activation

As shown in Figure 6.84, activation of hormone sensitive triacylglycerol lipase (HSTL) is accomplished by epinephrine stimulation process and that it overlaps with the same activation that stimulates glycogen breakdown and gluconeogenesis.

This coordination is very important. Each of the pathways stimulated by the epinephrine signaling system aims to provide the body with more materials to catabolize for energy - sugars and fatty acids. The HSTL is inhibited by dephospohrylation and this is stimulated by binding of insulin to its cell membrane receptor.

Perilipin

A protein playing an important roles in regulation of fat breakdown is perilipin. Perilipin associates with fat droplets and helps regulate action of HSTL, the enzyme catalyzing the first reaction in fat catabolism. When perilipin is not phosphorylated, it coats the fat droplet and prevents HSTL from getting access to it. Activation of protein kinase A in the epinephrine cascade, however, results in phosphorylation of both perilipin and HSTL. When this occurs, perilipin loosens its tight binding to the fat droplet, allowing digestion of the fat to begin by HSTL.

Perilipin expression is high in obese organisms and some mutational variants have been associated with obesity in women. Another mutation reduces perilipin expression and is associated with greater lipolysis (fat breakdown) in women. Mice lacking perilipin eat more food than wild-type mice, but gain 1/3 less weight when on the same diet.

Fat synthesis

Synthesis of fat requires action of acyl transferase enzymes, such as glycerol-3 O-phosphate acyl transferase, which catalyzes addition of fatty acids to the glycerol backbone (reaction #1 above). The process requires glycerol-3-phosphate (or DHAP) and three fatty acids. In the first reaction, glycerol-3-phosphate is esterified at position 1 with a fatty acid, followed by a duplicate reaction at position 2 to make phosphatidic acid (diacylglycerol phosphate). This molecule, which is an intermediate in the synthesis of both fats and phosphoglycerides, gets dephosphorylated to form diacylglycerol before the esterification of the third fatty acid to the molecule to make a fat.

Fatty acids released from adipocytes travel in the bloodstream bound to serum albumin. Arriving at target cells, fatty acids are taken up by membrane-associated fatty acid binding proteins, which help control cellular fatty acid uptake by transport proteins. Players in this process include CD36, plasma membrane-associated fatty acid-binding protein, and a family of fatty acid transport proteins (called FATP1-6).

Fatty acid oxidation

Upon arrival inside of target cells, fatty acids are oxidized in a process that chops off two carbons at a time to make acetyl-CoA, which is subsequently oxidized in the citric acid cycle. Depending on the size of the fatty acid, this process (called β-oxidation) will begin in either the mitochondrion (Figure 6.86) or the peroxisomes (see HERE).

Transport

To be oxidized in the mitochondrion, fatty acids must first be attached to coenzyme A (CoA-SH or CoA) and transported through the cytoplasm and the outer mitochondrial membrane. In the mitochondrion's intermembrane space, the CoA on the fatty acid is replaced by a carnitine (Figure 6.87) in order to be moved into the matrix. After this is done, the fatty acid linked to carnitine is transported into the mitochondrial matrix and in the matrix the carnitine is replaced again by coenzyme A. It is in the mitochondrial matrix where the oxidation occurs. The fatty acid linked to CoA (called an acyl-CoA) is the substrate for fatty acid oxidation.

Steps

The process of fatty acid oxidation (Figure 6.88) is fairly simple. The reactions all occur between carbons 2 and 3 (with #1 being the one linked to the CoA) and sequentially include the following steps 1) dehydrogenation to create FADH2 and a fatty acyl group with a double bond between carbons 2 and 3 in the trans configuration; 2) hydration across the double bond to put a hydroxyl group on carbon 3 in the L configuration; 3) oxidation of the hydroxyl group to make a ketone; and 4) thiolytic cleavage to release acetyl-CoA and a fatty acid two carbons shorter than the starting one.

Enzymes of β-oxidation

Two of the enzymes of β-oxidation are notable. The first is acyl-CoA dehydrogenase, which catalyzes the dehydrogenation in the first reaction and yields FADH2. The enzyme comes in three different forms – ones specific for long, medium, or short chain length fatty acids. The first of these is sequestered in the peroxisomes of animals (see below) whereas the ones that work on medium and shorter chain fatty acids are found in the mitochondria. Action of all three enzymes is typically needed to oxidize a fatty acid. Plants and yeast perform β-oxidation exclusively in peroxisomes.

The most interesting of the acyl-CoA dehydrogenases is the one that works on medium length fatty acids. This one, which is the one most commonly deficient in animals, has been associated with sudden infant death syndrome. Reactions two and three in β-oxidation are catalyzed by enoyl-CoA hydratase and 3-hydroxyacyl-CoA dehydrogenase, respectively. The latter reaction yields an NADH.

Thiolase

The second notable enzyme of β-oxidation is thiolase because this enzyme not only catalyzes the formation of acetyl-CoAs in β-oxidation, but also the joining of two acetyl-CoAs (essentially the reversal of the last step of β-oxidation) to form acetoacetyl-CoA – essential for the pathways of ketone body synthesis and cholesterol biosynthesis.

Similarity to citric acid cycle oxidation

It is worth noting that oxidation of fatty acids is chemically very similar to oxidation of the four carbon compounds of the citric acid cycle (Figure 6.89). In fatty acid oxidation, dehydrogenation between carbons 2 and 3 generates electrons which are donated to FAD to make FADH2 and a trans-bonded intermediate is formed.

The same thing happens in the citric acid cycle reaction catalyzed by succinate dehydrogenase - the trans-bonded molecule is fumarate. Addition of water in the second step of fatty acid oxidation occurs also in the next step of the citric acid cycle catalyzed by fumarase to create malate. Oxidation of the hydroxyl on carbon 3 in β-oxidation is repeated in the citric acid cycle reaction catalyzed by malate dehydrogenase yielding oxaloacetate.

Oxidation of odd chain fatty acids

Though most fatty acids of biological origin have even numbers of carbons, not all of them do. Oxidation of fatty acids with odd numbers of carbons ultimately produces an intermediate with three carbons called propionyl-CoA, which cannot be oxidized further in the β-oxidation pathway.

Metabolism of this intermediate is odd. Sequentially, the following steps occur (Figure 6.90) – 1) carboxylation to make D-methylmalonyl-CoA; 2) isomerization to L-methylmalonyl-CoA; 3) rearrangement to form succinyl-CoA. The last step of the process utilizes the enzyme methylmalonyl-CoA mutase, which uses the B12 coenzyme in its catalytic cycle. Succinyl-CoA can be metabolized in the citric acid cycle.

Peroxisomal oxidation

Long chain fatty acids (typically 22 carbons or more - Figure 6.91) have their oxidation initiated in the peroxisomes, due to the localization of the long acyl-CoA dehydrogenase in that organelle. Peroxisomal fatty acid oxidation is chemically similar to β-oxidation of mitochondria, but there are some differences in the overall process.

Differences

First, since there is no electron transport system in peroxisomes, the reduced electron carriers produced in oxidation there must have their own recycling process. Peroxisomes accomplish this by transferring electrons and protons from FADH2 to O2 to form hydrogen peroxide (H2O2). As a result of this, the lack of electron transport means no proton pump and, consequently, no ATP produced from FADH2 for peroxisomal fatty acid oxidation, making it less efficient that mitochondrial β-oxidation.

Electrons from NADH produced in the third step of the fatty acid oxidation must be shuttled to the cytoplasm and ultimately to the mitochondrion for ATP generation. Peroxisomal oxidation is increased for individuals on a high fat diet. In addition to long chain fatty acids, peroxisomes are also involved in oxidation of branched chain fatty acids, leukotrienes, and some prostaglandins.

Unsaturated fatty acid oxidation

Unsaturated fatty acids complicate the oxidation process a bit (see below), primarily because they have cis bonds, for the most part, if they are of biological origin, and these must be converted to the relevant trans intermediates for β-oxidation. Sometimes the bond must be moved down the chain, as well, in order to be positioned properly. Two enzymes (described below) handle all the necessary isomerizations and moves necessary to oxidize all of the unsaturated fatty acids (Figure 6.92).

Extra enzymes

As noted above, oxidation of unsaturated fatty acids requires two additional enzymes to the complement of enzymes for β-oxidation. If the β-oxidation of the fatty acid produces an intermediate with a cis bond between carbons three and four, cis-Δ3-enoyl-CoA isomerase will convert the bond to a trans bond between carbons two and three and β-oxidation can proceed as normal.

On the other hand, if β-oxidation produces an intermediate with a cis double bond between carbons four and five, the first step of β-oxidation (dehydrogenation between carbons two and three) occurs to produce an intermediate with a trans double bond between carbons two and three and a cis double bond between carbons four and five.

2,4 dienoyl-CoA reductase

The enzyme 2,4 dienoyl CoA reductase reduces this intermediate (using NADPH) to one with a single cis bond between carbons three and four. The newly created cis-bonded molecule is then identical to the one acted on by cis-Δ3-enoyl-CoA isomerase above, which converts it into a regular β-oxidation intermediate, as noted above.

α-oxidation

Yet another consideration for oxidation of fatty acids is α-oxidation. This pathway, which occurs in peroxisomes, is necessary for catabolism of fatty acids that have branches in their chains. For example, breakdown of chlorophyll's phytol group yields phytanic acid (Figure 6.93), which undergoes hydroxylation and oxidation on carbon number two (in contrast to carbon three of β-oxidation), followed by decarboxylation and production of an unbranched intermediate that can be further oxidized by the β-oxidation pathway. Though α-oxidation is a relatively minor metabolic pathway, the inability to perform the reactions of the pathway leads to Refsum's disease where accumulation of phytanic acid leads to neurological damage.

ω-oxidation of fatty acids

In addition to β-oxidation and α-oxidation of fatty acids, which occur in the mitochondria and peroxisomes of eukaryotic cells respectively, another fatty acid oxidation pathway known as ω-oxidation also occurs in the smooth endoplasmic reticulum of liver and kidney cells. It is normally a minor oxidation pathway operating on medium chain fatty acids (10-12 carbons), but gains importance 1) when β-oxidation is not functional or 2) for production of long chain intermediates, such as 20-HETE (20-hydroxyeicosatetraenoic acid), that can function in signaling.

Steps in the process involve 1) oxidation of the terminal methyl group of the fatty acid to an alcohol; 2) oxidation of the alcohol to an aldehyde, and 3) oxidation of the aldehyde group to a carboxylic acid (Figure 6.94). The first oxidation is catalyzed by a mixed function oxidase, and yields 20-HETE if the starting material is arachidonic acid. The last two reactions are catalyzed by alcohol dehydrogenase and each requires NAD+. After the last oxidation, the fatty acid has carboxyl groups at each end and can be attached to coenzyme A at either end and subsequently oxidized, ultimately yielding succinate.

Regulation of fatty acid oxidation

Breakdown of fatty acids is controlled at different levels. The first is by control of the availability of fatty acids from the breakdown of fat. As noted above, this process is by regulating the activity of hormone-sensitive triacylglycerol lipase (HSTL) activity by epinephrine (stimulates) and insulin (inhibits).

A second level of control of fatty acid availability is by regulation of carnitine acyl transferase (Figure 6.87 - see HERE). This enzyme controls the swapping of CoA on an acyl-CoA molecule for carnitine, a necessary step for the fatty acid to be imported into the mitochondrion for oxidation.

The enzyme is inhibited by malonyl-CoA, an intermediate in fatty acid synthesis. Thus, when fatty acids are being synthesized, import of them into the mitochondrion for oxidation is inhibited. Last, the last enzyme in the β-oxidation cycle, thiolase, is inhibited by acetyl-CoA.

Fatty acid synthesis

Synthesis of fatty acids occurs in the cytoplasm and endoplasmic reticulum of the cell and is chemically similar to the reverse of the β-oxidation process, but with a couple of key differences (Figure 6.95). The first of these occur in preparing substrates for the reactions that grow the fatty acid. Fatty acid synthesis occurs in the cytoplasm of eukaryotic cells. Transport of acetyl-CoA from the mitochondrial matrix occurs when it begins to build up. This happens when the citric acid cycle slows or stops from lack of exercise.

Two molecules can play roles in moving acetyl-CoA to the cytoplasm – citrate and acetylcarnitine. Joining of oxaloacetate with acetyl-CoA in the mitochondrion creates citrate which gets transported across the membrane, followed by action of citrate lyase in the cytoplasm of the cell to release acetyl-CoA and oxaloacetate. Additionally, when free acetyl-CoA accumulates in the mitochondrion, it may combine with carnitine and be transported out to the cytoplasm.

Fatty acid synthase

In animals, six different catalytic activities necessary to fully make palmitoyl-CoA are contained in a single complex called Fatty Acid Synthase. As shown in Figures 6.96 and 6.97, these include 1) transacylases (MAT) for swapping CoA-SH with ACP-SH on acetyl-CoA and malonyl-CoA; 2) a synthase (KS) to catalyze addition of the two carbon unit from the three carbon malonyl-ACP in the first step of the elongation process; 3) a reductase (KR) to reduce the ketone; 4) a dehydrase (DH) to catalyze removal of

water; 5) a reductase (ER) to reduce the trans double bond and 6) a thioesterase (TE) to cleave the finished palmitoyl -CoA into palmitic acid and CoA-SH.

In the middle of the complex is a site for binding the ACP portion of the growing fatty acid chain to hold it as the other part of the fatty acid is rotated into positions around the enzyme complex for each catalysis. In bacteria, these six activities are found on separate enzymes and are not part of a complex.

Cytoplasmic reactions

The process of making a fatty acid in the cytoplasm starts with two acetyl-CoA molecules. One is converted to malonyl-CoA by adding a carboxyl group. This reaction is catalyzed by the enzyme acetyl-CoA carboxylase (ACC), the only regulated enzyme of fatty acid synthesis (see below) and the only one separate from the fatty acid synthase. Next, both acetyl-CoA and malonyl-CoA have their CoA portions replaced by a carrier protein known as ACP (acyl-carrier protein) to form acetyl-ACP (catalyzed by acetyl-CoA : ACP transacylase - MAT in Figure 6.97) and malonyl-ACP (catalyzed by malonyl-CoA : ACP transacylase - MAT in Figure 6.97). Joining of a fatty acyl-ACP (in this case, acetyl-ACP) with malonyl-ACP splits out the carboxyl group from malonyl-ACP that was added to it and creates the acetoacyl-ACP intermediate (catalyzed by β-ketoacyl-ACP synthase - KS on Figure 6.97) .

From this point forward, the chemical reactions resemble those of β-oxidation reversed. First, the ketone is reduced to a hydroxyl using NADPH (catalyzed by β-ketoacyl-ACP reductase - KR on Figure 6.97). In contrast to the hydroxylated intermediate of β-oxidation, the intermediate here (D-β- hydroxybutyryl-ACP) is in the D-configuration.

Dehydration

Next, water is removed from carbons 2 and 3 of the hydroxyl intermediate in a reaction catalyzed by 2,3-trans-enoyl-ACP dehydrase - DH on Figure 6.97. This yields a trans doubled-bonded molecule. Last, the double bond is hydrogenated to yield a saturated intermediate by 2,3-trans-enoyl-ACP reductase - ER on Figure 6.97. This completes the first cycle of synthesis.

Additional cycles involve addition of more two-carbon units from malonyl-ACP to the growing chain until ultimately an intermediate with 16 carbons is produced (palmitoyl-ACP). At this point, a thioesterase cleaves the ACP from the palmitoyl-ACP to yield palmitic acid and the cytoplasmic synthesis ceases.

Regulation of fatty acid synthesis

Acetyl-CoA carboxylase, which catalyzes synthesis of malonyl-CoA, is the only regulated enzyme in fatty acid synthesis. Its regulation involves both allosteric control and covalent modification. The enzyme is known to be phosphorylated by both AMP Kinase and Protein Kinase A.

Dephosphorylation is stimulated by phosphatases activated by insulin binding. Dephosphorylation activates the enzyme and favors its assembly into a long polymer, while phosphorylation reverses the process. Citrate acts as an allosteric activator and may also favor polymerization. Palmitoyl-CoA allosterically inactivates it.

Elongation past 16 carbons

Elongation to make fatty acids longer than 16 carbons occurs in the endoplasmic reticulum and is catalyzed by enzymes described as elongases. Mitochondria also can elongate fatty acids, but their starting materials are generally longer than 16 carbons long.

The mechanisms in both environments are similar to those in the cytoplasm (a malonyl group is used to add two carbons, for example), but CoA is attached to the intermediates, not ACP. Further, whereas cytoplasmic synthesis employs the fatty acid synthase complex, the enzymes in these organelles are separable and not part of a complex.

Desaturation of fatty acids

Fatty acids are synthesized in the saturated form and desaturation occurs later - in the endoplasmic reticulum. Reactions to elongate the fatty acid (with elongases) may also occur to make unsaturated fatty acids of varying lengths. Desaturases are named according to the location of the double bonds they introduce in fatty acids. The delta (Δ) system numbers the carbon at the carboxyl end as number 1 and the omega (ω) number system numbers the carbon at the methyl end as number 1 (Figure 6.98). Humans have desaturases named as Δ5, Δ6, and Δ9. A Δ9 desaturase, for example, could convert stearic acid into oleic acid, because stearic acid (see HERE) is a saturated 18 carbon fatty acid and oleic acid is an 18 carbon fatty acid with only one double bond - at position Δ9.

Polyunsaturated fatty acids

Polyunsaturated fatty acids require the action of multiple enzymes and (in some cases) the action of elongases. Arachidonic acid, for example, is a 20 carbon fatty acid with four double bonds and its synthesis requires both an elongase (to increase the length of the fatty acid from 16 to 20) and multiple desaturases - one for each desaturated double bond.

Animals are limited in the fatty acids they can make, due to an inability of their desaturases to catalyze reactions beyond carbons Δ9. Thus, humans can make oleic acid, but cannot synthesize linoleic acid (Δ9,12) or linolenic acid (Δ9,12,15). Consequently, these two must be provided in the diet and are referred to as essential fatty acids.

Almost all desaturases make cis, not trans double bonds. There are a few minor exceptions to this, in cattle, for example (Figure 6.99). The trans fatty acids found in trans fat of prepared food are produced not by biological processes, but rather by the process of partial hydrogenation of unsaturated fats.

Unusual oxidation reaction

Removal of electrons and protons from a fatty acid to create a double bond is an oxidation reaction and these electrons, must have a destination. The path they take is a bit complex. It involves NAD(P)H, O2, two membrane-bound cytochromes, the membrane bound desaturase, and the fatty acid.

In the electron transfer, the O2 is reduced to two molecules of H2O. This reduction requires four electrons and four protons. Two electrons and two protons come from the fatty acid to form the double bond on it. Two electrons come from the NAD(P)H via the cytochromes and two protons come from the aqueous solution.

Prostaglandin synthesis

The pathway for making prostaglandins and related molecules, such as the leukotrienes, prostacyclin, and thromboxanes is an extension of the synthesis of fatty acids (Figure 6.100).

Prostaglandins, known as eicosanoids because they contain 20 carbons, are synthesized in cells from arachidonic acid whenever it has been cleaved from membrane lipids. Prostaglandins are important for many physiological phenomena in the body, including swelling and pain and reduction of their levels is a strategy of some painkillers, such as aspirin (see below). Inflammation arising from bee stings, for example, occurs because bee (and snake) venom contains mellitin, an activator of PLA2 activity (Figure 6.100). There are two strategies for reducing prostaglandin production and the pain associated with it.

Phospholipase A2

Action of phospholipase enzymes on glycerophospholipids produces fatty acids and either glycerol-3-phosphate or other substances. Figure 6.101 shows cleavage sites on phospholipids that are targeted by different phospholipases. Phospholipase A1 (PLA1), for example, cleaves the fatty acid from position one of the glycerophospholipid and phospholipiase D (PLD) cleaves the R group from the phosphate part of the molecule.

Since the fatty acid on position #2 (where PLA2 cuts) is most commonly unsaturated, PLA2 is an important phosopholipase for hydrolyzing the unsaturated fatty acid known as arachidonic acid from glycerophospholipids. Release of arachidonic acid from membranes is necessary for synthesis of prostaglandins.

Inhibition of the release of arachidonic acid from membranes is the mechanism of action of steroidal anti-inflammatory drugs. They block action of phospholipase A2 (PLA2 - Figure 6.101) which cleaves arachidonic acid from membrane lipids.

Lipocortin

Lipocortin (also called annexin) is a protein that inhibits action of PLA2. Synthesis of lipocortin is stimulated by glucocorticoid hormones, such as cortisol, and is used in some treatments to reduce swelling/inflammation when it is severe and untreatable by non-steroidal drugs.

Second strategy

Synthesis of the prostanoid compounds (prostaglandins, prostacyclin, and thromboxanes) depends on conversion of arachidonic acid to prostaglandins G2 and H2 by COX enzymes. A non-steroidal strategy for decreasing production of prostaglandins then is to inhibit the enzyme that catalyzes their synthesis from arachidonic acid (Figure 6.102). This enzyme is known as prostaglandin

synthase, but is more commonly referred to as a cyclooxygenase (or COX) enzyme.

COX enzymes come in at least two forms in humans - COX-1, COX-2. A third form known as COX-3 has been reported as a splice variant of COX-1, but information about it is unclear. COX-1 and COX-2 are very similar in structure (70 kD and 72 kD, respectively, and 65% amino acid sequence homology), but coded by different genes.

COX-1 is synthesized constitutively whereas COX-2 displays inducible expression behavior and has a more specific pattern of tissue expression. COX-2 enzymes are expressed in increasing amounts in areas of growth and inflammation.

Non-steroidal drugs

Molecules inhibiting cyclooxygenases are known as non-steroidal anti-inflammatory drugs (NSAIDs). Molecules in this class include aspirin, ibuprofen, vioxx, and celebrex.

Targeting inhibitors

Some NSAID inhibitors, such as aspirin, bind to all types of COX enzymes. Newer COX inhibitors target the COX-2 enzyme specifically because it was believed to be a better target for relief of joint pain than COX-1 enzymes which are synthesized by most cells. COX-2 enzymes are found more specifically in joints so the thinking was that specific inhibition of them would not affect the COX-1 enzymes which are important for producing prostaglandins that help maintain gastric tissue.

Numerous COX-2 - specific inhibitors were developed - celecoxib, etoricoxib, and rofecoxib (Vioxx), for example. Unfortunately, the COX-2 specific inhibitors are associated with some serious side effects, including a 37% increase in incidence of major cardiovascular events in addition to some of the gastrointestinal problems of NSAIDs.

Imbalance

The increased risk of heart attack, thrombosis, and stroke are apparently due to an imbalance between prostacyclin (reduced by inhibitors) and thromboxanes (not reduced by the inhibitors). Prostacyclin (made from prostaglandin H2 by prostacyclin synthase) is a special prostaglandin that inhibits activation of blood platelets in the blood clotting process and acts as a vasodilator. Thromboxanes counter prostacyclin, causing vasoconstriction and activating blood platelets for clotting. Due to imbalances in these opposite acting molecules resulting from COX-2-specific inhibition, Vioxx, was withdrawn from the market in September, 2004, due to health concerns.

Other compounds known to inhibit COX enzymes include some flavonoids, some components of fish oil, hyperforin, and vitamin D.

Connections to other pathways

There are several connections between fats and fatty acid metabolism and other metabolic pathways. Diacylglycerol (DAG - Figure 6.105), which is produced by removal of a phosphate from phosphatidic acid, is an intermediate in fat synthesis and also a messenger in some signaling systems. Phosphatidic acid, of course, is a branch intermediate in the synthesis of triacylglycerols and other lipids, including phosphoglycerides.

Fatty acids twenty carbons long based on arachidonic acid (also called eicosanoids) are precursors of the leukotrienes, prostaglandins, thromboxanes, and endocannabinoids.

Acetyl-CoA from β-oxidation can be assembled by the enzyme thiolase to make acetoacetyl-CoA, which is a precursor of both ketone bodies and the isoprenoids, a broad category of compounds that include steroid hormones, cholesterol, bile acids, and the fat soluble vitamins. In plants, acetyl-CoA can be made into carbohydrates in net amounts via the glyoxylate cycle.

Fat, obesity, and hunger

Obesity is an increasing problem in the western world. It is, in fact, the leading preventable cause of death worldwide. In 2014, over 600 million adults and 42 million children in the world were classified as obese, a condition when their body mass index is over 30 kg/m2 (Figure 6.106). The body mass index of a person is obtained by dividing a person's weight by the square of their height. At a simple level, obesity arises from consumption of calories in excess of metabolic need, but there are many molecular factors to consider.

Adipokines

Adipokines are adipose tissue-synthesized cytokines. The class of molecules includes leptin (first discovered adipokine) and hundreds of other such compounds. These include adiponectin (regulates glucose levels and fatty acid oxidation), apelin (control of blood pressure, angiogenesis promotion, vasodilator release, increased water intake), chemerin (stimulation of lipolysis, adipocyte differentiation, link to insulin resistance), and resistin (links to obesity, type II diabetes, LDL production in liver), among others.

Resistin

Resistin is an adipokine peptide hormone with numerous associated negative health effects. Injection of the hormone into mice results in increased resistance to insulin, a phenomenon of type 2 diabetes.

Resistin is linked to increased inflammation and serum levels of it correlate with increased obesity, though direct linkage of it to obesity is controversial. Resistin stimulates production of LDLs in the liver, supporting increased levels in the arteries. Resistin also adversely impacts the effects of statin drugs used to control levels of cholesterol in the body.

Leptin

Leptin is a peptide hormone (adipokine) made in adipose cells that negatively impacts hunger and regulates energy balance. It is countered by ghrelin, also known as the hunger hormone. Both hormones act in the hypothalamus where hunger is controlled. When leptin levels are higher due to higher levels of body fat, hunger is suppressed, but when levels of leptin are lower (less body fat), then appetite increases.

Notably, leptin is also made in places besides adipose tissue and leptin receptors are found in places besides the hypothalamus, so the hormone has other effects in the body. When sensitivity to leptin changes, increased obesity can result. In mice, deletion of leptin function by mutation results in mice with voracious appetites and extreme obesity. Deletion of the leptin receptor gene in mice results in the same phenotype. Eight humans with leptin mutations all suffer from extreme obesity in infancy.

Physiology

Leptin is produced primarily by cells in white adipose tissue, but is also made in brown adipose tissue, ovaries, skeletal muscle, stomach, mammary epithelial cells and bone marrow.

Leptin levels

Leptin levels in the body are highest between midnight and early morning, presumably to suppress appetite. Though it is produced by fat cells, levels of leptin in humans do not strictly reflect levels of fat. For example, early in fasting, leptin levels fall before fat levels fall. Sleep deprivation can reduce leptin levels, as can increasing levels of testosterone and physical exercise.

Increasing estrogen, however, increases leptin levels. Emotional stress and insulin can increase leptin levels. Obesity increases leptin levels, but doesn't fully suppress appetite. Leptin resistance in these individuals is an important consideration, lessening the effects of the hormone on appetite.

Blocking leptin action

In the medial hypothalamus, leptin stimulates satiety and in the lateral hypothalamus, leptin inhibits hunger. Lesions in the lateral hypothalamus that block the ability to sense hunger result in anorexia (there are other causes of anorexia, though) and lesions in the medial hypothalamus cause excess hunger (no satiety). Neuropeptide Y is a potent hunger promoter whose receptors in the arcuate nucleus can be bound and blocked by leptin. Leptin levels are more sensitive to decreasing food intake than increasing food intake meaning that in humans the hormone plays a bigger role with respect to appetite than to levels of fat in the body.

At the molecular level, binding of leptin to the Ob-Rb receptor causes down-regulation of synthesis of endocannabinoids, whose normal function is to increase hunger. High fructose diets have been associated with reduced levels of leptin and of leptin receptor.

Ghrelin

Ghrelin is a peptide hormone made by cells in the gastrointestinal tract when the stomach is empty. Stretching of the stomach reduces the expression of the hormone. Ghrelin exerts its effects on the central nervous system to increase appetite and it is an unusual peptide in being able to cross the blood-brain barrier. The ghrelin receptor in the brain is found on the same cells as the leptin receptor (arcuate nucleus). Leptin can counter the ghrelin effect by decreasing hunger.

Behavioral effects

Activation of ghrelin occurs after processing the zymogen form of the hormone (pre-proghrelin) followed by linkage of an octanoic acid to a serine at position 3. Circulating levels of ghrelin increase before eating and decrease afterwards. There appears to be a dose dependence for ghrelin on the amount of food consumed. Ghrelin increases food seeking behavior and there is a negative correlation between levels of ghrelin and weight.

Neuropeptide Y

Neuropeptide Y is a neuropeptide neurotransmitter produced by neurons of the sympathetic nervous system. It acts as a vasoconstrictor and favors growth of fat tissue. It appears to stimulate food intake, fat storage, relieve anxiety/stress, reduce pain perception, and lower blood pressure. Blockage of neuropeptide Y receptors in the brain of rats decreases food intake.

Stress effects

In mice and monkeys, repeated stress and high fat, high sugar diets stimulate neuropeptide Y levels and cause abdominal fat to increase.

High levels of neuropeptide Y may also help individuals to recover from post-traumatic stress disorder and to reduce the fear response. It may also protect against alcoholism. Mice lacking the ability to make neuropeptide Y have a higher voluntary consumption of alcohol and are less sensitive to its effects. The neuropeptide Y receptor is a G-protein-coupled receptor in the 7-transmembrane domain family.

Metabolism: Fats and Fatty Acids

564

YouTube Lectures

by Kevin

HERE & HERE

565

Figure 6.83 - Movement of dietary triglycerides

Image by Aleia Kim

Figure 6.82 - Trimyristin - A triacylglyceride

566

Figure 6.84 - Breakdown of fat in adipocytes

Image by Pehr Jacobson

567

Figure 6.85 - Synthesis of fat from phosphatidic acid (phosphatidate)

Image by Penelope Irving

Synthesis of Phosphatidic Acid from Glycerol-3-phosphate

1. Glycerol-3-phosphate + Acyl-CoA <=> Monoacylglycerol phosphate + CoA-SH

2. Monoacylglycerol phosphate + Acyl-CoA <=> Phosphatidic acid + CoA-SH

568

Figure 6.86 - Mitochondria - site of β-oxidation

Figure 6.87 - Transport of fatty acid (acyl group) across mitochondrial inner membrane

Image by Aleia Kim

569

Figure 6.88 - Four reactions in β-oxidation

Image by Aleia Kim

YouTube Lectures

by Kevin

HERE & HERE

570

Figure 6.89 - Similar reactions for fatty acid oxidation and oxidation of 4-carbon compounds in the citric acid cycle

Image by Aleia Kim

571

Figure 6.90 - Metabolism of propionyl-CoA

Image by Pehr Jacobson

572

In beta oxidation, it just occurred to me

The process all takes place ‘tween carbons two and three

Some hydrogens are first removed to FADH2

Then water adds across the bond, the H to carbon two

Hydroxyl oxidation’s next, a ketone carbon three

Then thiolase catalysis dissects the last two C’s

The products of the path, of course, are acetyl-CoAs

Unless there were odd carbons, hence propionyl-CoA

Figure 6.91 - Cerotic acid - A long chain fatty acid with 26 carbons

573

Figure 6.92 - Unsaturated fatty acid oxidation

YouTube Lectures

by Kevin

HERE & HERE

574

Figure 6.93 - Phytanic acid

575

Figure 6.94 - ω Oxidation

Wikipedia

576

Figure 6.95 - Fatty acid synthesis is the reverse of fatty acid oxidation chemically

577

Figure 6.96 - One round of fatty acid synthesis

image by Aleia Kim

Figure 6.97 - Fatty acid synthase complex

578

579

For fatty acid synthesis, I must reverse the path

Of breaking fatty acids, though you'll wonder 'bout the math

Each cycle of addition starts with carbons one two three

Yet products of reactions number carbons evenly

The reason is that CO2 plays peek-a-boo like games

By linking to an Ac-CoA then popping off again

Reactions are like oxidations 'cept they're backwards here

Reduction, dehydration, then two hydrogens appear

The product of the process is a 16 carbon chain

The bonds are saturated. No double ones remain

For them desaturases toil to put in links of cis

In animals to delta nine, but no more go past this

And last there's making longer ones eicosanoidic fun

They're made by elongases in the e. reticulum

Kevin Ahern

YouTube Lectures

by Kevin

HERE & HERE

Figure 6.98 - Carbon numbering schemes for fatty acids

Image by Pehr Jacobson

580

Stearoyl-CoA + 2 Cytochrome b5 (red) + O2 + 3 H+ + NADPH

Oleoyl-CoA + 2 Ferricytochrome b5 (ox) + 2 H2O + NADP+

Desaturase Reaction to Oxidize Stearic Acid

Figure 6.99 - Elaidic acid - A rare trans fatty acid in biology

581

Figure 6.100 - Eicosanoid synthesis pathways

Image by Pehr Jacobson

582

583

Figure 6.101 - Cleavage sites for four phospholipiases on a glycerophospholipid - phospholipases A1 (PLA1), A2 (PLA2), C (PLC), and D (PLD)

YouTube Lectures

by Kevin

HERE & HERE

584

Figure 6.102 - Catalytic activity of cyclooxygenase and peroxidase in making prostaglandins

Image by Pehr Jacobson

585

Figure 6.103 - Synthesis of prostaglandins from prostaglandin H2 (red)

Image by Pehr Jacobson

Figure 6.104 - Two NSAIDs

586

Figure 6.105 - Diacylglycerol

Figure 6.106 - Obesity worldwide - females (top) and males (bottom)

Wikipedia

587

Figure 6.107 Leptin

Wikipedia

588

YouTube Lectures

by Kevin

HERE & HERE

Figure 6.108 Neuropeptide Y

589

590

Figure 6.109 - Pre-proghrelin

591

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

592

When Acids Get Oxidized

To the tune of "When Johnny Comes Marching Home"

Metabolic Melodies Website HERE

The fatty acids carried by
CoA, CoA
Are oxidized inside the
mi-to-chon-dri-ay

They get to there as you have seen
By hitching rides on carnitine
Then it goes away
When acids get oxidized

Electrons move through membranes, yes
It's true, it's true
They jump from complex I onto
Co-Q, Co-Q

The action can be quite intense
When building proton gradients
And its good for you
When acids get oxidized

The protons pass through complex V
You see, you see
They do this to make lots of
A-TP, TP

The mechanism you should know
Goes through the stages L-T-O
So there's energy
When acids get oxidized

Recording by Tim Karplus

Lyrics by Kevin Ahern
Recording by Tim Karplus Lyrics by Kevin Ahern

593

When Acids Are Synthesized

To the tune of “When Johnny Comes Marching Home”

Metabolic Melodies Website HERE

The 16 carbon fatty acid, palmitate
Gets all the carbons that it needs from acetate
Which citric acid helps release
From mitochondri - matrices
Oh a shuttle's great
When acids are synthesized

Carboxylase takes substrate and it puts within
Dioxy carbon carried on a biotin
CoA's all gain a quick release
Replaced by larger ACPs
And it all begins
When acids are synthesized

A malonate contributes to the growing chain
Two carbons seven times around again, again
For saturated acyl-ates
There's lots of N-A-DPH
That you must obtain
When acids are synthesized

Palmitic acid made this way all gets released
Desaturases act to make omega-threes
The finished products big and small
Form esters with a glycerol
So you get obese
When acids are synthesized

Recording by Tim Karplus

Lyrics by Kevin Ahern
Recording by Tim Karplus Lyrics by Kevin Ahern

This page titled 6.3: Fats and Fatty Acids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6.4: Other Lipids

Source: BiochemFFA_6_4.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Sugars are the building blocks of carbohydrates, amino acids are the building blocks of proteins and nucleotides are the building blocks of the nucleic acids - DNA and RNA. Another crucial building block is **acetyl-CoA**, which is used to build many lipid substances, including fatty acids, cholesterol, fat soluble vitamins, steroid hormones, prostaglandins, endocannabinoids, and the bile acids. Indeed, acetyl-CoA goes into more different classes of molecule than any other building block.

Isoprenoids

We focus our attention here on a group of molecules made from acetyl-CoA that are known as as the isoprenoids. Isoprenoids are a large, diverse and ancient group of molecules that are found in all three domains of life. As noted earlier, they are components of membrane lipids in the cell membranes of archaebacteria, but beyond this, they serve an astonishing variety of functions. From photosynthetic pigments to plant defense compounds, from flavor compounds in cinnamon, mint, ginger and cloves to plant and animal hormones, from the cannabinoids in marijuana to the lycopene that gives tomatoes their color, and from heme to the quinones in the electron transport chain, isoprenoids are ubiquitous in cells. Isoprenoids derive their name from the fact that they are, in fact, made from five carbon building blocks called isoprenes that are derived from acetyl-CoA. The synthesis of the two isoprene units - isopentenyl pyrophosphate and dimethylallyl pyrophosphate is shown in Figure 6.111 and Figure 6.112.

The pathway leading up to isoprene synthesis overlaps with that of ketone body synthesis, for the two reactions (Figure 6.112), as has been discussed earlier in this book (see HERE). Thiolase catalyzes the initial reaction, joining together two acetyl-CoA molecules to make acetoacetyl-CoA. In the second reaction catalyzed by HMG-CoA synthase, a third acetyl-CoA is joined to form the six carbon compound known as hydroxymethyl glutaryl-CoA (HMG-CoA). Reaction three is an important one biologically and medically because of the enzyme catalyzing it - HMG-CoA reductase.

Statins

Medically, HMG-CoA reductase is the target of a class of drugs known as statins (Figure 6.113 & Movie 6.1), which are used to reduce cholesterol levels in people. These competitive inhibitors, which compete with HMG-CoA for binding have two effects. First, they reduce the production of mevalonate, which restricts the amount of substrate available to make cholesterol. Second, and perhaps more importantly, they increase production of LDL receptors in the liver, which favors uptake and destruction of LDLs, thus lowering serum cholesterol levels.

Regulation

Biologically, the HMG-CoA reductase enzyme is also of importance because it is the primary regulatory point in cholesterol synthesis. Control of it is complex. First, it is feedback inhibited by cholesterol itself. High levels of glucose in the blood activate the enzyme. Phosphorylation by AMP-activated protein kinase inhibits its activity. Interestingly, the same enzyme phosphorylates and inactivates acetyl-CoA carboxylase - the only regulatory enzyme controlling fatty acid synthesis. Transcription of the gene encoding HMG-CoA reductase

is enhanced by binding of the sterol regulatory element binding protein (SREBP) to the sterol recognition element (SRE) located near the gene coding sequence. As cholesterol levels rise, SREBP is proteolytically cleaved and transcription stops.

From HMG-CoA, the enzyme HMG-CoA reductase catalyzes the formation of mevalonate. This reaction requires NADPH and results in release of coenzyme A. Mevalonate gets phosphorylated twice and then decarboxylated to yield the five carbon intermediate known as isopentenyl-pyrophosphate (IPP). IPP is readily converted to the other important isoprenoid unit, dimethylallylpyrophosphate (DMAPP).

Isoprenes

These two five carbon compounds, IPP and DMAPP, are also called isoprenes (Figure 6.115) and are the building blocks for the synthesis of cholesterol and related compounds. This pathway proceeds in the direction of cholesterol starting with the joining of IPP and DMAPP to form geranyl-pyrophosphate. Geranyl-pyrophosphate combines with another IPP to make farnesyl-pyrophosphate, a 15-carbon compound.

Squalene

Two farnesyl-pyrophosphates join to create the 30-carbon compound known as squalene. Squalene, in a complicated rearrangement involving reduction and molecular oxygen forms a cyclic intermediate known as lanosterol (Figure 6.116) that resembles cholesterol. Conversion of lanosterol to cholesterol is a lengthy process involving 19 steps that occur in the endoplasmic reticulum.

The cholesterol biosynthesis pathway from lanosterol is a long one and requires significant amounts of reductive and ATP energy. As noted earlier (see HERE), cholesterol has an important role in membranes. It is also a precursor of steroid hormones and bile acids and its immediate metabolic precursor, 7-dehydrocholesterol (Figure 6.117), branches to form vitamin D (Figure 6.118).

All steroid hormones in animals are made from cholesterol and include the progestagens, androgens, estrogens, mineralocorticoids, and the glucocorticoids. The branch molecule for all of the steroid hormones is the cholesterol metabolite (and progestagen) known as pregnenalone (Figure 6.119). The progestagens are thus precursors of all of the other classes of steroid hormones.

The estrogens are derived from the androgens in an interesting reaction that required formation of an aromatic ring (Figure 6.120). The enzyme catalyzing this reaction is known as an aromatase and it is of medical significance. The growth of some tumors is stimulated by estrogens, so aromatase inhibitors are prescribed to prevent the formation of estrogens and slow tumor growth. Two commonly used inhibitors include exemestane (a suicide inhibitor - Figure 6.121) and anastrozole (a competitive inhibitor).

Other fat-soluble vitamins

Synthesis of other fat soluble vitamins and chlorophyll also branches from the isoprenoid synthesis pathway at geranyl pyrophosphate. Joining of two geranylgeranyl pyrophosphates occurs in plants and bacteria and leads to synthesis of lycopene, which, in turn is a precursor of β-carotene, the final precursor of Vitamin A (see below also). Vitamins E and K, as well as chlorophyll are all also synthesized from geranylgeranyl pyrophosphate.

Bile acid metabolism

Another metabolic pathway from cholesterol leads to the polar bile acids, which are important for the solubilization of dietary fat during digestion. Converting the very non-polar cholesterol to a bile acid involves oxidation of the terminal carbon on the side chain off the rings. Other alterations to increase the polarity of these compounds include hydroxylation of the rings and linkage to other polar compounds.

Common bile acids include cholic acid, chenodeoxycholic acid, glycocholic acid, taurocholic acid, and deoxycholic acid (Figure 6.123). Another important consideration about bile acids is that their synthesis reduces the amount of cholesterol available and promotes uptake of LDLs by the liver. Normally bile acids are recycled efficiently resulting in limited reduction of cholesterol levels. However, inhibitors of the recycling promote reduction of cholesterol levels.

Vitamin A Synthesis

Vitamin A is important for many cellular functions related to growth, differentiation and organogenesis during embryonic development, tissue maintenance, and vision, to name a few.

There are three main active forms of the vitamin, retinal, retinol and retinoic acid, each with its own set of functions. Retinal, complexed with the protein, opsin, is found in the rod cells of the retina and is necessary for vision. Retinol and retinoic acid both

function as signaling molecules that can modulate gene expression during development.

Synthesis of vitamin A occurs as a branch in synthesis of isoprenoids. Addition of isopentenyl pyrophosphate to farnesyl pyrophosphate creates a 20-carbon intermediate, geranylgeranyl pyrophosphate (GGPP - Figure 6.124).

Joining of two GGPPs creates a 40 carbon intermediate that is unstable and decomposes to phytoene. Desaturases oxidize two single bonds in phytoene, creating lycopene.

Lycopene is a linear 40 carbon unsaturated molecule found in tomatoes and other red vegetables and it gives them their color. Cyclization of end portions of lycopene give rise to β-carotene, the precursor of vitamin A (retinal/retinol - Figure 6.124).

β-carotene is found in carrots and other orange vegetables, and is converted in the body to vitamin A. Catalytic action by β-Carotene 15,15' monooxygenase cleaves β-carotene to form retinal (the aldehyde form used in vision).

The enzyme retinol dehydrogenase catalyzes reduction of retinal to retinol (storage form). Oxidation of retinal creates another important retinoid known as retinoic acid. This form of vitamin A cannot be reduced back to retinal and thus cannot be used for vision or storage.

Instead, retinoic acid has roles in embryonic development. Retinoic acid acts through binding to the Retinoic Acid Receptor (RAR). RAR binds to DNA and affects transcription of several important sets of genes important for differentiation. These include the Hox genes, which control anterior/posterior patterning in early embryonic development.

Sphingolipid synthesis

Synthesis of sphingolipids, which are found primarily in brain and nerve tissue, begins with palmitoyl-CoA and serine that combine to make an 18-carbon amine called 3-keto-sphinganine (Figure 6.125). Reduction of that by NADPH yields dihydrosphingosine and addition of a fatty acid from an acyl-CoA yields N-acylsphinganine, which is a ceramide (Figure 6.126). A ceramide can be converted into a cerebroside by addition of a glucose from UDP-glucose (Figure 6.127).

If a few other simple sugars are added to the cerebroside, a globoside is created. If, instead of adding sugar, a phosphocholine is added from phosphatidylcholine, then sphingomyelin is created (Figure 6.127). If a complex set of sugars are added to to a cerebroside, then a ganglioside results (Figure 6.127).

Sphingolipid breakdown

In the overall metabolism of sphingolipids, the greatest problems arise with their catabolism. Figure 6.128 illustrates the numerous genetic diseases arising from mutations in DNA coding for some of these enzymes. All are lysosomal storage diseases and many of these are quite severe. GM1 glandiosidoses (arising from inability to breakdown GM1 gangliosides) cause severe neurodegeneration and seizures. Individuals suffering from them typically die by age 3. Tay-Sachs disease usually causes death by age 4, though late-onset forms of the disease in adults are known.

With Gaucher's disease, three different types have been described with widely varying effects. In some, the disease is fatal by age four and in others, it does not manifest until teens or even adulthood. Fabry's disease patients can live into their 50s, on average.

Glycerophospholipid metabolism

Glycerophospholipids are the major components of membranes. Synthesis of glycerophospholipids begins with glycerol-3-phosphate. In the first reaction, glycerol-3-phosphate gains a fatty acid at position one from an acyl-CoA, followed by a duplicate reaction at position two to make phosphatidic acid (Figure 6.129). This molecule, which can branch to other reactions to form fats, is an important intermediate in the synthesis of many glycerophospholipids. Glycerophospholipid compounds can often be made by more than one pathway. The nucleotide CDP plays an important role in glycerophospholipid synthesis, serving as part of an activated intermediate for synthesis of phosphatidyl compounds. This is necessary, because formation of the phosphodiester bonds of these compounds requires higher energy input.

Cells use two strategies to accomplish this. Both involve CDP. In the first, CTP combines with phosphatidic acid to make CDP-diacylglcyerol with release of a pyrophosphate. The reaction is catalyzed by phosphatidate cytidylyltransferase.

CDP-diacylglycerol then serves as an activated intermediate to donate the phosphotidate part of itself to another molecule. The reaction below illustrates one example

The second strategy is to make a CDP derivative of the group being added to phosphatidic acid. An example is shown next

Then the CDP donates the phosphocholine to a diacylglycerol to made phosphatidylcholine and CMP

Synthesis of other important glycerophospholipids follows from these basic strategies. Phosphatidylethanolamine can be easily made from phosphatidylserine by decarboxylation.

Phosphatidylethanolamine can serve as a precursor in an alternative pathway for making phosphatidylcholine (SAM = S-Adenosyl Methioinine / SAH = S-Adenosyl Homocysteine)

Phosphatidylserine and phosphatidylethanolamine can swap groups reversibly in the reaction below

Similarly, phosphatidylserine and phosphatidylcholine can be interchanged as follows:

Phosphatidylglycerol can be made from glycerol-3-phosphate and CDP-diacylglycerol

Cardiolipin, which is essentially a diphosphatidyl compound can be made by joining CDP-diacylglyerol with phosphatidylglycerol

Phosphtidylinositol can be made from CDP-diacylglycerol and inositol .

Heme synthesis

The porphyrin ring found in the hemes of animals, fungi, and protozoa (Figure 6.130) is synthesized starting from very simple compounds (Figure 6.131). The process is a bit complicated, occurring between the cytoplasm and the mitochondrion. The first step is the creation of δ-aminolevulinic acid (also called aminolevulinic acid or dALA) from glycine and Succinyl-CoA.

Joining of two δ-aminolevulinic acid molecules together with splitting out of two molecules of water yields porphobilinogen.

Joining of four molecules of porphobilinogen together yields hydroxylmethylbilane (Figure 6.132).

Next, a series of reactions involving 1) loss of water; 2) loss of four molecules of carbon dioxide; 3) loss of two more carbon dioxides, loss of six protons and electrons and (finally) 4) addition of Fe++ with loss of two protons yields heme. Individual heme molecules may be further processed.

Two enzymes in heme synthesis are sensitive to the presence of lead, and this is one of the primary causes of lead toxicity in humans. Inhibition of the enzymes leads to 1) anemia and 2) accumulation of δ-aminolevulinic acid, which can be harmful to neurons in development, resulting in learning deficiencies in children.

Porphyria

Defects in enzymes of the pathway can also lead to porphyrias, diseases in which one or more of the intermediates in the heme synthesis pathway accumulate due to deficiency of the enzyme necessary to convert the accumulating material into the next molecule in the pathway. The accumulation of purplish intermediates gave the diseases the name porphyria from the Greek word for purple.

Severe porphyrias can lead to brain damage, nerve damage, and mental disturbances. The “madness” of King George III may have been due to a form of porphyria. In other manifestations of the disease, cutaneous porphyrias cause skin problems on exposure to light. This need, for patients with certain forms of porphyria, to avoid light, coupled with the fact that porphyrias can be treated by blood tranfusions, may have led to the legend of vampires.

Breakdown of heme

Catabolism of heme (Figure 6.133) begins in macrophages within the spleen . Targets for degradation are hemes within damaged red blood cells, which get removed from the blood supply due to their appearance. It is because of this system, for example, that

sickle cell anemia is classified as an anemia (decrease in red blood cells or hemoglobin in the blood). After cells have sickled, they lose their shape and are more likely to be removed from the blood by this process, leaving the patient weakened from low blood cell counts.

The first biochemical step in catabolism is conversion of heme to biliverdin. This reaction is catalyzed by heme oxygenase and requires electrons from NADPH. In the process, Fe++ is released. Interestingly, carbon monoxide is also produced and it acts as a vasodilator.

Next, biliverdin is converted to bilirubin by biliverdin reductase and is secreted from the liver into bile. Bacteria in the intestine convert bilirubin to urobilinogens, some of which is absorbed intestinal cells and transported into kidneys and excreted. The yellow color of urine arises from the compound known as urobilin, which is an oxidation product of urobilinogen. The remainder of the urobilinogens are converted in the intestinal tract to stercobilinogen whose oxidation product is stercobilin and it gives the color associated with feces.

This page titled 6.4: Other Lipids is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6.5: Amino Acids and the Urea Cycle

Source: BiochemFFA_6_5.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

In contrast to some of the metabolic pathways described to this point, amino acid metabolism is not a single pathway. The 20 amino acids have some parts of their metabolism that overlap with each other, but others are very different from the rest. In discussing amino acid metabolism, we will group metabolic pathways according to common metabolic features they possess (where possible). First, we shall consider the anabolic pathways.

Transamination

Before beginning discussion of the pathways, it is worthwhile to discuss a reaction common to the metabolism of most of the amino acids and other nitrogen-containing compounds and that is transamination. In cells, nitrogen is a nutrient that moves from one molecule to another in a sort of hand-off process. A common transamination reaction is shown on the next page.

A specific reaction of this type is shown in Figure 6.134.

Glutamate and glutamine play central roles in transamination, each containing one more amine group than α-ketoglutarate and glutamate, respectively. Transamination reactions, as noted earlier, occur by a ping-pong mechanism and involve swaps of amines and oxygens in Schiff base reactions. Two amino acids, glutamine and asparagine are the products of gaining an amine in their respective R-groups in reactions involving ammonium ion.

Synthesis varies

It is also important to recognize that organisms differ considerably in the amino acids that they can synthesize. Humans, for example, cannot make 9 of the 20 amino acids needed to make proteins, and the number of these that can be synthesized in needed amounts varies between adults and children.

Amino acids that cannot be made by an organism must be in the diet and are called essential amino acids. Non-essential amino acids are those an organism can make in sufficient quantities (Figure 6.135). Though amino acids do not have a common pathway of metabolism, they are often organized in “families” of amino acids with overlapping metabolic reactions common to members of each group. To designate amino acid families in the text we will use a blue font for headings to distinguish them.

α-ketoglutarate family

This family of amino acids arises from α-ketoglutarate of the citric acid cycle. It includes the amino acids glutamic acid, glutamine, proline, and arginine. It is also called the glutamate family, since all the amino acids in it derive from glutamate.

Glutamate

α-ketoglutarate is readily converted to glutamate in transamination reactions, as noted above. It can also be produced by the enzyme glutamate dehydrogenase, which catalyzes the reaction below (in reverse) to make glutamate.

In the forward direction, the reaction is a source of ammonium ion, which is important both for the urea cycle and for glutamine metabolism. Because it is a byproduct of a citric acid cycle intermediate, glutamate can therefore trace its roots to any of the intermediates of the cycle. Citrate and isocitrate, for example, can be thought of as precursors of glutamate. In addition, glutamate can be made by transamination from α-ketoglutarate in numerous transamination reactions involving other amino acids.

Glutamine

Synthesis of glutamine proceeds from glutamate via catalysis of the enzyme glutamine synthetase, one of the most important regulatory enzymes in all of amino acid metabolism (Figure 6.136).

Regulation of the enzyme is complex, with many allosteric effectors. It can also be controlled by covalent modification by adenylylation of a tyrosine residue in the enzyme (Figure 6.137). In the figure, PA and PD are regulatory proteins facilitating conversion of the enzyme.

Ammonia used in the reaction catalyzed by glutamate synthetase commonly arises from nitrite reduction, amino acid breakdown, or photorespiration. Because it builds ammonia into an amino acid, glutamine synthetase helps reduce the concentration of toxic ammonia - an important consideration in brain tissue. Some inhibitors of glutamine synthetase are, in fact, the products of glutamine metabolism. They include histidine, tryptophan, carbamoyl phosphate, glucosamine-6-phosphate, CTP, and AMP. The glutamate substrate site is a target for the inhibitors alanine, glycine, and serine. The ATP substrate site is a target for the inhibitors GDP, AMP, and ADP. Complete inhibition of the enzyme is observed when all of the substrate sites of the multi-subunit enzyme are bound by inhibitors. Lower levels of inhibitors results in partial or full activity, depending on the actual amounts.

Proline

Synthesis of proline starts with several reactions acting on glutamate. They are shown below in the green text box.

The L-glutamate-5-semialdehyde, so produced, is a branch point for synthesis of proline or ornithine. In the path to make proline, spontaneous cyclization results in formation of 1-pyrroline-5-carboxylic acid (Figure 6.138).

This, in turn, is reduced to form proline by pyrroline-5-carboxylate reductase.

Arginine

Arginine is a molecule synthesized in the urea cycle and, thus, all urea cycle molecules can be considered as precursors. Starting with citrulline, synthesis of arginine can proceed as shown on the next page. The urea cycle can be seen HERE.

An alternate biosynthetic pathway for making arginine from citrulline involves reversing the reaction catalyzed by nitric oxide synthase. It catalyzes an unusual five electron reduction reaction that proceeds in the following manner

Yet another way to synthesize arginine biologically is by reversal of the arginase reaction of the urea cycle

Arginine can also be made starting with glutamate. This 5 step pathway leading to ornithing is illustrated at the top of the next page (enzymes in blue). Ornithine, as noted above can readily be converted to arginine.

The last means of making arginine is by reversing the methylation of asymmetric dimethylarginine (ADMA - Figure 6.140). ADMA is a metabolic byproduct of protein modification. It interferes with production of nitric oxide and may play a role in cardiovascular disease, diabetes mellitus, erectile dysfunction, and kidney disease.

Serine family

Serine is a non-essential amino acid synthesized from several sources. One starting point is the glycolysis intermediate, 3-phosphoglycerate, (3-PG) in a reaction catalyzed by 3-PG dehydrogenase.

Transamination by phosphoserine aminotransferase produces O-phosphoserine. The phosphate is then removed by phosphoserine phosphatase, to make serine. These reactions are shown below. Phosphoserine phosphatase is missing in the genetic disease known as Williams-Beuren syndrome.

Serine can also be derived from glycine and vice versa. Their metabolic paths are intertwined as will be seen below. Serine is important for metabolism of purines and pyrimidines, and is the precursor for glycine, cysteine, and tryptophan in bacteria, as well as for sphingolipids and folate. Serine in the active site of serine proteases is essential for catalysis. A serine in the active site of acetylcholinesterases is the target of nerve gases and insecticides.

Covalent modification target

Serine in proteins can be the target of glycosylation or phosphorylation. D-serine is the second D-amino acid known to function in humans. It serves as a neuromodulator for NMDA receptors, by serving as a co-agonist, together with glutamate. D-serine is being studied as a schizophrenia treatment in rodents and as a possible biomarker for Alzheimers.

Glycine

As noted, glycine's metabolism is intertwined with that of serine. This is apparent in the reaction catalyzed by serine hydroxymethyltransferase.

Notably, the previous reaction is also needed for recycling of folate molecules, which are important for single carbon reactions in nucleotide synthesis.

Vertebrates can also synthesize glycine in their livers using the enzyme glycine synthase.

Glycine is a very abundant component of collagen. It is used in the synthesis of purine nucleotides and porphyrins. It is an inhibitory neurotransmitter and is a co-agonist of NMDA receptors with glutamate. Glycine was detected in material from Comet Wild 2.

Cysteine

Cysteine can be synthesized from several sources. One source is the metabolism of the other sulfur-containing amino acid, methionine. This begins with formation of S-Adenosyl-Methionine (SAM), catalyzed by methionine adenosyltransferase.

SAM is a methyl donor for methyl transfer reactions and that is the next step in the pathway - donation of a methyl group (catalyzed by transmethylase)

SAH (S-Adenosylhomocysteine) is cleaved by S-adenosylhomocysteine hydrolase,

Homocysteine can be recycled back to methionine by action of methionine synthase

On the path to making cysteine, homocysteine reacts as follows (catalyzed by cystathionine β-synthase).

Last, cystathionase catalyzes release of cysteine

β-ketobutyrate can be metabolized to propionyl-CoA and then to succinyl-CoA to be used ultimately in the citric acid cycle.

Another route to making cysteine is a two-step process that begins with serine, catalyzed first by serine-O-acetyltransferase

and then by cysteine synthase

Cysteine can be also released from cystine by cystine reductase

Finally, cysteine can be made from cysteic acid by action of cysteine lyase

Aspartate family

Metabolism of aspartic acid is similar to that of glutamate. Aspartic acid can arise from transamination of a citric acid cycle intermediate (oxaloacetate).

Aspartate can also be generated from asparagine by the enzyme asparaginase.

Further, aspartate can be produced by reversal of a reaction in the urea cycle (see HERE)

Aspartate is also a precursor to four amino acids that are essential in humans. They are methionine, isoleucine, threonine, and lysine. Because oxaloacetate can be produced from aspartate, aspartate is an important intermediate for gluconeogenesis when proteins are the energy source.

Asparagine

Asparagine, too, is an amino acid produced in a simple transamination reaction. In this case, the precursor is aspartate and the amine donor is glutamine (catalyzed by asparagine synthetase)

Methionine

Metabolism of methionine overlaps with metabolism of the other sulfur-containing amino acid, cysteine. Methionine is not made in humans (essential) so the pathway shown in Figure 6.141 is from bacteria.

The process begins with phosphorylation of aspartate. Numbers for each catalytic step in the figure are for the enzymes that follow:

1 - Aspartokinase
2 - Aspartate-semialdehyde dehydrogenase
3 - Homoserine dehydrogenase
4 - Homoserine O-transsuccinylase
5 - Cystathionine-γ-synthase
6 - Cystathionine-β-lyase
7 - Methionine synthase

Though humans cannot make methionine by the pathway shown in the figure, they can recycle methionine from homocysteine (a product of S-adenosylmethionine metabolism). This reaction requires the enzyme methionine synthase and Vitamin B12 as a co-factor.

An alternative pathway of converting homocysteine to methionine involves a prominent liver enzyme, betaine-homocysteine methyltransferase. This enzyme catalyzes the reaction below.

In this reaction, a methyl group is transferred to homocysteine from glycine betaine to make the methionine. Glycine betaine is a trimethylated amine of glycine found in plants. It is a byproduct of choline metabolism.

Bacteria, mitochondria, and chloroplasts use a modified form of methionine, N-formyl-methionine (Figure 6.142), as the first amino acid incorporated into their proteins. Formylation of methionine occurs only after methionine has been attached to its tRNA for translation. Addition of the formyl group is catalyzed by the enzyme methionyl-tRNA formyltransferase

Threonine

Though threonine is chemically similar to serine, the metabolic pathway leading to threonine does not overlap with that of serine. As seen in the figure, aspartate is a starting point for synthesis. Two phosphorylations/dephosphorylations and two reductions with electrons from NADPH result in production of threonine.

Enzymes in Figure 6.143 are as follows:

1 Aspartokinase
2 β-aspartate semialdehyde dehydrogenase
3 Homoserine dehydrogenase
4 Homoserine kinase
5 Threonine synthase

Breakdown of threonine produces acetyl-CoA and glycine. It can also produce α-ketobutyrate, which can be converted to succinyl-CoA for oxidation in the citric acid cycle.

Lysine

To get from aspartate to lysine, nine reactions and two non-enzymatic steps are involved, as seen in Figure 6.144. Enzymes involved in lysine biosynthesis include (numbers correspond to numbered reactions in Figure 6.144):

1 - Aspartokinase
2 - Aspartate-semialdehyde dehydrogenase
3 - 4-hydroxy-tetrahydrodipicolinate synthase
4 - 4-hydroxy-tetrahydrodipicolinate reductase
5 - 2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
6 - Succinyl-diaminopimelate transaminase

7 - Succinyl-diaminopimelate desuccinylase
8 - Diaminopimelate epimerase
9 - Diaminopimelate decarboxylase

Low in cereal grains

Lysine is the essential amino acid found in the smallest quantity in cereal grains, but is found abundantly in legumes. Besides its synthesis and breakdown, lysine can be methylated, acetylated, hydroxylated, ubiquitinated, sumoylated, neddylated, biotinylated, pupylated, and carboxylated within proteins containing it. Hydroxylation of lysine is important for strengthening collagen and acetylation/methylation of lysine in histone proteins play roles in control of gene expression and epigenetics. Besides being used to make proteins, lysine is important for calcium absorption, recovery from injuries, and for production of hormones.

Oral lysine has been used as a treatment for herpes infections (cold sores) but its efficacy is not established and it is not clear by what mechanism is would reduce the duration of the infection or reduce the number of outbreaks of viral infection..

Aromatic amino acids

The aromatic amino acids, tryptophan, phenylalanine, and tyrosine can all be made starting with two simple molecules - PEP and erythrose-4-phosphate (Figure 6.145). All three aromatic amino acids are also important sources of hormones, neurotransmitters, and even the skin pigment melanin.

Tryptophan synthesis

The proteogenic amino acid with the largest R-group, tryptophan is an essential amino acid distinguished structurally by its indole group. The amino acid is made in bacteria and plants from shikimic acid or anthranilate and serine is used in its synthesis.

Erythrose-4-phosphate and phosphenolpyruvate (PEP) also serve as building blocks of tryptophan. The pathway of its synthesis is shown in Figures 6.146 to 6.148.

Erythrose-4-phosphate and phosphoenolpyruvate (PEP) are joined and then, after one hydrolysis, one dehydration, one oxidation and one reduction, the product is shikimic acid (Figure 6.147).

Shikimic acid is converted to chorismic acid in three steps, as shown in Figure 6.147. Finally, synthesis of tryptophan from chorismic acid is shown in Figure 6.148.

Regulation

Regulation of tryptophan synthesis in bacteria occurs partly via a process called attenuation that operates through the trp operon. In this mechanism, low levels of tryptophan slow ribosomal movement (and translation) through the operon. This is particularly important because bacteria can have transcription and translation occurring simultaneously. Slowing translation due to low tryptophan levels allows a transcription termination mechanism to be inhibited. Since translation only slows when tryptophan is in short supply, premature termination of transcription occurs when tryptophan is abundant (see also HERE).

Besides its importance for making proteins, tryptophan is an important precursor of serotonin (neurotransmitter), melatonin (hormone), niacin (vitamin), and auxin (plant hormone). The two pathways leading from tryptophan to three of these molecules is shown in Figure 6.149.

Melatonin

Melatonin is a compound made from tryptophan that is found in a wide spectrum of biological systems, including plants, animals, fungi, and bacteria. In animals, it acts as a hormone for circadian rhythm synchronization, signaling the onset of darkness each day. It has effects on the timing of sleep, seasonal effects, and can affect blood pressure, among other physiological phenomena. It can cross cell membranes, as well as the blood-brain barrier. Melatonin is a potent anti-oxidant and provides protective functions for nucleic acids. It is used sometimes to help in treatment of sleep disorders. Some reports have indicated that children with autism have abnormal melatonin pathways with low levels of the hormone.

Blue light

Melatonin production is affected by blue light and may be linked to sleep abnormalities for people using computer monitors after dark. To protect against this, some computer programs are available that reduce the screen's blue light output in the evenings. Special eyeglasses that block blue light are also available. Though melatonin is linked to sleep in some animals (including humans), nocturnal animals are activated by increasing melatonin levels. Varying day/night lengths during the year alter melatonin production and provide biological signals of the seasons. These are especially important in the seasonal coloring and breeding habits of some animals. Melatonin is present in cherries, bananas, grapes, rice, cereals, olive oil, wine, and beer.

Serotonin

Serotonin, or 5-hydroxytryptamine, is a monoamine neurotransmitter derived from tryptophan. Blood platelets store serotonin and release it when they bind to a clot, causing vasoconstriction. Serotonin plays a role in cognitive functions and enhances memory and learning. Serotonin is widely thought to be a contributor to feelings of happiness and well-being. Some common anti-depressant drugs, including Prozac, Paxil and Zoloft, act to modulate action of serotonin at synapses.

Niacin

Niacin is also known as Vitamin B3 and nicotinic acid. Niacin can be made from tryptophan and people who have the inability to absorb tryptophan in the digestive system exhibit symptoms similar to niacin deficiency.

Extreme deficiency of niacin in the diet leads to the disease known as pellagra, while insufficient amounts of niacin in the diet are linked with nausea, anemia, headaches, and tiredness. A diet that is primarily composed of grains like corn can lead to niacin deficiency, because the niacin in these sources is not readily bioavailable. Treatment of the grain with alkali, as in the traditional Mexican practice of soaking corn in lime, can make the niacin more easily absorbed from food.

Niacin is related to pyridine and the amide form of it is nicotinamide, an important component of NAD+/NADH and NADP+/NADPH. The last pairs of molecules are essential as electron acceptors/carriers for most cellular oxidation-reduction reactions.

Auxins

Auxins are plant growth hormones derived from tryptophan. The most important of these is indole-3-acetic acid (Figure 6.151). Auxins are involved in almost every aspect of plant growth and development. They activate proteins, such as expansins and various enzymes that modify the structure of cell wall components, to loosen the cell walls of a plant and stimulate elongation of cells. In the presence of cytokinins, auxins stimulate cell division. Auxins are also involved in the maintenance of meristems and in cell patterning and organogenesis. Auxins are crucial for establishing root primordia as well as for elongation of root hairs. Auxins play important roles in organizing the xylem and phloem of plants, and it has long been known that plant callus tissue can be made to differentiate into shoots or roots, depending on the relative concentrations of auxins and cytokinins supplied in the medium.

Agrobacterium tumefaciens, a bacterium which infects a wide variety of plants, inserts its own DNA, including genes necessary for the synthesis of plant hormones, into its host's cells. The subsequent overproduction of auxins stimulates the growth of tumors (called crown galls) on the plant (Figure 6.153).

Phenylalanine

Phenylalanine is an essential, hydrophobic amino acid in humans that is a precursor of tyrosine and since tyrosine is a precursor of several important catecholamines, phenylalanine is, thus, a precursor of them as well.

PKU

Phenylalanine is linked to the genetic disease phenylketonuria (PKU) which arises from an inability to metabolize the amino acid in people lacking (or deficient in) the enzyme phenylalanine hydroxylase. If left untreated, the disease can cause brain damage and even death, but if detected early, it can be easily managed by carefully monitoring dietary intake of the amino acid. Because of this, newborns are routinely tested for PKU. Phenylalanine is a component of the artificial sweetener known as aspartame (Nutrasweet - Figure 6.154) and is consequently dangerous for people suffering from this disorder.

Biosynthesis of phenylalanine in bacteria overlaps with synthesis of tryptophan. The branch occurs at chorismic acid where the enzyme chorismate mutase catalyzes a molecular rearrangement to produce prephenate.

Proton attack on prephenate results in loss of water and carbon dioxide to yield phenylpyruvate.

Transamination of phenylpyruvate yields phenylalanine.

Alternatively, phenylalanine can obtain its amine group in a transamination reaction from alanine.

Hydroxylation of phenylalanine by aromatic amino acid hydroxylase (phenylalanine hydroxylase) yields tyrosine.

Tyrosine

Because tyrosine is made from phenylalanine and the latter is an essential amino acid in humans, it is not clear whether to classify tyrosine as essential or non-essential. Some define it as a conditionally essential amino acid. Others simply categorize it as non-essential.

As noted above, tyrosine can arise as a result of hydroxylation of phenylalanine. In addition, plants can synthesize tyrosine by oxidation of prephenate followed by transamination of the resulting 4-hydroxyphenylpyruvate (Figure 6.155).

The hydroxyl group on tyrosine is a target for phosphorylation by protein kinase enzymes involved in signal transduction pathways (Figure 6.156). When located in membranes, these enzymes are referred to as receptor tyrosine kinases and they play important roles in controlling cellular behavior/response.

In photosystem II of chloroplasts, tyrosine, at the heart of the system, acts as an electron donor to reduce oxidized chlorophyll. The hydrogen from the hydroxyl group of tyrosine is lost in the process, requiring re-reduction by four core manganese clusters.

Tyrosine is also important in the small subunit of class I ribonucleotide reductases where it forms a stable radical in the catalytic action of the enzyme (see HERE).

Tyrosine metabolites

Tyrosine is a precursor of catecholamines, such as L-dopa, dopamine, norepinephrine, and epinephrine (Figure 6.157). The thyroid hormones triiodothyronine (T3) and thyroxine (T4) are also synthesized from tyrosine. As shown in Figure 6.158, this involves a series of iodinations of tyrosines side-chains of a protein known as thyroglobulin. Combinations of iodinated tyrosines give rise to thyroxine and triiodothyronine. These are subsequently cleaved from the protein and released into the bloodstream.

Oxidation and polymerization of tyrosine is involved in synthesis of the family of melanin pigments. Tyrosine is involved in the synthesis of at least two types - eumelanin and pheomelanin (Figure 6.159).

Another molecule derived from tyrosine is the benzoquinone portion of Coenzyme Q (CoQ). This pathway requires the enzyme HMG-CoA Reductase and since this enzyme is inhibited by cholesterol-lowering statin drugs, CoQ can be limited in people being treated for high cholesterol levels.

Dopamine

Dopamine plays several important roles in the brain and body. A member of the catecholamine and phenethylamine families, its name comes from the fact that it is an amine made by removing a carboxyl group from L-DOPA. Dopamine is synthesized in the brain and kidneys. It is also made in plants, though its function in plants is not clear. Conversion of dopamine to norepinephrine (Figure 6.157) requires vitamin C.

Dopamine is a neurotransmitter, being released by one nerve cell and then traveling across a synapse to signal an adjacent nerve cell. Dopamine plays a major role in the brain's reward-mediated behavior. Rewards, such as food or social interaction, increase dopamine levels in the brain, as do addictive drugs. Other brain dopamine pathways are involved in motor control and in managing the release of various hormones.

Chemical messenger

Outside the nervous system, dopamine is a local chemical messenger. In blood vessels, it inhibits norepinephrine release and causes vasodilation. In the kidneys, it increases sodium excretion and urine output. It reduces gastrointestinal motility and protects intestinal mucosa in the digestive system and in the immune system, it reduces lymphocyte activity. The effect dopamine has on the pancreas is to reduce insulin production. With the exception of the blood vessels, dopamine is synthesized locally and exerts its effects near the cells that release it.

Epinephrine

Epinephrine (also called adrenalin) is a catecholamine chemically related to norepinephrine that is a hormone with medical applications. It is used to treat anaphylaxis, cardiac arrest, croup, and, in some cases, asthma, when other treatments are not working, due to its ability to favor bronchodilation.

Epinephrine is the drug of choice for treating anaphylaxis. The compound may be given through inhalation, by intravenous injection, or subcutaneous injection and exerts effects through the α- and β-adrenergic receptors. In the body, it is produced and released by adrenal glands and some neurons.

Effects

Physiological effects of epinephrine may include rapid heart beat, increased blood pressure, heart output, pupil dilation, blood sugar concentration and increased sweating. Other physical effects may include shakiness, increased anxiety, and an abnormal heart rhythm.

Norepinephrine

Norepinephrine (also called noradrenalin) is a catecholamine molecule that acts as a hormone and neurotransmitter. It is chemically similar to epinephrine, differing only in the absence of a methyl group on its amine. Norepinephrine is made and released by the central nervous system (locus coeruleus of the brain) and the sympathetic nervous system. The compound is released into the blood stream from adrenal glands and affects α- and β-adrenergic receptors.

Norepinephrine is at its lowest levels during sleep and at its highest levels during stress (fight or flight response). The primary function of norepinephrine is to prepare the body for action. It increases alertness, enhances memory functions, and helps to focus attention. Norepinephrine increases heart rate and blood pressure, increases blood glucose and blood flow to skeletal muscle and decreases flow of blood to the gastrointestinal system.

Medical considerations

Norepinephrine may be injected to overcome critically low blood pressure and drugs countering its effects are used to treat heart conditions. α-blockers, for example, are used to battle cardiovascular and psychiatric disorders. β-blockers counter a different set of norepinephrine's effects than α-blockers and are used to treat glaucoma, migraine headaches and other cardiovascular problems.

Pyruvate family

The family of amino acids derived from pyruvate has four members, each with a simple aliphatic side chain no longer than four carbons. The simplest of these is alanine.

Alanine

Alanine is the amino acid that is most easily produced from pyruvate. The simple transamination catalyzed by alanine transaminase produces alanine from pyruvate.

Alternative pathways for synthesis of alanine include catabolism of valine, leucine, and isoleucine.

Glucose-alanine cycle

The glucose-alanine cycle is an important nitrogen cycle related to the Cori cycle that occurs between muscle and liver cells in the body (see HERE). In it, breakdown of glucose in muscles leads to pyruvate. When nitrogen levels are high, pyruvate is transaminated to alanine, which is exported to hepatocytes.

In the liver cells, the last transamination of the glucose-alanine cycle occurs. The amine group of alanine is transferred to α-ketoglutarate to produce pyruvate and glutamate. Glucose can then be made by gluconeogenesis from pyruvate. Importantly, breakdown of glutamate yields ammonium ion, which can be made into urea for excretion, thus reducing the body's load of potentially toxic amines. This pathway may be particularly important in the brain.

Another way of removing excess ammonium from a tissue is by attaching it to glutamate to make glutamine. Glutamate is a neurotransmitter, so having an alternative way of removing amines (glucose-alanine cycle) is important, especially in the brain.

Leucine

Like valine and isoleucine, leucine is an essential amino acid in humans. In adipose tissue and muscle, leucine is used in sterol synthesis. It is the only amino acid to stimulate muscle protein synthesis, and as a dietary supplement in aged rats, it slows muscle degradation. Leucine is an activator of mTOR, a protein which, when inhibited, has been shown to increase life span in Saccharomyces cerevisiae, C. elegans, and Drosophila melanogaster.

Metabolism of leucine, valine, and isoleucine (also called Branched Chain Amino Acids - BCAAs) starts with decarboxylation of pyruvate and attachment of the two-carbon hydroxyethyl fragment to thiamine pyrophosphate (Figure 6.161). Metabolism of isoleucine proceeds with attachment of the hydroxylated two carbon piece (hydroxyethyl-TPP) to α-ketobutyrate and is covered in the section describing that amino acid (see HERE).

Metabolism of valine and leucine proceeds with attachment of the hydroxyethyl piece from TPP to another pyruvate to create α-acetolactate. Rearrangement of α-acetolactate by acetolactate mutase makes 3-hydroxy-3-methyl-2-oxobutanoate.

Reduction with NAD(P)H by acetohydroxy acid isomeroreductase yields α,β-dihydroxyisovalerate.

Loss of water, catalyzed by dihydroxyacid dehydratase produces α-ketoisovalerate.

This molecule is a branch point for synthesis of leucine and valine. Addition of an acetyl group from acetyl-CoA yields α-isopropylmalate (catalyzed by α-isopropylmalate synthase).

Rearrangement, catalyzed by isopropylmalate dehydratase, gives rise to β-isopropylmalate.

Oxidation by isopropylmalate dehydrogenase and NAD+, gives α-ketoisocaproate.

Transamination of it (catalyzed by leucine aminotransferase and using glutamate) gives the final product of leucine (top of next column).

Valine

An essential amino acid in humans, valine is derived in plants from pyruvate and shares part of its metabolic synthesis pathway with leucine and a small slice of it with isoleucine. Metabolism of all three amino acids starts with decarboxylation of pyruvate and attachment of the two-carbon hydroxyethyl fragment to thiamine pyrophosphate (Figure 6.161), as noted above.

As seen earlier, α-ketoisovalerate is the molecule at the point in the metabolic pathway where synthesis of valine branches from that of leucine. In fact, α-ketoisovalerate is only one step away from valine. Transamination of α-ketoisovalerate catalyzed by valine isoleucine aminotransferase gives valine.

Isoleucine

Synthesis of isoeleucine (an essential amino acid in humans) begins in plants and microorganisms with pyruvate and α-ketobutyrate (a byproduct of threonine metabolism - threonine deaminase - Figure 6.162).

Metabolism of isoleucine proceeds with attachment to α-ketobutyrate of the hydroxyethyl-TPP product of pyruvate decarboxylation to form α-aceto-α-hydroxybutyrate. The reaction is catalyzed by acetolactate synthase. Rearrangement and reduction by acetohydroxy acid isomeroreductase and NAD(P)H yields α,β-dihydroxy-β-methylvalerate. Shown on next page.

Loss of water (catalyzed by dihydroxy acid dehydratase) gives α-keto-β-methylvalerate.

Transamination (using glutamate and valine isoleucine transaminase) yields isoleucine.

Interestingly, several of the enzymes of valine metabolism catalyze reactions in the isoleucine pathway. Though the substrates are slightly different, they are enough like the valine intermediates that they are recognized as substrates.

Isoleucine has a second asymmetric center within it, but only one isomeric form of the four possible ones from the two centers is found biologically.

Regulation of synthesis

Regulation of synthesis of the branched chain amino acids (BCAAs - valine, leucine, and isoleucine) is complex. The key molecule in the regulation is α-ketobutyrate, which is synthesized in cells as a breakdown product of threonine. The enzyme catalyzing its synthesis is threonine deaminase (Figure 6.162), which is allosterically regulated. The enzyme is inhibited by its own product (isoleucine) and activated by valine, a product of a parallel pathway.

Thus, when valine concentration is high, the balances shifts in favor of production of isoleucine and since isoleucine competes with valine and leucine for hydroxyethyl-TPP, synthesis of these two amino acids goes down. When isoleucine concentration increases, threonine deaminase is inhibited, shifting the balance back to production of valine and leucine.

Attenuation

Another control mechanism for regulation of leucine synthesis occurs in bacteria and is known as attenuation. In this method, accumulation of leucine speeds the process of translation of a portion of the mRNA copy of the leucine operon (coding sequences for enzymes necessary to make leucine). This, in turn, causes transcription of the genes of the leucine operon to terminate prematurely, thus stopping production of the enzymes necessary to make leucine.

When leucine levels fall, translation slows, preventing transcription from terminating prematurely and allowing leucine metabolic enzymes to be made. Thus, leucine levels in the cell control the synthesis of enzymes necessary to make it.

Histidine family

Synthesis of histidine literally occurs in a class by itself - there are no other amino acids in its synthesis family. The amino acid is made in plants (Arabidopsis, in this case) by a pathway that begins with ribose-5-phosphate. The overall pathway is show in the green text boxes on the next two pages. Abbreviations used in the boxes are shown below.

Enzyme names

1 = Ribose-phosphate diphosphokinase
2= ATP-phosphoribosyltransferase
3 = Phosphoribosyl-ATP pyrophospohydrolase
4 = Phosphoribosyl-AMP cyclohydrolase
5 = ProFAR-I (N'-[(5'phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide isomerase)
6 = Imidazole glycerol-phosphate synthase (IGPS)
7 = Imidazole glycerol-phosphate dehydratase
8 = Histidinol-phosphate aminotransferase
9 = Histidinol-phosphate phosphatase
10 = Histidinol dehydrogenase

Abbreviations used

1 - PRPP = Phosphoribosyl Pyrophosphate
2. PRATP = Phosphoribosyl ATP
3. PRAMP = Phosphoribosyl AMP
4. ProFAR = (N′-[(5′-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide) ribonucleotide
5. PRFAR = (N′-[(5-phosphoribulosyl)formimino]-5-aminoimidazole-4-carboxamide) ribonucleotide
6. IGP = Imidazole glycerol-phosphate
7. AICAR = 5′-phosphoribosyl-4-carboximide-5-aminoimidazole
8. IAP = Imidazole acetol-phosphate
9. α-KG = α-ketoglutarate

Histidine is a feedback inhibitor of ATP-phosphoribosyltransferase and thus helps to regulate its own synthesis. Histidine is the only amino acid to contain an imidazole ring. It is ionizable and has a pKa of about 6. As a result, histidine's R-group can gain/lose a proton at pH values close to cellular conditions.

Selenocysteine

A cysteine analog commonly referred to as the 21st amino acid, selenocysteine (Figure 6.163) is an unusual amino acid

occasionally found in proteins. Although it is rare, selenocysteine has been found in proteins in bacteria, archaea and eukaryotes.

In contrast to amino acids such as phosphoserine, hydroxyproline, or acetyl-lysine, which arise as a result of post-translational modifications, selenocysteine is actually built into growing peptide chains in ribosomes during the process of translation.

No codon specifies selenocysteine, so to incorporate it into a protein, a tRNA carrying it must bind to a codon that normally specifies STOP (UGA). This alternative reading of the UGA is dependent on formation of a special hairpin loop structure in the mRNA encoding selenoproteins.

Selenium is rather toxic, so cellular and dietary concentrations are typically exceedingly low. About 25 human proteins are known to contain the amino acid. These include five glutathione peroxidases, and three thioredoxin reductases. Iodothyronine deiodinase, a key enzyme that converts thyroxine to the active T3 form, also contains selenocysteine in its active site. All of these proteins contain a single selenocysteine.

A eukaryotic protein known as selenoprotein P, found in the blood plasma of animals, contains ten selenocysteine residues and is thought to function as an antioxidant and/or in heavy metal detoxification. Besides selenocysteine, at least two other biological forms of a seleno-amino acid are known. These include 1) selenomethionine (Figure 6.164), a naturally occurring amino acid in Brazil nuts, cereal grains, soybeans, and grassland legumes and 2) methylated forms of selenocysteine, such as Se-methylselenocysteine, are found in Astragalus, Allium, and Brassica species.

Stop codon

The specifics of the process of translation will be described elsewhere in the book, but to get selenocysteine into a protein, the tRNA carrying selenocysteine pairs with a stop codon (UGA) in the mRNA in the ribosome. Thus, instead of stopping translation, selenocysteine can incorporated into a growing protein and translation continues instead of stopping.

Four genes are involved in preparation of selenocysteine for incorporation into proteins. They are known as sel A, sel B, sel C, and sel D. Sel C codes for the special tRNA that carries selenocysteine. The amino acid initially put onto the selenocysteine tRNA is not selenocysteine, but rather serine. Action of sel A and sel D are necessary to convert the serine to a selenocysteine.

An intermediate in the process is selenophosphate, which is the selenium donor. It is derived from H2Se, the form in which selenium is found in the cell. The tRNA carrying selenocysteine has a slightly different structure than other tRNAs, so it requires assistance in translation. The sel B gene encodes for an EF-Tu-like protein that helps incorporate the selenocysteine into the protein during translation.

Recoding the UGA

Using UGA codons to incorporate selenocysteine into proteins could wreak havoc if done routinely, as UGA, in fact, almost always functions as a stop codon and is only rarely used to code for selenocysteine. Fortunately, there is a mechanism to ensure that the reading of a UGA codon as selenocysteine occurs only when the mRNA encodes a selenoprotein.

Unusual structures in mRNAs

The mRNAs for selenocysteine-containing proteins form unusual mRNA structures around the UGA codon that make the ribosome “miss” it as a stop codon and permit the tRNA with selenocysteine to be incorporated instead.

Pyrrolysine

Like selenocysteine, pyrrolysine is a rare, unusual, genetically encoded amino acid found in some cells. Proteins containing it are enzymes involved in methane metabolism and so far have been found only methanogenic archaeans and one species of bacterium. The amino acid is found in the active site of the enzymes containing it. It is sometimes referred to as the 22nd amino acid.

Synthesis of the amino acid biologically begins with two lysines. One is converted to (3R)-3-Methyl-D-ornithine, which is attached to the second lysine. After elimination of an amine group, cyclization, and dehydration, L-pyrrolysine is produced. Pyrrolysine is attached to an unusual tRNA (pylT gene product) by action of the aminoacyl tRNA synthetase encoded by the pylS gene. This unusual tRNA can pair with the UAG stop codon during translation and allow for incorporation of pyrrolysine into the growing polypeptide chain during translation in a manner similar to incorporation of selenocysteine.

Urea cycle

The urea cycle holds the distinction of being the first metabolic cycle discovered - in 1932, five years before the citric acid cycle. It is an important metabolic pathway for balancing nitrogen in the bodies of animals and it takes place primarily in the liver and kidney.

Organisms, like humans, that excrete urea are called ureotelic. Those that excrete uric acid (birds, for example) are called uricotelic and those that excrete ammonia (fish) are ammonotelic. Ammonia, of course, is generated by metabolism of amines and is toxic, so managing levels of it is critical for any organism. Excretion of ammonia by fish is one reason that an aquarium periodically requires cleaning and replacement of water.

Liver failure can lead to accumulation of nitrogenous waste and exacerbates the problem. As shown in Figure 1.166, the cycle contains five reactions, with each turn of the cycle producing a molecule of urea. Of the five reactions, three occur in the cytoplasm and two take place in the mitochondrion. (The reaction making carbamoyl phosphate, catalyzed by carbamoyl phosphate synthetase is not shown in the figure.)

Ornithine synthesis

Though the cycle doesn't really have a starting point, a common place to begin discussion is with the molecule of ornithine. As discussed elsewhere in this book, ornithine intersects the metabolic pathways of arginine and proline.

Ornithine is found in the cytoplasm and is transported into the mitochondrion by the ornithine-citrulline antiport of the inner mitochonrial membrane. In the matrix of the mitochondrion, two reactions occur relevant to the cycle. The first is formation of carbamoyl phosphate from bicarbonate, ammonia, and ATP catalyzed by carbamoyl phosphate synthetase I.

Carbamoyl phosphate then combines with ornithine in a reaction catalyzed by ornithine transcarbamoylase to make citrulline.

The citrulline is transported out to the cytoplasm by the ornithine-citrulline antiport mentioned above. In the cytoplasm, citrulline combines with L-aspartate using energy of ATP to make citrullyl-AMP (an intermediate) followed by argininosuccinate. The reaction is catalyzed by argininosuccinate synthase.

Next, fumarate is split from argininosuccinate by argininosuccinate lyase to form arginine.

Water is used by arginase to cleave arginine into urea and ornithine, completing the cycle.

Urea is less toxic than ammonia and is released in the urine. Some organisms make uric acid for the same reason.

It is worth noting that aspartic acid, ammonia, and bicarbonate enter the cycle and fumarate and urea are produced by it. Points to take away include 1) ammonia is converted to urea using bicarbonate and the amine from aspartate; 2) aspartate is converted to fumarate which releases more energy than if aspartate were converted to oxaloacetate, since conversion of fumarate to malate to oxaloacetate in the citric acid cycle generates an NADH, but direct conversion of aspartate to oxaloacetate does not; and 3) glutamate and aspartate are acting as shuttles to funnel ammonia into the cycle. Glutamate, as will be seen below, is a scavenger of ammonia.

Urea cycle regulation

The urea cycle is controlled both allosterically and by substrate concentration. The cycle requires N-acetylglutamate (NAG) for allosteric activation of carbamoyl phosphate synthetase I. The enzyme that catalyzes synthesis of NAG, NAG synthetase, is activated by arginine and glutamate. Thus, an indicator of high amine levels, arginine, and an important shuttler of amine groups, glutamate, stimulates the enzyme that activates the cycle.

The reaction catalyzed by NAG synthetase is

At the substrate level, all of the other enzymes of the urea cycle are controlled by the concentrations of substrates they act upon. Only at high concentrations are the enzymes fully utilized.

Complete deficiency of any urea cycle enzyme is fatal at birth, but mutations resulting in reduced expression of enzymes can have mixed effects. Since the enzymes are usually not limiting for these reactions, increasing substrate can often overcome reduced enzyme amounts to a point by simply fully activating enzymes present in reduced quantities.

Ammonia accumulation

However, if the deficiencies are sufficient, ammonium can accumulate and this can be quite problematic, especially in the brain, where mental deficiencies or lethargy can result. Reduction of ammonium concentration relies on the glutamate dehydrogenase reaction (named for the reverse reaction).

Additional ammonia can be taken up by glutamate in the glutamine synthetase reaction.

The result of these reactions is that α-ketoglutarate and glutamate concentrations will be reduced and the concentration of glutamine will increase. For the brain, this is a yin/yang situation. Removal of ammonia is good, but reduction of α-ketoglutarate concentration means less energy can be generated by the citric acid cycle. Further, glutamate is, itself, an important neurotransmitter and a precursor of another neurotransmitter - γ-aminobutyric acid (GABA).

Energy generation

From an energy perspective, the urea cycle can be said to break even or generate a small amount of energy, if one includes the energy produced in releasing ammonia from glutamate (one NADH). There are two NADHs produced (including the one for converting fumarate to oxaloacetate), which give 4-6 ATPs, depending on how efficiently the cell performs electron transport and oxidative phosphorylation.

The cycle takes in 3 ATPs and produces 2 ADPs and one AMP. Since AMP is equivalent to 2 ATP, the cycle uses 4 ATP. Thus, the cycle either breaks even in the worst case or generates 2 ATPs in the best case.

Amino acid catabolism

Amino acids are divided according to the pathways involved in their degradation. There are three general categories. Ones that yield intermediates in the glycolysis pathway are called glucogenic and those that yield intermediates of acetyl-CoA or acetoacetate are called ketogenic. Those that involve both are called glucogenic and ketogenic. These are shown in Figures 6.167 and 6.168.

As seen in the two figures, amino acids largely produce breakdown products related to intermediates of the citric acid cycle or glycolysis, but this isn't the complete picture. Some amino acids, like tryptophan, phenylalanine, and tyrosine yield hormones or neurotransmitters on further metabolism (as noted earlier). Others like cysteine and methionine must dispose of their sulfur and all of the amino acids must rid themselves of nitrogen, which can happen via the urea cycle, transamination, or both.

Tyrosine catabolism

Breakdown of tyrosine (Figure 6.169) is a five step process that yields acetoacetate and fumarate. Enzymes involved include 1) tyrosine transaminase; 2) p-hydroxylphenylpyruvate dioxygenase; 3) homogentisate dioxygenase; 4) maleylacetoacetate cis-trans-isomerase; and 5) 4-fumaryl acetoacetate hydrolase.

Breakdown of leucine is a multi-step process ultimately yielding the ketone body acetoacetate and acetyl-CoA. Branched chain amino acids (BCAAs - valine, leucine, and isoleucine) rely on Branched Chain AminoTransferase (BCAT) followed by Branched Chain α-ketoacid dehydrogenase (BCKD) for catabolism.

Breakdown of isoleucine yields intermediates that are both ketogenic and glucogenic. These include acetyl-CoA and propionyl-CoA.

Breakdown of valine is a multi-step process ultimately yielding propionyl-CoA.

This page titled 6.5: Amino Acids and the Urea Cycle is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

6.6: Nucleotides

Source: BiochemFFA_6_6.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Diverse functions of nucleotides

Nucleotides are most often thought of as the building blocks of the nucleic acids, DNA and RNA. While this, is, of course, a vital function, nucleotides also play other important roles in cells. Ribonucleoside triphosphates like ATP, CTP, GTP and UTP are necessary, not just for the synthesis of RNA, but as part of activated intermediates like UDP-glucose in biosynthetic pathways. ATP is also the universal “energy currency” of cells, and coupling of energetically unfavorable reactions with the hydrolysis of ATP makes possible the many reactions in our cells that require an input of energy. Adenine nucleotides serve as components of NAD(P)+ and FAD. Nucleotides can also serve as allosteric and metabolic regulators. The synthesis and breakdown pathways for nucleotides and the molecules derived from them are thus, of vital importance to cells. Regulation of nucleotide synthesis, especially for deoxyribonucleotides, is important to ensure that the four nucleotides are made in the right proportions, as imbalances in nucleotide concentrations can lead to increases in mutation rates.

Pathways of nucleotide metabolism are organized in two major groups and one minor one. These include, respectively, metabolism of 1) purines; 2) pyrimidines; and 3) deoxyribonucleotides. Each group can be further subdivided into pathways that make nucleotides from simple precursors (de novo pathways) and others that use pieces of nucleotides to reassemble full ones (salvage pathways). Notably, de novo synthesis pathways for all of the nucleotides begin with synthesis of ribonucleotides. Deoxyribonucleotides are made from the ribonucleotides.

Purine nucleotide metabolism

Synthesis of purine nucleotides by the de novo pathway begins with addition of a pyrophosphate to carbon 1 of ribose-5-phosphate, creating phosphoribosylpyrophosphate (PRPP). The reaction is catalyzed by PRPP synthetase. Some number the purine metabolic pathway starting with the next reaction. We have therefore given this reaction the number of zero in Figure 6.172.

In the next step (reaction 1 in Figure 6.172), the pyrophosphate is replaced by an amine from glutamine in a reaction catalyzed by PRPP amidotransferase (PPAT). The product is 5-phosphoribosylamine (5-PRA).

PPAT is an important regulatory enzyme for purine biosynthesis. The end products of the pathway, AMP and GMP both inhibit the enzyme and PRPP activates it. Interestingly, full inhibition of the enzyme requires binding of both AMP and GMP.

Binding of only one of the two nucleotides allows the enzyme to remain partially active so that the missing nucleotide can be synthesized. Through this enzyme, the relative amounts of ATP and GTP are controlled.

5-PRA is very unstable chemically (half-life of 38 seconds at 37°C), so it has been proposed that it is shuttled directly from PRPP amidotransferase to GAR synthetase for the next reaction.

In this reaction (#2), glycine is added to the growing structure above the ribose-5-phosphate to create glycineamide ribonucleotide (GAR). This reaction, which requires ATP, is catalyzed, as noted, by the enzyme GAR synthetase.

In reaction #3, a formyl group is transferred onto the GAR from N10-formyl-tetrahydrofolate (N10-formyl-THF or fTHF) by phosphoribosylglycinamide formyltransferase (GART).

Next, the double bonded oxygen in the ring is replaced with an amine in a reaction catalyzed by phosphoribosylformylglycinamidine synthase (PFAS) that uses glutamine and produces glutamate. The reaction requires energy from ATP (top of next column).

In humans the GAR synthetase, phosphoribosylglycinamide formyltransferase, and the enzyme catalyzing the next reaction (#5), AIR synthetase activities are all on the same protein known as trifunctional purine biosynthetic protein adenosine-3.

In reaction #6, carboxylation of AIR occurs, catalyzed by phosphoribosylaminoimidazole carboxylase (PAIC)

Aspartic acid is then added to donate its amine group and fumarate will be lost in the reaction that follows this one. The enzyme involved here is phosphoribosyl-aminoimidazole-succinocarboxamide synthase (PAICS)

In the next reaction, the carbon shell of aspartate is released (as fumarate) and the amine is left behind. The reaction is catalyzed by adenylosuccinate lyase (ADSL).

Reaction #9 involves another formylation reaction, catalyzed by phosphoribosylaminoimidazolecarboxamide formyltransferase (ATIC-E1).

Next, inosine monophosphate synthase (ATIC-E2) catalyzes release of water to form the first molecule classified as a purine - inosine monophosphate or IMP).

Though it doesn't appear in DNA, IMP does, in fact, occur in the anticodon of many tRNAs where its ability to pair with numerous bases is valuable in reading the genetic code.

IMP is a branch point between pathways that lead to GMP or AMP. The pathway to GMP proceeds via catalysis by IMP dehydrogenase as follows:

In the last step of GMP synthesis, GMP synthase catalyzes a transamination to form GMP using energy from ATP.

The energy source being ATP makes sense, since the cell is presumably making GMP because it needs guanine nucleotides. If the cell is low on guanine nucleotides, GTP would be in short supply.

Adenine nucleotide synthesis

Synthesis of AMP from IMP follows. First, adenylosuccinate synthetase catalyzes the addition of aspartate to IMP, using energy from GTP.

Then, adenylosuccinate lyase splits fumarate off to yield AMP.

In humans, the bifunctional purine biosynthesis protein known as PURH contains activities of the last two enzymes above.

Abbreviations used above

- PRPP = Phosphoribosyl Pyrophosphate
- 5-PRA = 5-phosphoribosylamine
- GAR = glycineamide ribonucleotide
- fGAR = Phosphoribosyl-N-formylglycineamide
- THF = Tetrahydrofolate
- fTHF = N10-formyl-Tetrahydrofolate
- fGAM = 5'-Phosphoribosylformylglycinamidine
- AIR = 5-Aminoimidazole ribotide
- CAIR = 5'-Phosphoribosyl-4-carboxy-5-aminoimidazole
- SAICAR = Phosphoribosylamino-imidazolesuccinocarboxamide
- AICAR = 5-Aminoimidazole-4-carboxamide ribonucleotide
- FAICAR = 5-Formamidoimidazole-4-carboxamide ribotide
- IMP = inosine monophosphate

Regulation

It is worth repeating that synthesis of GMP from IMP requires energy from ATP and that synthesis of AMP from IMP requires energy from GTP. In addition, the enzymes converting IMP into intermediates in the AMP and GMP pathways are each feedback inhibited by the respective monophosphate nucleotide. Thus, IMP dehydrogenase is inhibited by GMP (end product of pathway branch) and adenylosuccinate synthetase is inhibited by AMP, the end product of that pathway branch.

Purine nucleotide levels are balanced by the combined regulation of PRPP amidotransferase , IMP dehydrogenase,

adenylosuccinate synthetase and the nucleotides AMP and GMP. The importance of the regulatory scheme of purines is illustrated by two examples. First imagine both AMP and GMP are abundant. When this occurs, PRPP amidotransferase will be completely inhibited and no purine synthesis will occur.

Partial activity

High levels of GMP and low levels of AMP would result in PRPP amidotransferase being slightly active, due to the fact GMP will fill one allosteric site, but low AMP levels will mean second allosteric site will likely be unfilled. This lowered (but not completely inhibited) activity of PRPP amidotransferase will allow for limited production of 5-PRA and the rest of the pathway intermediates, so it will remain active.

At the IMP branch, however, the high levels of GMP will inhibit IMP dehydrogenase, thus shutting off that branch and allowing all of the intermediates to be funneled into making AMP. When the AMP level rises high enough, AMP binds to PRPP amidotransferase and along with GMP, shuts off the enzyme. A reversal will occur if AMP levels are high, but GMP levels are low.

Proper balance

Regulated in this way, AMP and GMP levels can be maintained in a fairly narrow concentration range. Properly balancing nucleotide levels in cells is critical. It is likely for this reason that cells have numerous controls on the amount of each nucleotide made.

Other mechanisms

Cells have two other ways of balancing GMP and AMP nucleotides. First, the enzyme GMP reductase will convert GMP back to IMP using electrons from NADPH.

The IMP, in turn, can then be made into AMP if its concentration is low. Second, AMP can be converted back to IMP by the enzyme AMP deaminase. In this case, the IMP can then be made into GMP.

It is important to maintain appropriate proportions of the different nucleotides. Excess or scarcity of any nucleotide of any nucleotide can result in an increased tendency to mutation.

To convert AMP to ATP and GMP to GTP requires action of kinase enzymes. Each monophosphate nucleotide form has its own specific nucleoside monophosphate kinase. For adenine-containing nucleotides (ribose forms and deoxyribose forms), adenylate kinase catalyzes the relevant reaction.

The adenylate kinase reaction is reversible and is used to generate ATP when the cell's ATP concentration is low. When ATP is made from 2 ADPs in this way, AMP levels increase and this is one way the cell senses that it is low on energy.

Guanosine monophosphates also have their own kinase and it catalyzes the reaction at the top of the next page.

Other monophosphate kinases for UMP and CMP use ATP in a similar fashion.

In going from the diphosphate form to the triphosphate form, the picture is simple - one enzyme catalyzes the reaction for all diphosphates (ribose and deoxyribose forms). It is known as nucleoside diphosphate kinase or (more commonly) NDK or NDPK and it catalyzes reactions of the form

where X and Y refer to any base.

Purine salvage reactions

Not all nucleotides in a cell are made from scratch. The alternatives to de novo syntheses are salvage pathways. Salvage reactions to make purine nucleotides start with attachment of ribose to purine bases using phosphoribosylpyrophosphate (PRPP).

The enzyme catalyzing this reaction is known as hypoxanthine/guanine phosphoribosyltransferase (HGPRT - Figure 6.175) and is interesting from an enzymological as well as a medical perspective. First, the enzyme is able to catalyze both of the next two important salvage reactions - converting hypoxanthine to IMP or guanine to GMP.

HGPRT is able to bind a variety of substrates at its active site and even appears to bind non-natural substrates, such as acyclovir preferentially over its natural ones.

Medical perspective

From a medical perspective, reduction in levels of HGPRT leads to hyperuricemia, a condition where uric acid concentration increases in the body. Complete lack of HGPRT is linked to Lesch-Nyhan syndrome, a rare, inherited disease in high uric acid concentration throughout the body is associated with severe accompanying neurological disorders.

Reduced production of HGPRT occurs frequently in males and has a smaller consequence (gout) than complete absence. Interestingly, gout has been linked to a decreased likelihood of contracting multiple sclerosis, suggesting uric acid may help prevent or ameliorate the disease.

Expression of HGPRT is stimulated by HIF-1, a transcription factor made in tissues when oxygen is limiting, suggesting a role for HGPRT under these conditions.

Adenine salvage

The enzyme known as adenine phosphoribosyltransferase (APRT) catalyzes the reaction corresponding to HGPRT for salvaging adenine bases.

Pyrimidine nucleotide metabolism

The de novo pathway for synthesizing pyrimidine nucleotides has about the same number of reactions as the purine pathway, but also has a different strategy. Whereas the purines were synthesized attached to the ribose sugar, pyrimidine bases are made apart from the ribose and then attached later.

The first reaction is catalyzed by carbamoyl phosphate synthetase (Figure 6.176).

Two different forms are found in eukaryotic cells. Form I is found in mitochondria and form II is in the cytoplasm.

The reaction catalyzed by carbamoyl phosphate synthetase is the rate limiting step in pyrimidine biosynthesis and corresponds to reaction 1 in Figure 6.178.

Balance

The enzyme is activated by ATP and PRPP and is inhibited by UMP. This helps to balance pyrimidine vs. purine concentrations. High concentrations of a purine (ATP) activates the synthesis of pyrimidines. PRPP increases in concentration as purine concentration increases, so it too helps to establish that balance. UMP is an end product of pyrimidine metabolism, so the process is self-limiting. The next enzyme in the pathway, aspartate transcarbamoylase (ATCase) also plays a role in the same balance, as we will see. The reaction it catalyzes is shown below and is reaction 2 in Figure 6.178.

ATCase is a classic enzyme exhibiting allosteric regulation and feedback inhibition, having both homotropic and heterotropic effectors (Figure 6.179 and see HERE). With 12 subunits (6 regulatory and 6 catalytic units), the enzyme exists in two states - a low activity T-state and a high activity R-state. Binding of the aspartate substrate to the active site shifts the equilibrium in favor of the R-state.

Aspartate is a homotropic effector of the enzyme, because it acts allosterically on the enzyme and is a substrate for it as well. Similarly, binding of ATP to the regulatory units favors the R-state, whereas binding of CTP to the regulatory units favors the T-state. ATP and CTP are heterotropic effectors of the enzyme because they are not substrates for it, but act allosterically.

Regulation

As was seen with the first enzyme of the pathway, high concentration of purine nucleotides stimulates synthesis of pyrimidines and high concentration of pyrimidines turns off the pathway that synthesizes them.

Dihydroorotase catalyzes reaction 3 and is found in the cytoplasm, as is ATCase.

Reaction 4 occurs in the mitochondrion, so the product of reaction 3, dihydroorotate, must be transported into the mitochondrion from the cytoplasm. In reaction 4, dihydroorotate is oxidized to orotate. The enzyme catalyzing the reaction is dihydroorotate dehydrogenase.

Reaction #5, catalyzed by orotate phosphoribosyl transferase, involves connection of orotate to ribose to yield a nucleotide - orotidine-5'-monophosphate (OMP).

Last, OMP is converted to uridine-5'-monophosphate (UMP) by action of a fascinating enzyme known as OMP decarboxylase.

OMP decarboxylase is frequently cited as an example for the incredible ability of an enzyme to speed a reaction. The decarboxylation of OMP, if allowed to proceed in the absence of an enzyme takes about 78 million years. In the presence of OMP decarboxylase, the reaction takes place in 18 milliseconds, a speed increase of about 1017. Remarkably, the enzyme accomplishes this without any cofactors or coenzymes of any kind.

The mechanism of action of the enzyme is shown in Figure 6.180. In mammals, the activities of OMP decarboxylase and orotate phosphoribosyl transferase are contained on the same protein.

A monophosphate kinase (UMP/CMP kinase) catalyzes conversion of UMP to UDP.

The same enzyme will also phosphorylate CMP to CDP and dCMP to dCDP. Like the reaction of adenylate kinase, the reaction above, when run in the reverse direction, can be a source of ATP when the cell is low on energy.

The next step, catalyzed by NDPK, uses energy of any triphosphate nucleotide (XTP) to produce UTP from UDP.

CTP Synthase

UTP is the substrate for synthesis of CTP via catalysis by CTP synthase.

This enzyme is inhibited by its product, ensuring too much CTP is not made and activated by physiological concentrations of ATP, GTP, and glutamine. One human isozyme, CTPS 1, has been shown to be inactivated by phosphorylation by glycogen synthase kinase 3.

CTP synthase has two domains and is a heterodimer (Figure 6.183). It exists as an inactive monomer at low enzyme concentrations or in the absence of UTP and ATP. One domain of the enzyme cleaves the amine group from glutamine and transfers it internally to the UTP. The other domain (synthase domain) binds ATP and initiates the mechanism shown in Figure 6.184 for making CTP.

CTP is the only nucleotide synthesized de novo directly as a triphosphate, since it arises directly from UTP. Since deoxyribonucleotides are made from ribonucleoside diphosphates, it means deoxycytidine nucleotides must either be made preferentially from salvage nucleotides or CTP must be dephosphorylated first.

One enzyme that can do this is a membrane-bound enzyme known as apyrase, which sequentially converts CTP to CDP and then CMP.

Pyrimidine salvage reactions

Pyrimidine salvage synthesis allows cells to remake pyrimidine triphosphate nucleotides starting from either the C or U pyrimidine bases, nucleosides, or nucleotides. Figures 1.85 & 6.186 depict salvage pathway reactions. As is apparent in Figure 1.86, there are multiple ways of making the same molecules. For example, uracil can be made into uridine by reaction 11 or by reaction 12.

The figure depicts not only the synthesis of CTP and UTP from basic components, but also shows how these nucleotides can be broken down into smaller pieces.

In many cases, the same enzyme works on cytidine, uridine, and deoxycytidine molecules.

Enzymes of note

There are several enzymes of note in the salvage pathway. Seven enzymes, for example, work on both uracil and cytosine containing nucleosides/nucleotides. These include NTP phosphatase (reaction 2), NDPK (reaction 3), apyrase (reaction 4), NDP phosphatase (reaction 5), UMP/CMP kinase (reaction 6), pyrimidine-specific 5' nucleotidase (reaction 7), and uridine/cytidine

kinase (reaction 8). The enzymes for reactions 6 and 8 can also use deoxyribonucleosides/deoxyribonucleotides as substrates.

Cytidine deaminase (reaction #9) converts cytidine to uridine by removing an amine group from the cytosine base and thus is a counter for the UTP to CTP reaction catalyzed by CTP synthetase. Countered reactions allow cells to balance concentrations of nucleosides/nucleotides in either direction if they should get out of balance.

Two other reactions in the figure are worth mentioning. Both UTP and CTP are converted in the breakdown process to UMP and CMP, respectively. Both of these reactions are important for deoxyribonucleotide metabolism. In each case, the monophosphate derivatives are phosphorylated, creating diphosphate derivatives (UDP and CDP) that are substrates for RNR that yield dUDP and dCDP, respectively. dUDP is phosphorylated to dUTP and then pyrophosphate is removed by dUTPase to yield dUMP. dUMP is a substrate for thymidine synthesis (see HERE). dCDP is converted to dCTP by NDPK

Deoxyribonucleotide metabolism

Deoxyribonucleotides, the building blocks of DNA, are made almost exclusively from ribonucleoside diphosphates. A single enzyme called ribonucleotide reductase (RNR) is responsible for the conversion of each of these to a deoxy form (Figure 6.187). The enzyme's substrates are ribonucleoside diphosphates (ADP, GDP, CDP, or UDP) and the products are deoxyribonucleoside diphosphates (dADP, dGDP, dCDP, or dUDP). Thymidine nucleotides are synthesized from dUDP.

RNR has two pairs of two identical subunits - R1 (large subunit) and R2 (small subunit). R1 has two allosteric binding sites and a catalytic site. R2 forms a tyrosine radical necessary for the reaction mechanism of the enzyme.

There are three classes of RNR enzymes and they differ in the nature or means of generating a radical used in the enzyme's catalytic mechanism. Class I RNRs are found in eukaryotes, eubacteria, bacteriophages, and viruses. They all use a ferrous iron center that loses an electron (converting to ferric iron) to generate a free radical on a tyrosine ring. These enzymes only work in aerobic conditions.

Class II RNRs use 5'-deoxyadenosyl cobalamin (vitamin B12) to generate a radical and work under aerobic or anaerobic conditions. They are found in eubacteria, archaebacteria, and bacteriophages. Class III RNRs generate a glycine radical using S-adenosyl methionine (SAM) and an iron-sulfur center. They work under anaerobic conditions and are used by archaebacteria, eubacteria, and bacteriophages. Substrates for class I enzymes are ribonucleoside diphosphates. Class II enzymes work on ribonucleoside diphosphates or ribonucleoside triphosphates. Class III enzymes work on ribonucleoside triphosphates.

In class I enzymes, RNR is an iron-dependent dimeric enzyme with each monomeric unit containing a large subunit (known as α or R1) and a small subunit (known as β or R2). The R1 subunit contains regulatory binding sites for allosteric effectors (see below), whereas the R2 subunit houses a tyrosine residue that forms a radical critical to the reaction mechanism of the enzyme. Electrons needed in the reaction are transmitted from NADPH to the enzyme by one of two pathways, reducing a disulfide bond in the enzyme to two sulfhydryls. In the first transfer mechanism, NADPH passes electrons to glutathione, which passes them to glutaredoxin, which then donates them to the RNR enzyme used in the reaction. In the second mechanism, NADPH passes electrons to FAD, which uses them to reduce thioredoxin, which then passes the electrons to RNR with the same end result as in the first pathway - reduction of a suflhydryl in RNR.

In the reaction mechanism (Figure 6.188), a tyrosine side chain in the R2 unit must be radicalized to start. This electronic change is transmitted through the small R2 subunit to the active site of the large R1 subunit. Several aromatic amino acid side chains are thought to play a role in that process. Iron atoms in the R2 subunit assist in creation and stabilization of the radical. The tyrosine radical contains an unpaired electron delocalized across its aromatic ring.

Transfer of the electronic instability to the R1 unit results in radicalization of a cysteine (to form a thiyl radical) at the active site. The thiyl radical, thus formed, abstracts a hydrogen atom (proton plus electron) from carbon 3 of ribose on the bound ribonucleoside diphosphate, creating a radical carbon atom. Radicalization of carbon #3 favors release of the hydroxyl group on carbon #2 as water. The extra proton comes from the sulfhydryl of the enzyme's cysteine. In the next step of the process, a proton and two electrons from the same cysteine are transferred to carbon #2 and then carbon #3 takes back the proton originally removed from it to yield a deoxyribonucleoside diphosphate. The enzyme's thiyl group gains an electron from R2 and the disulfide bond created in the reaction must be reduced by electrons from NADPH again in order to catalyze again.

Regulation

In addition to RNR's unusual reaction mechanism, the enzyme also has a complex system of regulation, with two sets of allosteric binding sites, both found in the R1 subunit. Because a single enzyme, RNR, is responsible for the synthesis of all four

deoxyribonucleotides, it is necessary to have mechanisms to ensure that the enzyme produces the correct amount of each dNDP. This is a critical consideration, since imbalances in DNA precursors can lead to mutation.

Consequently, the enzyme must be responsive to the levels of the each deoxyribonucleotide, selectively making more of those that are in short supply, and preventing additional synthesis of those that are abundant. These demands are met by having two separate control mechanisms on the enzyme - one that determines which substrate will be acted on, and another that controls the enzyme's activity.

Two allosteric sites

RNR is allosterically regulated via two molecular binding sites - a specificity binding site (binds dNTPs and induces structural changes in the enzyme that determines which substrates preferentially bind at the catalytic site and an activity control site (controls whether or not enzyme is active). The activity control site functions like a simple on/off switch - ATP activates catalysis, dATP inactivates it. (One subset of class I enzymes, however, is not affected by dATP.)

The inactivation of RNR by dATP is an important factor in the disease known as Severe Combined Immunodeficiency Disease (SCID). In SCID, the salvage enzyme adenosine deaminase is deficient, leading to a rise in concentration of dATP in cells of the immune system. dATP shuts down RNR in these cells, thus stopping their proliferation and leaving the affected individual with a very weak or no immune system.

Allosteric effectors

When dTTP is abundant (Figure 6.189), it binds to RNR's specificity site and inhibits binding and reduction of CDP and UDP but stimulates binding and reduction of GDP at the active site of the enzyme. Conversely, binding of ATP or dATP at the specificity site stimulates binding and reduction of CDP and UDP at the active site. Last, binding of dGTP to the specificity site (specificity site B) induces binding and reduction of ADP at the active site.

Students sometimes confuse the active site of RNR with the activity control site (sometimes called the activity site). The active site is where the reaction is catalyzed, and could better be called the catalytic site, whereas the activity site is an allosteric binding site for ATP or dATP that controls whether the enzyme is active. High levels of dATP are an indicator that sufficient dNTPs are available, so the enzyme gets inhibited to stop production of more. Low levels of dATP allow binding of ATP and activation of the enzyme.

In addition to regulation by deoxyribonucleotides and ATP, RNR can be directly inhibited by hydroxyurea.

dTTP synthesis

Synthesis of dTTP by the de novo pathway involves a multi-step process from UDP to dTTP. It begins with UDP, which is converted to dUDP by RNR. dUDP is phosphorylated by NDPK to yield dUTP, which is quickly broken down by dUTPase to produce dUMP. The remaining reactions are shown in Figure 6.190.

Important enzymes in the pathway include dUTPase and thymidylate synthetase. dUTPase is important for keeping the concentration of dUTP low so it does not end up in DNA. DNA polymerase can use dUTP just as it does dTTP, and incorporate it into a DNA strand, across from adenine nucleotides.

Thymidylate synthetase is important because it is a target (directly and indirectly) for anticancer therapies. As shown in Figure 6.191, a methyl group from N5,N10-methylene-tetrahydrofolate (often called tetrahydrofolate) is donated to dUMP, making dTMP and dihydrofolate (DHF). Folate molecules are in limited quantities in cells and must be recycled, because if they are not, then the reaction to make dTMP cannot occur. Recycling of dihydrofolate to tetrahydrofolate occurs by the reaction shown in Figure 6.192.

The enzyme involved in the conversion of dihydrofolate to tetrahydrofolate, dihydrofolate reductase (DHFR - Figure 6.192), is one target of anticancer drugs because by stopping the regeneration of tetrahydrofolate from dihydrofolate (otherwise a dead end), one can stop production of thymidine nucleotides and, as a result, halt DNA synthesis, thus preventing a cancer cell from dividing. Competitive inhibitors of DHFR include methotrexate (Figure 6.194) or aminopterin. Cells contain numerous folates for performing one carbon metabolism and the pathways by which they are all recycled is shown in Figure 6.193.

5-fluorouracil

Yet another important inhibitor of thymidine synthesis is used to treat cancer. This compound, 5-fluorouracil (Figure 6.195 and Movie 6.3) is a suicide inhibitor of thymidylate synthase.

Salvage synthesis

Besides synthesis from simple precursors, nucleotides can also be made from pieces of existing ones. This is particularly relevant, since consumption of food introduces to the body a large collection of proteins, lipids, and nucleic acids that are all more efficiently recycled than degraded. For proteins, the process is simple. Digestion converts them into constituent building blocks (amino acids) and these are re-assembled into proteins of the consuming organism using the genetic code.

Nucleotides

The multi-component structure of nucleotides, though (base, sugar, phosphate) means subsections of them may be re-utilized. Phosphate is recycled simply by entering the phosphate pool of the cell. It is typically built back into triphosphate forms (ultimately) by oxidative phosphorylation and kinase actions. Salvage of bases is different for purines and pyrimidines and is discussed separately HERE and HERE.

Nucleotide catabolism

Besides salvage and being built into nucleic acids, nucleotides can also be broken down into simpler component molecules. Some of these molecules, such as uric acid, can have significant impact on organisms (see HERE).

Purine catabolism

Breakdown of purine nucleotides starts with nucleoside monophosphates, which can be produced by breakdown of an RNA, for example, by a nuclease (Figure 6.196).

Metabolism of AMP and GMP converge at xanthine. First, AMP is dephosphorylated by nucleotidase to create adenosine, which is then deaminated by adenosine deaminase to yield inosine. Alternatively, AMP can be deaminated by AMP deaminase to yield IMP.

IMP is also an intermediate in the synthesis pathway for purine anabolism. Dephosphorylation of IMP (also by nucleotidase) yields inosine. Inosine has ribose stripped from it by action of purine nucleotide phosphorylase to release hypoxanthine. Hypoxanthine is oxidized to xanthine in a hydrogen peroxide-generating reaction catalyzed by xanthine oxidase.

Catabolism of GMP proceeds independently, though similarly. First, phosphate is removed by nucleotidase to yield guanosine. Guanosine is stripped of ribose to yield free guanine base, which is deaminated by guanine deaminase (also called guanase) to produce xanthine.

Xanthine oxidase enters the picture a second time in the next reaction catalyzing a second reaction by a similar mechanism to the hypoxanthine oxidation described previously. It is shown on the next page.

Uric acid

Uric acid is problematic in some higher organisms (including humans) because it is not very soluble in water. Consequently it precipitates out of solution, forming crystals (Figure 6.198). Those crystals can accumulate in joints and (frequently) in the big toe. Such a condition is known as gout.

Interestingly, there may be a negative correlation between gout and contracting multiple sclerosis. This protective effect may be due to the antioxidant protection afforded by uric acid. Uric acid is the primary excretion form of nitrogen for birds. Dalmation dogs also excrete uric acid instead of urea and may suffer from joint pain as a result of gout-like conditions.

Gout is treated with a hypoxanthine analog known as allopurinol (Figure 6.199). It inhibits action of xanthine oxidase, which favors increase in the concentration of hypoxanthine. The latter is used in salvage synthesis to make additional purines.

Uric acid can be excreted into the urine (in humans) or broken down into allantoin by the uricase enzyme. Since humans lack the enzyme to make allantoin (urea in humans is produced by the urea cycle), its presence in the body means it was produced by non-enzymatic means. This is taken to be an indicator of oxidative stress, since it allantoin is produced non-enzymatically by oxidation of uric acid.

Pyrimidine catabolism

Catabolism of uridine and thymidine nucleotides is shown above (Figure 6.200). Catabolism of cytidine nucleotides proceeds

through uridine by deamination of cytosine. The free bases, thymine and uracil, are released by the enzyme ribosylpyrimidine nucleosidase In the reductive pathway, uracil and thymine reduction by NADPH gives dihydrothymine and dihydrouracil respectively. Addition of water to these creates 3-ureidoisobutyrate and 3-ureidopropionate respectively. Hydrolysis of both these intermediates yields ammonium ion and carbon dioxide (which are made into urea) plus 3-aminoisobutyrate for the thymine pathway and β-alanine for the product of the uracil pathway. 3-aminoisobutyrate is produced during exercise and activates expression of thermogenic genes in white fat cells.

β-alanine is a rate-limiting precursor of carnosine, a dipeptide of histidine and β-alanine (Figure 6.201). Carnosine functions as an antioxidant that scavenges reactive oxygen species. It also acts as an anti-glycating agent to prevent against attachment of sugar molecules to proteins. These are factors in degenerative diseases and may play a role in aging.

Sugars

Last, but not least, the sugars ribose and deoxyribose can be recycled (ribose) or catabolized (ribose and deoxyribose). In the case of ribose, it can be reattached to bases by phosphorylase enzymes, such as uridine phosphorylase, or converted into PRPP for the same purpose, to create nucleosides. Ribose-5-phosphate is an intermediate in the pentose phosphate pathway, allowing it to be converted into other sugars or broken down in glycolysis.

Deoxyribose-5-phosphate can be broken into two pieces by deoxyribose-5-phosphate aldolase. The products of this reaction are glyceraldehyde-3-phosphate and acetaldehyde. The former can be oxidized in glycolysis and the latter can be converted into acetyl-CoA for further metabolism.

This page titled 6.6: Nucleotides is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7: INFORMATION PROCESSING

The nature of biological information, how it is copied and passed on, how it is read and interpreted, and how it gives rise to the cellular activities that we can observe, is the subject of this chapter. Another kind of information is also considered, towards the end of the chapter- the molecular information that cells receive from, and send to, each other. Overlaid on the instructions in the genes, this information provides cells with ongoing clues about both their own inner state and the environment around them. The interplay of these two kinds of information is responsible for the form and behavior of all living organisms.

7.1: Prelude to Information Processing

7.2: Genes and Genomes

7.3: DNA Replication

7.4: DNA Repair

7.5: Transcription

7.6: RNA Processing

7.7: Translation

7.8: Gene Expression

7.9: Signaling

This page titled 7: Information Processing is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.1: Prelude to Information Processing

“The blueprints for the construction of one human being requires only a meter of DNA and one tiny cell. ... even Mozart started out this way.” — L.L. Larison Cudmore

As creatures used to regarding ourselves as exceptional, humans must surely be humbled to realize that the instructions, for making one of our own, reside in a molecule so simple that scientists, for a very long time, did not believe could possibly contain enough information to build even a simple cell. But a large body of evidence, built up over the past century, supports Larison Cudmore’s assertion that the information for making you and me (and all the other kinds of living things in the world) is encoded in DNA. Tying in with Mendel’s observations about how characteristics are passed on from one generation to the next, the discovery that there was a molecule that carried this information, altered for ever how people thought about heredity.

The elucidation of the structure of DNA provided greater insights into how traits might be encoded in a molecule, and the ways in which the information is used by cells. As we learn more about this topic, scientists have remarked on how the information in our DNA resembles the programs that drive computers. While this analogy is a simplification, there is definitely a sense in which, as Richard Dawkins put it, “the machine code of the genes is uncannily computer-like”, with information in our DNA directly determining the properties of the proteins that run our cells. We know, as Ada Yonath described it, that, “DNA is a code of four letters; proteins are made up of amino acids which come in 20 forms. So the ribosome is a very clever machine that reads one language and operates in another. “

If this sounds strange, it is even more intriguing to realize DNA is copied and passed on from cell to cell, from one generation to the next. There is an unbroken line of inheritance from the first cell to every organism alive today. In the words of Lewis Thomas, “All of today’s DNA, strung through all the cells of the earth, is simply an extension and elaboration of [the] first molecule.”

The nature of this information, how it is copied and passed on, how it is read and interpreted, and how it gives rise to the cellular activities that we can observe, is the subject of this chapter. Another kind of information is also considered, towards the end of the chapter- the molecular information that cells receive from, and send to, each other. Overlaid on the instructions in the genes, this information provides cells with ongoing clues about both their own inner state and the environment around them. The interplay of these two kinds of information is responsible for the form and behavior of all living organisms.

This page titled 7.1: Prelude to Information Processing is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.2: Genes and Genomes

Source: BiochemFFA_7_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy



Introduction

For many years, scientists wondered about the nature of the information that directed the activities of cells. What kind of molecules carried the information, and how was the information passed on from one generation to the next? Key experiments, done between the 1920s and the 1950s, established convincingly that this genetic information was carried by DNA. In 1953, with the elucidation of the structure of DNA, it was possible to begin investigating how this information is passed on, and how it is used.

Genomes

We use the word “genome” to describe all of the genetic material of the cell. That is, a genome is the entire sequence of nucleotides in the DNA that is in all of the chromosomes of a cell. When we use the term genome without further qualification, we are generally referring to the chromosomes in the nucleus of a eukaryotic cell. As you know, eukaryotic cells have organelles like mitochondria and chloroplasts that have their own DNA (Figure 7.1 & 7.2). These are referred to as the mitochondrial or chloroplast genomes to distinguish them from the nuclear genome.

Starting in the 1980s, scientists began to determine the complete sequence of the genomes of many organisms, in the hope of better understanding how the DNA sequence specifies cellular functions. Today, the complete genome sequences have been determined for thousands of species from all domains of life, and many more are in the process of being worked out by groups of scientists across the world.

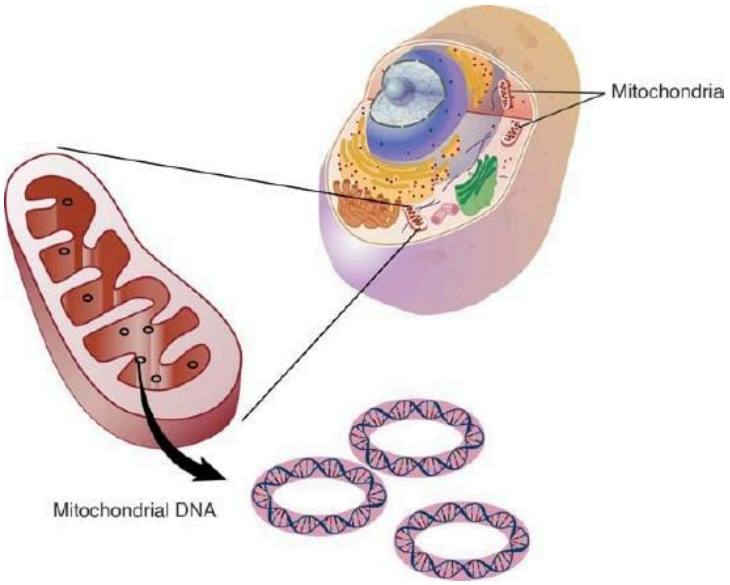


Figure 7.1 - Mitochondria carry their own genome

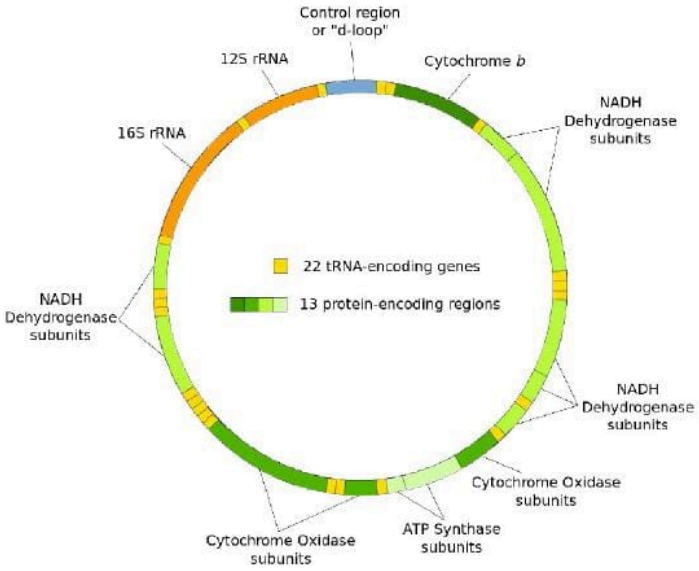


Figure 7.2 - The human mitochondrial genome - Wikipedia

Global genome initiative

The Global Genome Initiative, a collaborative effort to sequence at least one species from each of the 9,500 described invertebrate, vertebrate, and plant families is one of many such ventures. The information from these various efforts is collected in enormous online repositories, so that it is freely available to scientists. As the sequence databases compile ever more information, the fields of computational biology and bioinformatics have arisen, to analyze and organize the data in a way that helps biologists understand what the information in DNA means in the cellular context.

Genes

It has been known for many years that phenotypic traits are controlled by specific regions of the DNA that were termed “genes”. Thus, DNA was envisioned as a long string of nucleotides, in which certain regions, the genes, were separated by non-coding regions that were simply referred to as intergenic sequences (inter=between; genic=of genes). Early experiments in molecular biology suggested a simple relationship between the DNA sequence of a gene and its product, and led scientists to believe that each gene carried the information for a single protein. Changes, or mutations in the base sequence of a gene would be reflected in changes in the gene product, which in turn, would manifest itself in the phenotype or observable trait. This simple picture,

while still useful, has been modified by subsequent discoveries that demonstrated that the use of genetic information by cells is somewhat more complicated. Our definition of a gene is also evolving to take new knowledge into consideration.

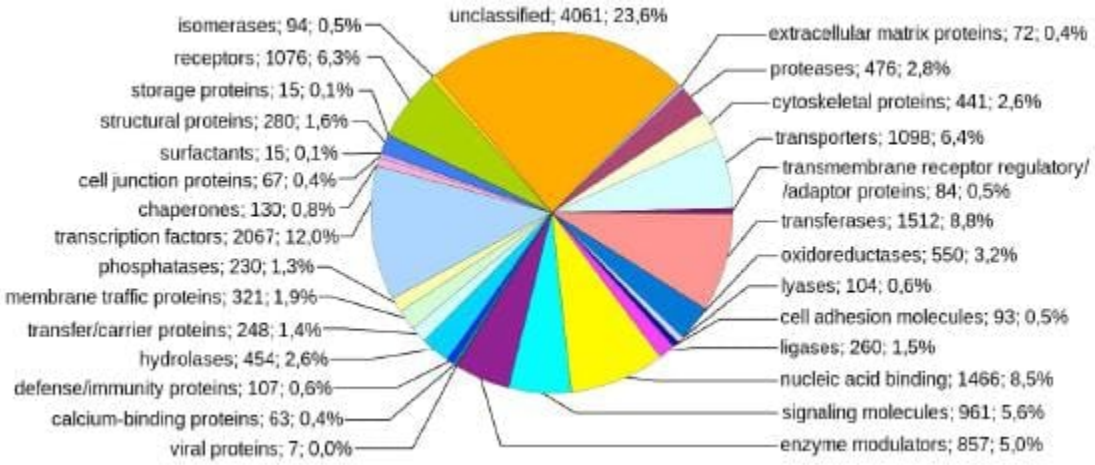


Figure 7.3 - From genomes to genes - Wikipedia

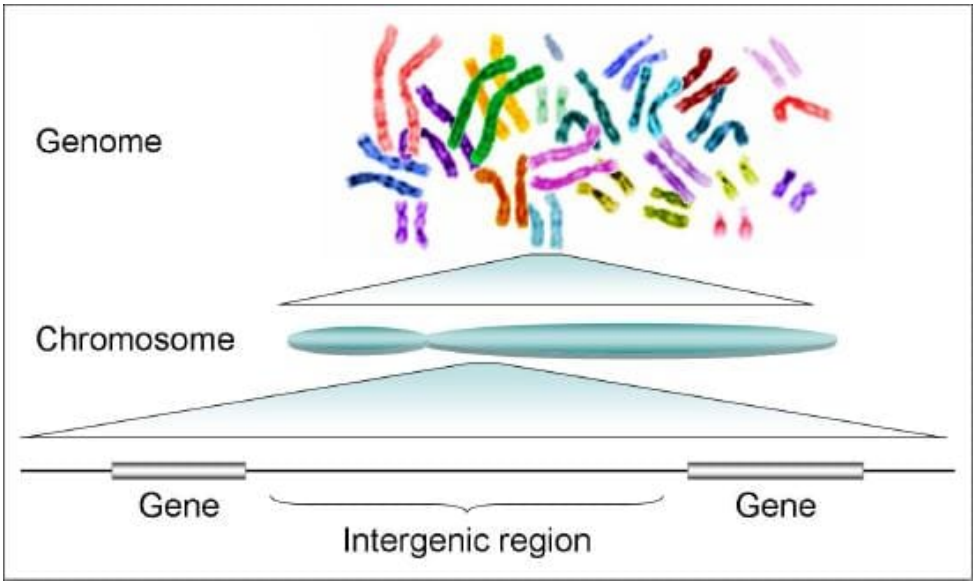


Figure 7.4 - Human genes sorted by class

Matters of size

A common-sense assumption about genomes would be that if genes specify proteins, then the more proteins an organism made, the more genes it would need to have, and thus, the larger its genome would be. Comparison of various genomes shows, surprisingly, that there is not necessarily a direct relationship between the complexity of an organism and the size of its genome (Figure 7.5). To understand how this could be true, it is necessary to recognize that while genes are made up of DNA, all DNA does not consist of genes (for purposes of our discussion, we define a gene as a section of DNA that encodes an RNA or protein product). In the human genome, less than 2% of the total DNA seems to be the sort of coding sequence that directs the synthesis of proteins. For many years, non-coding DNA in genomes was believed to be useless, and was described as “junk DNA” although it was perplexing that there seemed to be so much “useless” sequence. Recent discoveries have, however, demonstrated that much of this so-called junk DNA may play important roles in evolution, as well as in regulation of gene expression.

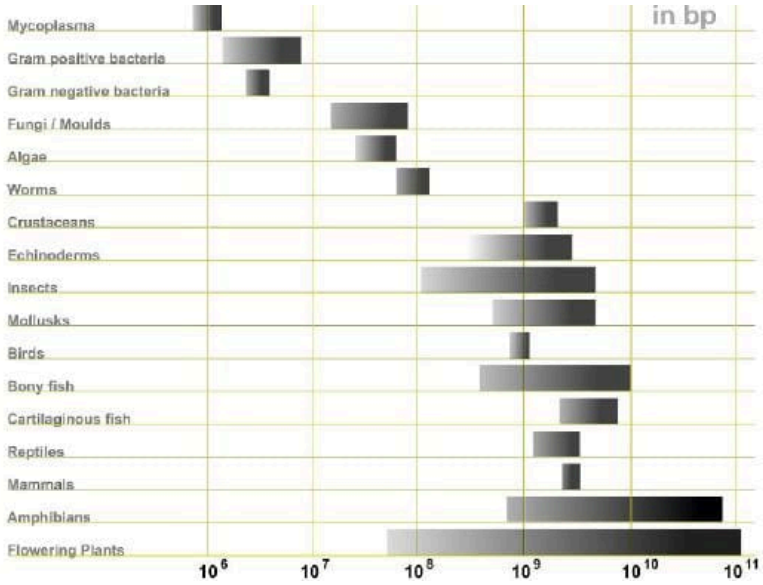


Figure 7.5 - Sizes of various genomes - Wikipedia

Introns

So, what is all the non-coding DNA doing there? We know that even coding regions in our DNA are interrupted by non-coding sequences called introns. This is true of most eukaryotic genomes. An examination of genes in eukaryotes shows that non-coding intron sequences can be much longer than the coding sections of the gene, or exons. Most exons are relatively small, and code for fewer than a hundred amino acids, while introns can vary in size from several hundred base-pairs to many kilobase-pairs (thousands of base-pairs) in length. For many genes in humans, there is much more of intron sequence than coding (a.k.a. exon) sequence. Intron sequences account for roughly a quarter of the genome in humans.

Other non-coding sequences

What other kinds of non-coding sequences are there? One function for some DNA sequences that do not encode RNA or proteins is in specifying when and to what extent a gene is used, or expressed. Such regions of DNA are called regulatory regions and each gene has one or more regulatory sequences that control its expression. However, regulatory sequences do not account for all the rest of the DNA in our genomes, either.

Transposable sequences

Surprisingly, almost half of the human genome appears to consist of several kinds of repetitive sequences. Many of the repetitive sequences are known to be transposable elements (transposons), sections of DNA that can move around within the genome. Sometimes referred to as “jumping genes” these transposable elements can move from one chromosomal location to another, either through a simple “cut and paste” mechanism that cuts the sequence out of one region of the DNA and inserts it into another location, or through a process called retrotransposition involving an RNA intermediate.

LINES & SINES

There are millions of copies of each of two major classes of such transposable elements, the LINEs (Long Interspersed Elements) and SINEs (Short Interspersed Elements) in our genomes.

LINEs and SINEs are both a kind of transposable element called retrotransposons, sequences that are copied into RNA, then reverse transcribed back into DNA before being inserted into new locations. This movement is typically not sequence specific, meaning that the transposons can be inserted randomly in the genome, in many cases within coding regions. As might be expected, this can disrupt the function of the gene. Transposons may also insert within regulatory regions, and change the expression of the genes they control. As a major cause of mutation in genomes, transposons play an important role in evolution.

Finally, recent findings have shown that much of the genome is transcribed into RNAs, even though only about 2% encodes

proteins. What are the RNAs that do not encode proteins? Ribosomal RNAs (Figure 7.7) and transfer RNAs, together with the small nuclear RNAs that function in splicing, account for some of these non-translated transcripts, but not all. The remaining RNAs are regulatory RNAs, small molecules that play an important role in regulating gene expression. As we understand more about genomes, it is becoming evident that the so-called “junk” DNA is anything but.

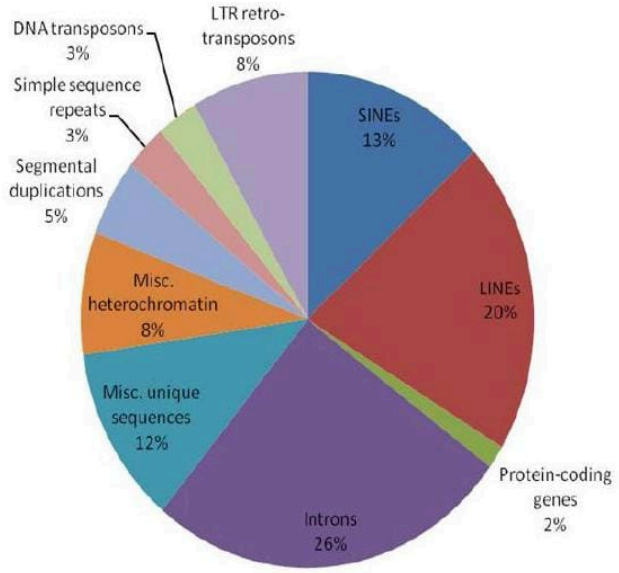


Figure 7.6 - Components of the human genome - Wikipedia

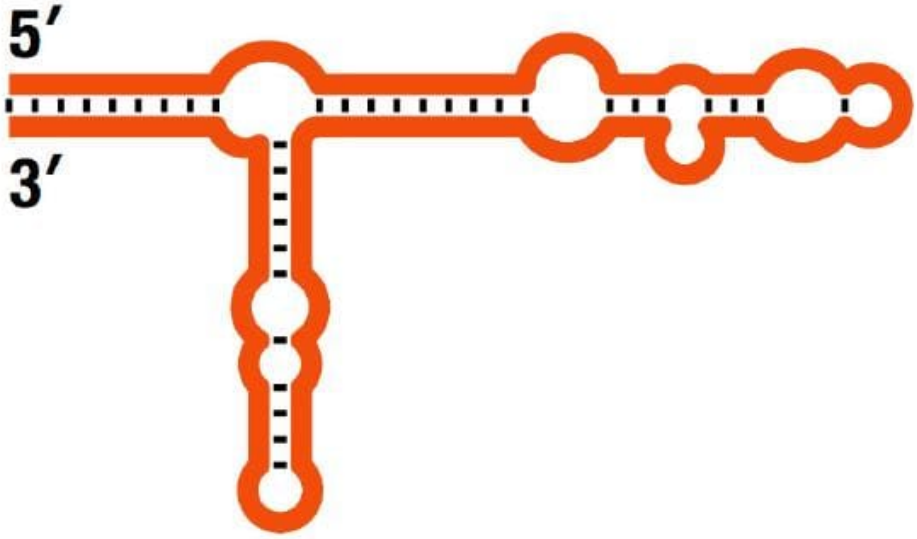


Figure 7.7 - 5S rRNA structure

This page titled 7.2: Genes and Genomes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.3: DNA Replication

Source: BiochemFFA_7_2.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy



Copying instructions

The only way to make new cells is by the division of pre-existing cells. Single-celled organisms undergo division to produce more cells like themselves, while multicellular organisms arise through division of a single cell, generally the fertilized egg. Each time a cell divides, all of its DNA must be copied faithfully so that a copy of this information can be passed on to the daughter cell. This process is called DNA replication. It is the means by which genetic information can be transmitted down generations of cells, and it ensures that every new cell has a complete copy of the genome. In the next section, we will examine the process by which the DNA of a cell is completely and accurately copied.

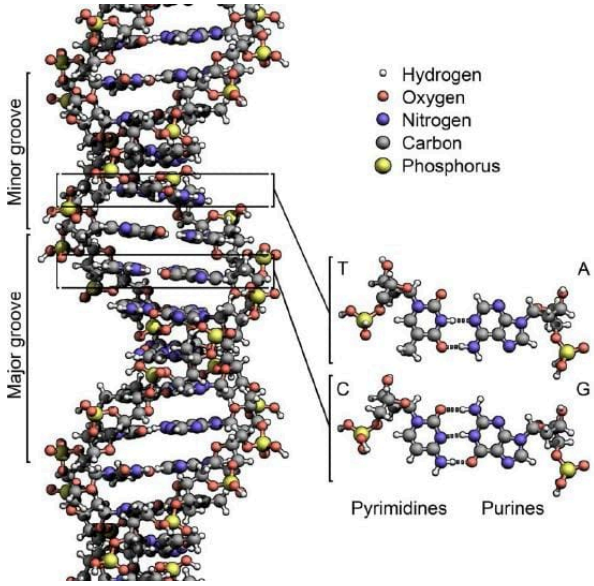


Figure 7.8 - Watson & Crick's DNA structure - Wikipedia

The structure of DNA elucidated by Watson and Crick in 1953 immediately suggested a mechanism by which double-stranded DNA could be copied to give two identical copies of the DNA. They proposed that the two strands of the DNA molecule, which are held together by hydrogen bonds between the base-paired nucleotides, would separate and each serve as a template on which a complementary strand could be assembled (Figure 7.8). The base-pairing rules would ensure that this process would result in the production of two identical DNA molecules. The beautiful simplicity of this scheme was shown to be correct in subsequent experiments by Meselson and Stahl, that demonstrated that DNA replication was semi-conservative, i.e., that after replication, each of the two resulting DNA molecules was made up of one old strand and one new strand that had been assembled across from it (Figure 7.9).

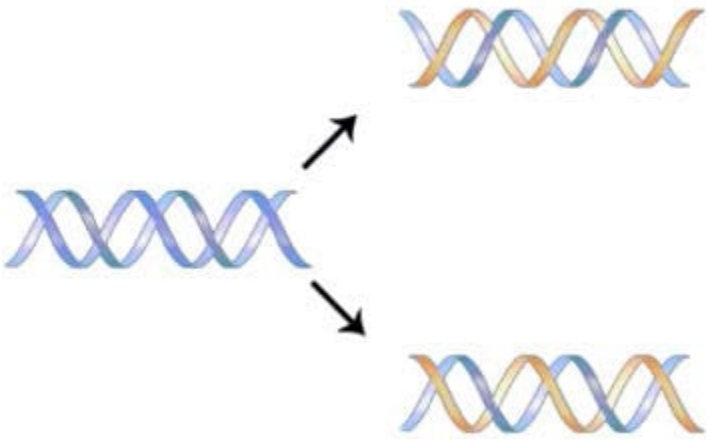


Figure 7.9 - Semi-conservative replication - Image by Martha Baker

Building materials

What are the ingredients necessary for building a new DNA molecule? As noted above, the original, or parental DNA molecule serves as the template. New DNA molecules are assembled across from each template by joining together free DNA nucleotides as directed by the base pairing rules, with As across from Ts and Gs across from Cs.

The nucleotides used in DNA synthesis are deoxyribonucleoside triphosphates or dNTPs. As can be inferred from their name, such nucleotides have a deoxyribose sugar and three phosphates, in addition to one of the four DNA bases, A, T, C or G (Figure 7.10).

When dNTPs are added into a growing DNA strand, two of those phosphates will be cleaved off, as described later, leaving the nucleotides in a DNA molecule with only one phosphate per nucleotide. This reaction is catalyzed by enzymes known as DNA polymerases, which create phosphodiester linkages between one nucleotide and the next.

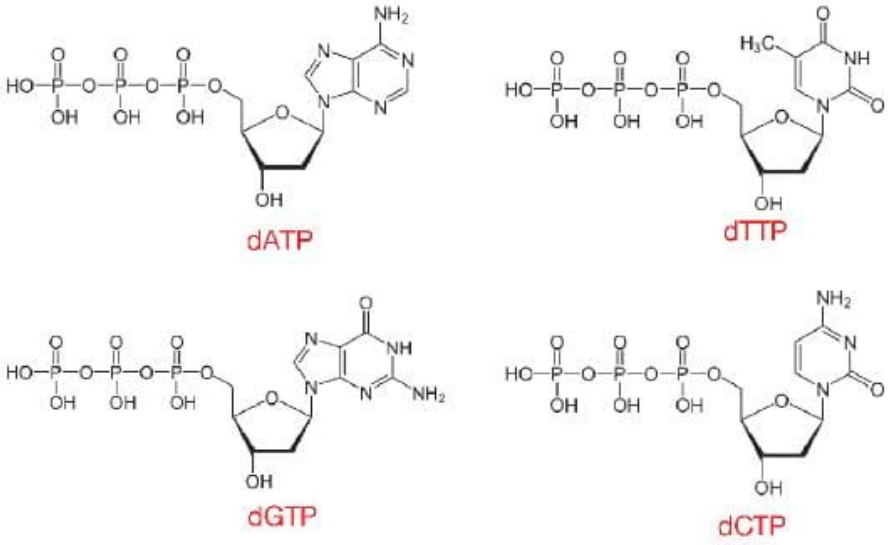


Figure 7.10 - The building blocks of DNA

Challenges

Before examining the actual process of DNA replication, it is useful to think about what it takes to accomplish this task successfully. Consider the challenges facing a cell in this process:

- The sheer number of nucleotides to be copied is enormous: e.g., in human cells, on the order of several billions.
- A double-helical parental DNA molecule must be unwound to expose single strands of DNA that can serve as templates for the synthesis of new DNA strands.
- Unwinding must be accomplished without introducing topological distortion into the molecule.
- The unwound single strands of DNA must be kept from coming back together long enough for the new strands to be synthesized.
- DNA polymerases cannot begin synthesis of a new DNA strand de novo and require a free 3' OH to which they can add deoxynucleotides.
- DNA polymerases can only extend a strand in the 5' to 3' direction. The 5' to 3' growth of both new strands means that one of the strands is made in pieces.
- The use of RNA primers requires that the RNA nucleotides must be removed and replaced with DNA nucleotides and the resulting DNA fragments must be joined.
- The copying of all the parental DNA must be accurate, so that mutations are not introduced into the newly made DNA.

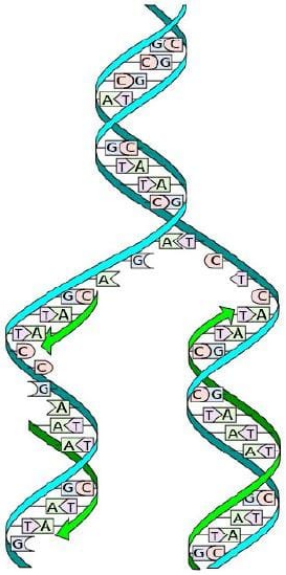


Figure 7.11 - Replication of DNA - Wikipedia

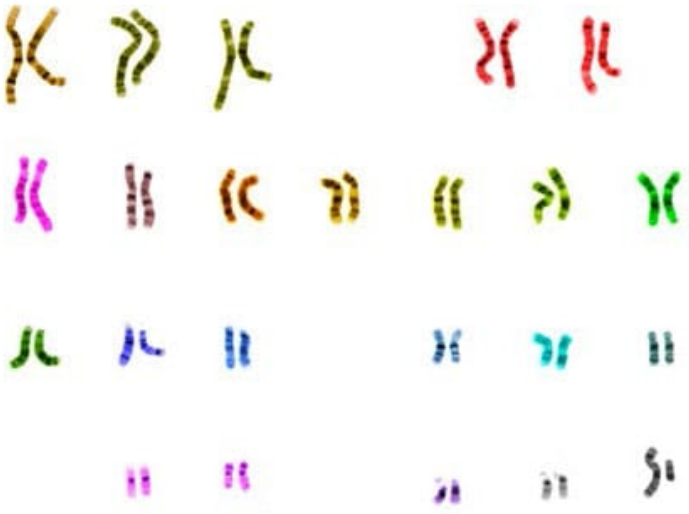


Figure 7.12 - Human chromosomes

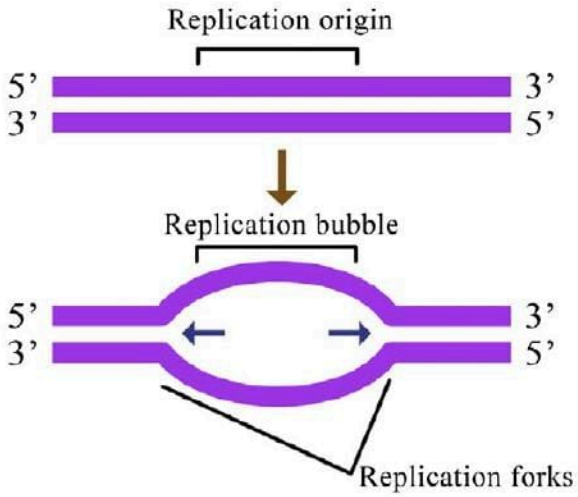


Figure 7.13 - Bidirectional replication from an origin - Martha Baker

Prokaryotic DNA Replication	Eukaryotic DNA replication
Occurs inside the cytoplasm	Occurs inside the nucleus
Only one origin of replication per molecule of DNA	Have many origins of replication in each chromosome
Origin of replication is about 100-200 or more nucleotides in length	Each origin of replication is formed of about 150 nucleotides
Replicaton occurs at one point in each chromosome	Replication occurs at several points simultaneously in each chromosome
Only have one origin of replicaton	Has multiple origins of replication
Initiation is carried out by protein DnaA and DnaB	Initiation is carried out by the Origin Recognition Complex
Topoisomerase is needed	Topoisomerase is needed
Replicaton is very rapid	Replication is slow

Figure 7.14 - Prokaryotic vs. eukaryotic DNA replication - Wikipedia

Addressing challenges

With this in mind, we can begin to examine how cells deal with each of these challenges. Our understanding of the process of DNA replication is derived from studies using bacteria, yeast, and other systems. These investigations have revealed that DNA replication is carried out by the action of a large number of proteins that act together as a complex protein machine. Numerous proteins involved in replication have been identified and characterized, including multiple different DNA polymerases in both prokaryotes and eukaryotes. Although the specific proteins involved are different in bacteria and eukaryotes, it is useful to understand the basic considerations that are relevant in all cells. A generalized account of the steps in DNA replication is presented below, focused on the challenges mentioned above.

- The sheer number of nucleotides to be copied is enormous: e.g., in human cells, on the order of several billions.

Cells, whether bacterial or eukaryotic, have to replicate all of their DNA before they can divide. In cells like our own, the vast amount of DNA is broken up into many chromosomes, each of which is composed of a linear strand of DNA (Figure 7.12). In cells like those of E. coli, there is a single circular chromosome.

In either situation, DNA replication is initiated at sites called origins of replication. These are regions of the DNA molecule that are recognized by special proteins called initiator proteins that bind the DNA. In E.coli, origins have small regions of A-T-rich sequences that are “melted” to separate the strands, when the initiator proteins bind to the origin or replication. As you may remember, A-T base-pairs, which have two hydrogen bonds between them are more readily disrupted than G-C base-pairs which have three apiece (Figure 7.15).

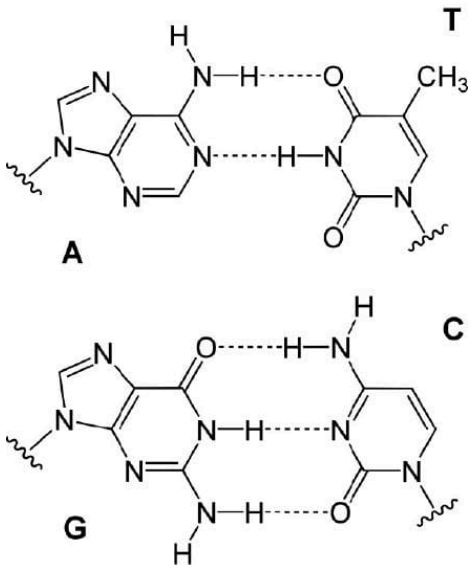


Figure 7.15 - A-T and G-C base pairs

How many origins of replication are there on a chromosome? In the case of E. coli, there is a single origin of replication on its circular chromosome. In eukaryotic cells there may be many thousands of origins of replication, with each chromosome having hundreds (Figure 7.16). DNA replication is, thus, initiated at multiple points along each chromosome in eukaryotes. Electron micrographs of replicating DNA from eukaryotic cells show many replication bubbles on a single chromosome. This makes sense in light of the large amount of DNA that there is to be copied in cells like our own, where beginning at one end of each chromosome and replicating all the way through to the other end from a single origin would simply take too long. This is despite the fact that the DNA polymerases in human cells are capable of building new DNA strands at the very respectable rate of about 50 nucleotides per second!

- A double-helical parental molecule must be unwound to expose single strands of DNA that can serve as templates for the synthesis of new DNA strands.

Unwinding

Once a small region of the DNA is opened up at each origin of replication, the DNA helix must be unwound to allow replication to proceed. The unwinding of the DNA helix requires the action of an enzyme called helicase.

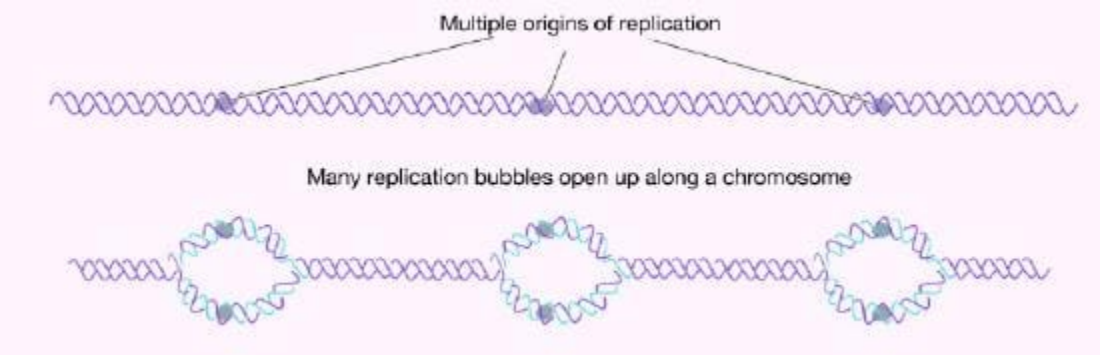


Figure 7.16 - Multiple replication bubble on a eukaryotic chromosome - Image by Martha Baker

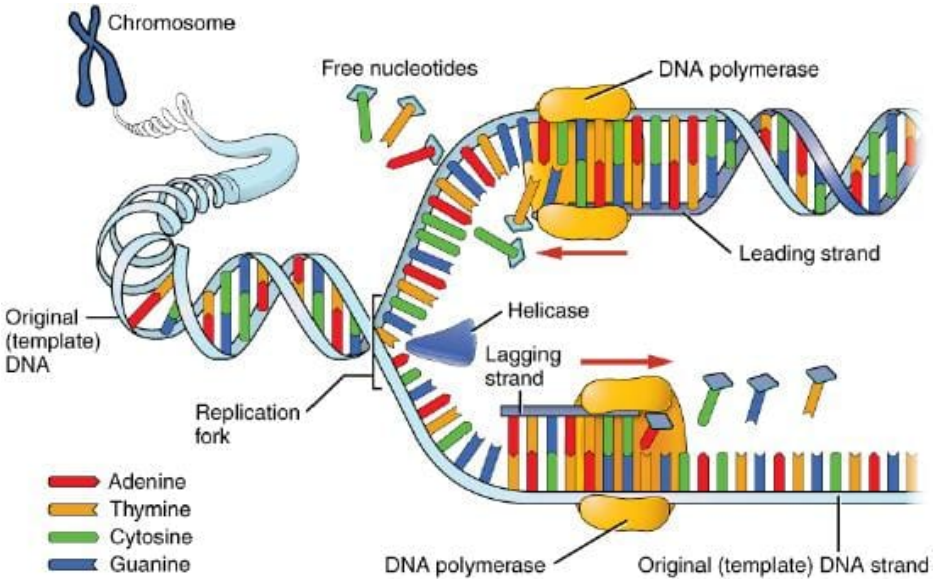


Figure 7.17 - View of a eukaryotic DNA replication fork with helicase shown in blue - Wikipedia

Helicase uses the energy released when ATP is hydrolyzed, to break the hydrogen bonds between the bases in DNA and separate the two strands (Figure 7.17). Note that a replication bubble is made up of two replication forks that "move" or open up, in opposite directions. At each replication fork, the parental DNA strands must be unwound to expose new sections of single-stranded template.

- This unwinding must be accomplished without introducing topological distortion into the molecule.

What is the effect of unwinding one region of the double helix? Local unwinding of the double helix causes over-winding (increased positive supercoiling) ahead of the unwound region.

The DNA ahead of the replication fork has to rotate, or it will get twisted on itself and halt replication. This is a major problem, not only for circular bacterial chromosomes, but also for linear eukaryotic chromosomes, which, in principle, could rotate to relieve the stress caused by the increased supercoiling.

Topoisomerases

The reason this is problematic is that it is not possible to rotate the entire length of a chromosome, with its millions of base-pairs, as the DNA at the replication fork is unwound. How, then, is this problem solved? Enzymes called topoisomerases can relieve the topological stress caused by local “unwinding” of the extra winds of the double helix. They do this by cutting one or both strands of the DNA and allowing the strands to swivel around each other to release the tension before rejoining the ends. In E. coli, the topoisomerase that performs this function is called gyrase.

- The separated single strands of DNA must be kept from coming back together so the new strands to be synthesized.

Single-strand DNA binding protein

Once the two strands of the parental DNA molecule are separated, they must be prevented from going back together to form double-stranded DNA. To ensure that unwound regions of the parental DNA remain single-stranded and available for copying, the separated strands of the parental DNA are bound by many molecules of a protein called single-strand DNA binding protein (SSB - Figure 7.18).

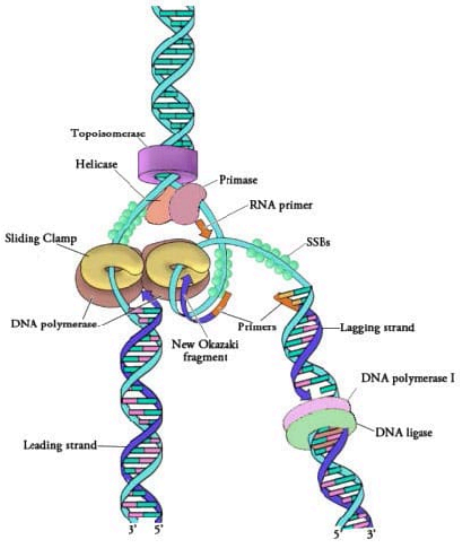


Figure 7.18 - Proteins at a prokaryotic DNA replication fork - Image by Martha Baker

- DNA polymerases cannot begin synthesis of a new DNA strand de novo and require a free 3' OH to which they can add DNA nucleotides.

Although single-stranded parental DNA is now available for copying, DNA polymerases cannot begin synthesis of a complementary strand de novo. This simply means that DNA polymerases can only add new nucleotides on to the 3' end of a pre-existing chain, and cannot start a chain of nucleotides on their own. Because of this limitation, some enzyme other than a DNA polymerase must first make a small region of nucleic acid, complementary to the parental strand, that can provide a free 3' OH to which DNA polymerase can add a deoxyribonucleotide. This task is accomplished by an enzyme called a **primase**, which assembles a short stretch of RNA base-paired to the parental DNA template. This provides a short base-paired region, called the RNA primer, with a free 3'OH group to which DNA polymerase can add the first new DNA nucleotide (Figure 7.12).

Sliding clamp

Once a primer provides a free 3'OH for extension, other proteins get into the act. These proteins are involved in loading the DNA polymerase onto the primed template and keeping it associated with the DNA. The first of these is the clamp loader. As its name suggests, the clamp loader helps to load a protein complex called the sliding clamp onto the DNA at the replication fork (Figure 7.19 and 7.20). The sliding clamp, a multi-subunit ring-shaped protein, is then joined by the DNA Polymerase. The function of the sliding clamp is to keep the polymerase associated with the replication fork - in fact, it has been described as a seat-belt for the DNA polymerase. The sliding clamp ensures that the DNA polymerase is able to synthesize long stretches of new DNA before it dissociates from the template. The property of staying associated with the template for a long time before dissociating is known as the processivity of the enzyme. In the presence of the sliding clamp, DNA polymerases are much more processive, making replication faster and more efficient.

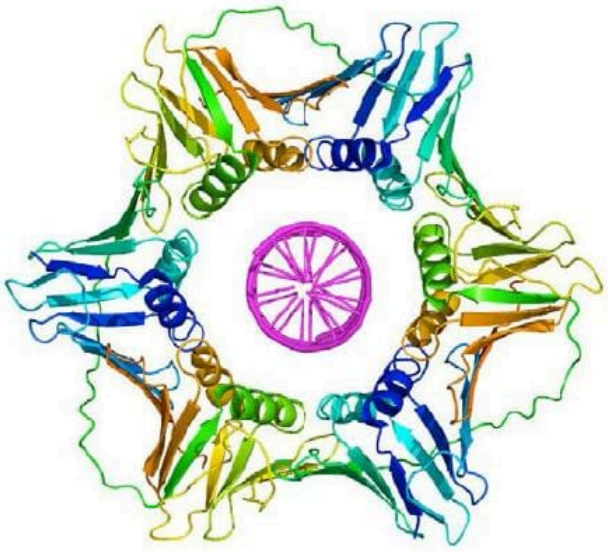


Figure 7.19 - Top view of the sliding clamp (outside) surrounding DNA strand (inside)

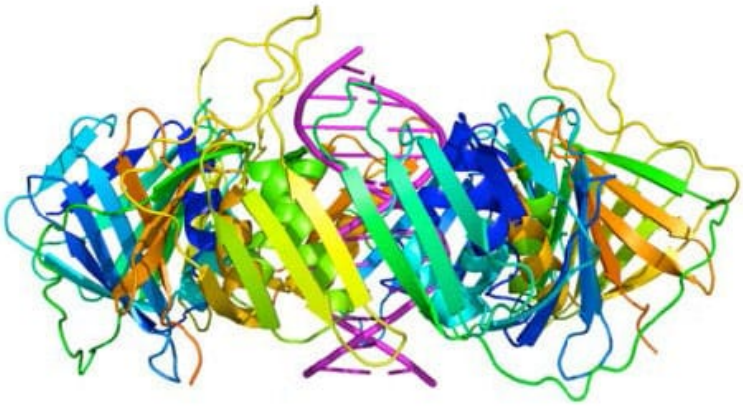


Figure 7.20 - Side view of the sliding clamp surrounding a DNA strand (in purple)

Extending the primer

The DNA polymerase is now poised to start synthesis of the new DNA strand (in E. coli, the primary replicative polymerase is called DNA polymerase III). As you already know, the synthesis of new DNA is accomplished by the addition of new nucleotides complementary to those on the parental strand. DNA polymerase catalyzes the reaction by which an incoming deoxyribonucleotide, complementary to the template, is added onto the 3' end of the previous nucleotide, starting with the 3'OH on the end of the RNA primer. The importance of the 3'OH group lies in the nature of the reaction that builds a chain of nucleotides.

The reaction catalyzed by the DNA polymerase occurs through the nucleophilic attack by the 3'OH group at the end of a nucleic acid strand on the α phosphate of the incoming dNTP (Figure 7.21). The immediate hydrolysis of the pyrophosphate that is cleaved off the incoming dNTP drives the reaction forward. The sequential addition of new nucleotides at the 3' end of the growing chain of DNA accounts for the fact that the strand grows in a 5' to 3' direction.

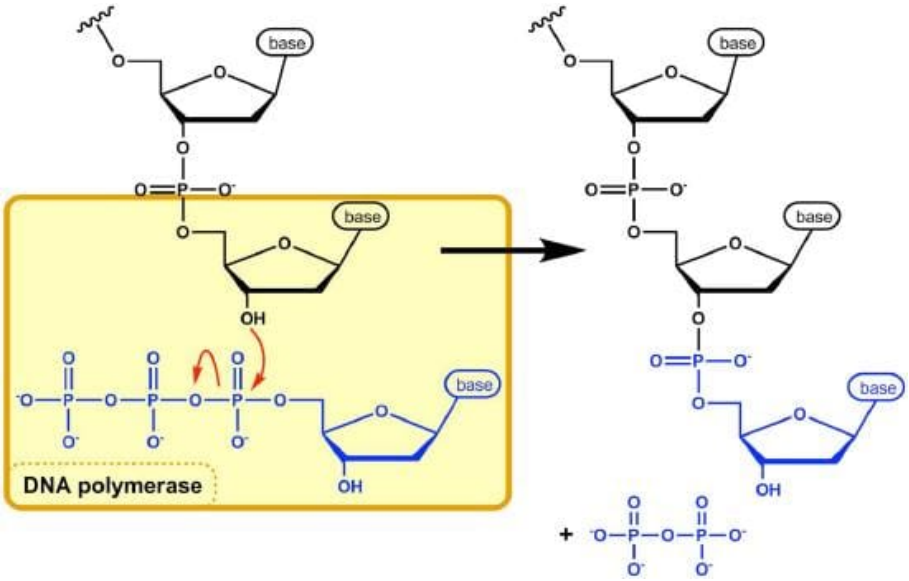


Figure 7.21 - Mechanism of DNA chain growth - Wikipedia

The 5' phosphate on each incoming nucleotide is joined by the DNA polymerase to the 3' OH on the end of the growing nucleic acid chain, to make a phosphodiester bond. Each added nucleotide provides a new 3'OH, allowing the chain to be extended for as long as the DNA polymerase continues to synthesize the new strand. As we already noted, the new DNA strands are synthesized by the addition of DNA nucleotides to the end of an RNA primer. The new DNA molecule thus has a short piece of RNA at the beginning.

- DNA polymerases can only extend a strand in the 5' to 3' direction. The 5' to 3' growth of both new strands means that one of the strands is made in pieces.

Leading strand

We know that DNA polymerases can only build a new DNA strand in the 5' to 3' direction. We also know that the two parental strands of DNA are antiparallel. This means that at each replication fork, one new strand, called the leading strand can be synthesized continuously in the 5' to 3' direction because it is being made in the same direction that the replication fork is opening up.

Lagging strand

The synthesis of the other new strand, called the lagging strand, also proceeds in the 5' to 3' direction. But because the template strands are running in opposite directions, the lagging strand is being extended in the direction opposite to the opening of the replication fork (Figure 7.22). As the replication fork opens up, the region behind the original start point for the lagging strand will need to be copied. This means another RNA primer must be laid down and extended. This process repeats itself as the replication fork opens up, with multiple RNA primers laid down and extended, producing many short pieces that are later joined. These short nucleic acid pieces, each composed of a small stretch of RNA primer and about 1000-2000 DNA nucleotides, are called Okazaki fragments, for Reiji Okazaki, the scientist who first demonstrated their existence.

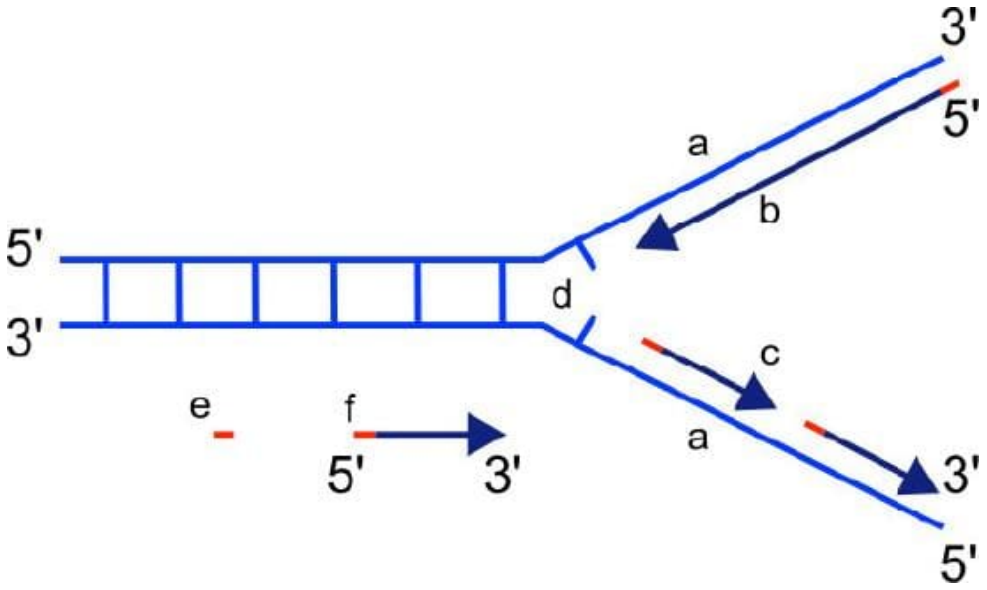


Figure 7.22 - Bidirectional DNA replication with template strand (a), leading strand (b), Okazaki fragments (c), replication fork (d), RNA primer (e), and DNA extended RNA primers (f) - Wikipedia

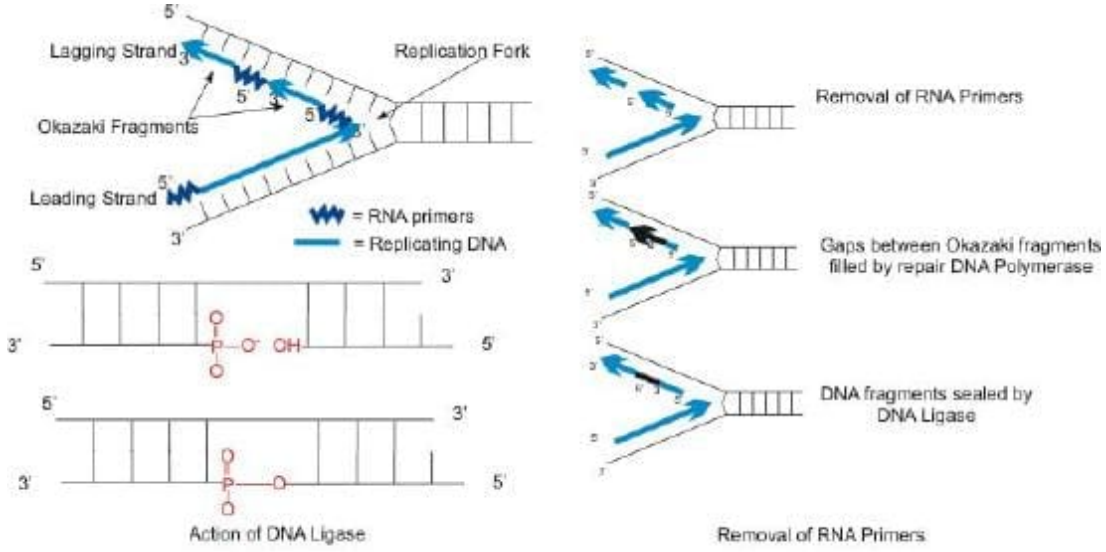


Figure 7.23 - Removal of RNA primers and joining of Okazaki fragments - Wikipedia

- The use of RNA primers requires that the RNA nucleotides must be removed and replaced with DNA nucleotides.

Primer removal

We have seen that each newly synthesized piece of DNA starts out with an RNA primer, effectively making a new nucleic acid strand that is part RNA and part DNA. The newly made DNA strand cannot be allowed to have pieces of RNA attached. So, the RNA nucleotides must be removed and the gaps filled in with DNA nucleotides (Figure 7.23). This is done by DNA polymerase I in E. coli. This enzyme begins adding DNA nucleotides at the end of each Okazaki fragment. However, the end of one Okazaki fragment is adjacent to the RNA primer at the beginning of the next Okazaki fragment. DNA polymerase I has an exonuclease activity acting in the 5' to 3' direction that removes the RNA nucleotides ahead of it, while the polymerase activity replaces the RNA nucleotides with dNTPs. Once all the RNA nucleotides have been removed, the lagging strand is made up of stretches of DNA. The DNA pieces are then joined together by the enzyme DNA ligase.

The steps outlined above essentially complete the process of DNA replication. But one issue still remains.

- Ensuring accuracy in the copying of so much information

Accuracy

How accurate is the copying of information by DNA polymerase? As you are aware, changes in DNA sequence (mutations) can change the amino acid sequence of the encoded proteins and that this is often, though not always, deleterious to the functioning of the organism. When billions of bases in DNA are copied during replication, how do cells ensure that the newly synthesized DNA is a faithful copy of the original information?

DNA polymerases, as we have noted earlier work fast (averaging 50 bases a second in human cells and up to 200 times faster in E. coli). Yet, both human and bacterial cells seem to replicate their DNA quite accurately. This is because replicative DNA polymerases have a proofreading function that enables the polymerase to detect when the wrong base has been inserted across from a template strand, back up and remove the mistakenly inserted base, before continuing with synthesis (Figure 7.24).

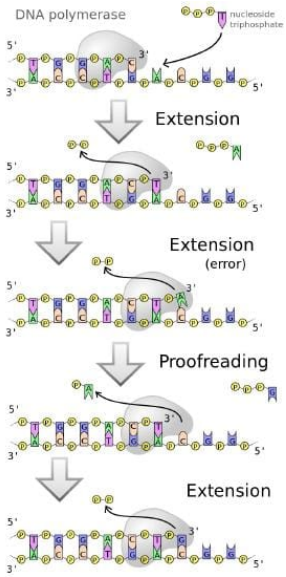


Figure 7.24 - Error corrected by DNA polymerases

Multiple activities

This is possible because most DNA polymerases are dual-function enzymes. They can extend a DNA chain by virtue of their 5' to 3' polymerase activity. Some polymerases like DNA polymerase I can also remove RNA primers in the 5’ to 3’ direction, though that is not a common activity of polymerases. Many polymerases, however have an ability to backtrack and remove the last inserted base because they possess a 3' to 5' exonuclease activity.

The exonuclease activity of a DNA polymerase allows it to excise a wrongly inserted base, after which the polymerase activity inserts the correct base and proceeds with extending the strand.

In other words, the DNA polymerase is monitoring its own accuracy (also termed its fidelity) as it makes new DNA, correcting mistakes immediately before moving on to add the next base. This mechanism, which operates during DNA replication, corrects many errors as they occur, reducing by about a 100-fold the mistakes made when DNA is copied.

DNA polymerases

As noted earlier, both prokaryotic and eukaryotic cells have multiple DNA polymerases. In E.coli, for example, DNA polymerase III is the major replicative polymerase (a.k.a. replicase) while DNA polymerase I is responsible for DNA repair as well as removal of RNA primers and their replacement with DNA nucleotides during replication. DNA polymerase II plays a role in restarting replication after DNA damage stalls replication, while DNA polymerases IV and V are both required in trans-lesion, or bypass, synthesis, which allows replication past sites of DNA damage.

Eukaryotic polymerases

In eukaryotes, there are over fifteen different DNA polymerases. The primary replicative polymerases in the nucleus are ð and ε. DNA polymerase α is also important for replication because it has primase and repair activities. Replication is initiated in eukaryotic cells by DNA polymerase α, which binds to the initiation complex at the origin and lays down an RNA primer, followed by about 25 nucleotides of DNA. It is then replaced by another polymerase, in a step called the pol switch. DNA polymerase ð or ε then continues synthesizing DNA, depending on the strand. The role of polymerase ε appears to be synthesis of the leading strand due to its high processivity and accuracy, whereas polymerase ð extends Okazaki fragments on the lagging strands. Proteins analogous to the clamp loader and sliding clamp are also present. The protein RFC plays the role of clamp loader, while another protein, PCNA acts like the sliding clamp. Several other DNA polymerases like β, γ and μ function in repairing gaps. Yet others are involved in trans-lesion synthesis following DNA damage and are associated with hypermutation.

Despite their diversity, DNA polymerases share some common structural features. X-ray crystallographic studies have shown that these enzymes have a structure that has been compared to a human right hand (Figures 7.25 & 7.26). The «palm» of the hand forms a cleft in which the DNA lies. The cleft is also the where the polymerase catalytic activity resides. This is where the incoming nucleotide is added on to the growing chain. «The fingers» position the DNA in the active site, while the «thumb» holds the DNA as it exits the polymerase. A separate domain contains the exonuclease (proofreading) activity of the enzyme.

The enzyme alternates between its polymerizing activity and its proofreading activity. When a mismatched base pair is in the polymerase catalytic site, the 3'end of the growing strand is moved from the polymerase site to the exonuclease active site (Figure 7.26). The mispair at the end is removed by the exonuclease, followed by repositioning of the 3' end in the polymerase active site to continue synthesis.

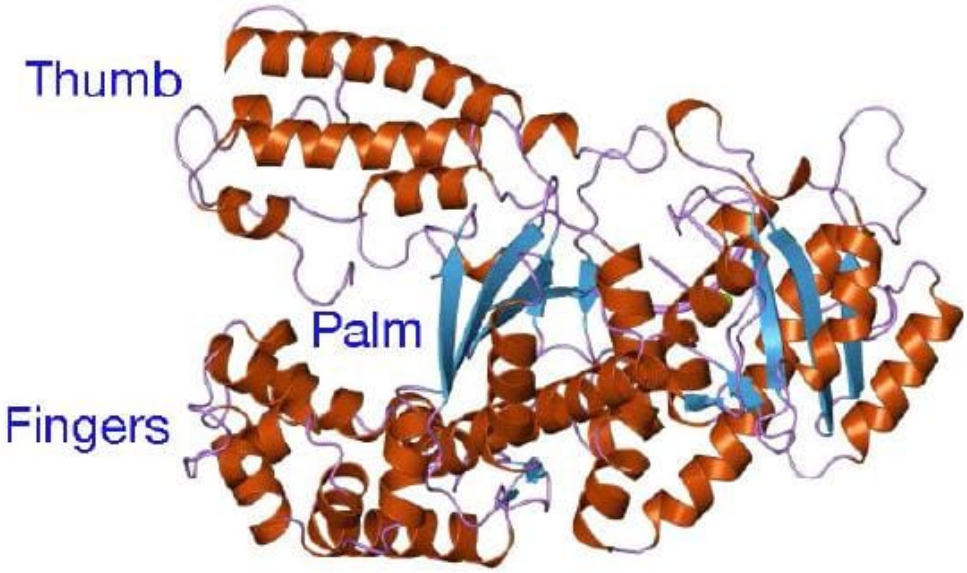


Figure 7.25 DNA polymerase I with hand structure

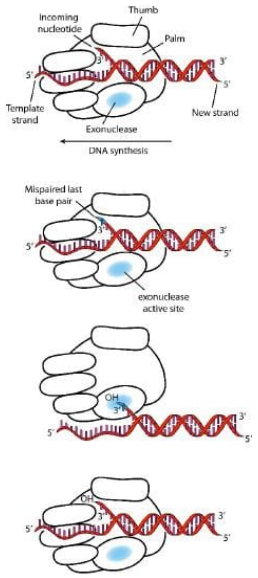


Figure 7.26 - Proofreading by DNA polymerase -Image by Pehr Jacobson

Name	Name	M.W.	Other Activity	Function
α	alpha	165,000	Primase	Repair
β	beta	39,000	Lyase	Base excision repair
δ	delta	100,000	3' exonuclease	Replication, DNA repair
ε	epsilon	225,000	3' exonuclease	Replication, DNA repair
γ	gamma	140,000	3' exonuclease	Mitochondrial replication
η	eta	78,000		Bypass of thymine dimers
I	iota	80,000	Lyase	Repair?
K	kappa	76,000		Nucleotide excision repair
λ	lambda	66,000	Lyase	Repair/recombination
μ	mu	55,000		Repair/recombination
ν	nu	100,000		Repair?
θ	theta	290,000	Helicase	Repair? Hypermutation?
ζ	zeta	353,000		Repair/hypermutation

Figure 7.27 - Mammalian DNA polymerases

Termination of replication

In circular bacterial chromosomes, there are specific sequences known as terminator or Ter sites. These are multiple short sequences that serve as termination sites, allowing the replication forks traveling clockwise and anticlockwise across the circular chromosome to meet at one of the sites.

The binding of a protein, Tus, at a Ter site prevents further movement of the replication fork and ends replication. The parental and newly made circular DNA are, at this point topologically interlinked and must be separated with the help of topoisomerase.

The end-replication problem

There is no fixed site for termination in linear eukaryotic chromosomes. As the replication forks reach the ends of the chromosome, the leading strand can be synthesized all the way to the end of the template strand. On the lagging strand, the need for an RNA primer to start synthesis creates a challenge. When the RNA primer at the extreme end of the lagging strand is removed, there is a small stretch of the template strand that cannot be copied. As a result, in each round of replication a short sequence at the ends of the chromosome will be lost. Over time, with many cycles of replication, chromosomes would become noticeably shorter. This shortening of chromosomes has been observed in vitro, in cultured mammalian somatic cells. It is also

seen in intact organisms, with increasing age.

Telomeres

What effect does the loss of sequence from the ends of the chromosomes have on cells? We know that the ends of chromosomes are characterized by structures called telomeres (Figure 7.28). Telomeres are made up of many copies of short repeated sequences (in humans, the repeat is TTAGGG) and special proteins that specifically bind to these sequences. This structure of telomeres is useful in distinguishing the ends of chromosomes from double-strand breaks in DNA, thus preventing the DNA repair mechanisms in cells from joining chromosomes end to end.

The other advantage of the repeated sequences, which do not encode proteins, is that losing some of the repeats does not lead to loss of important coding information. Thus, the repeats act as a sort of buffer zone, where the loss of sequence does not doom the cell. However, the shortening of chromosomes cannot continue indefinitely. After a certain number of replication cycles, cells are known to stop dividing and enter a state known as replicative senescence. This suggests that the shortening of the telomeres serves as a sort of clock, with the extent of shrinkage of the chromosomes serving as a measure of aging. Eventually cells that enter senescence will die.

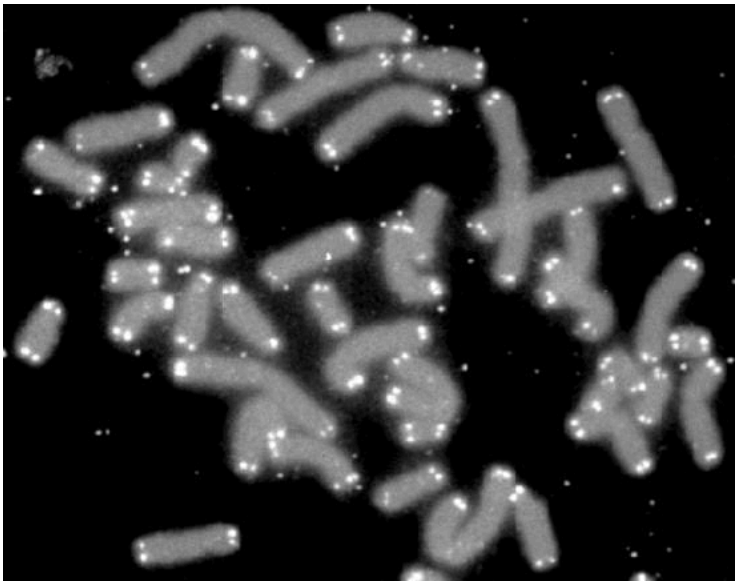


Figure 7.28 - Chromosomes with telomeres marked in white

Problems with sequence loss

Even if our cells are able to function with shorter chromosomes during our lifetimes, this leaves us with another problem. If our chromosomes grow shorter with age, then presumably our children, who inherit our chromosomes will be born with shorter chromosomes than we started with. They, in turn, would have their chromosomes shrink as they grew older, and their children would have even shorter chromosomes. Over the course of multiple generations, this would lead to the point where further chromosome shrinkage would result in cells that would enter senescence very early in life and die soon after. This obviously does not happen. Generation after generation of children are born with full-length chromosomes, so there is a mechanism that must ensure that at least in the reproductive cells, chromosomes do not get shorter.

To understand this mechanism, it is necessary to first examine the end of a newly made DNA molecule (Figure 7.29). While the leading strand, which grows in the same direction as the movement of the replication fork, can copy its template all the way to the end, the lagging strand encounters a problem. RNA primers are, as we noted, needed to start each of the Okazaki fragments of the lagging strand. The primers must be removed later, and the RNA nucleotides replaced with DNA nucleotides. When the RNA primer across from the end of the parental strand is removed, the RNA nucleotides cannot be replaced by DNA nucleotides because the DNA polymerase has no primer to start from. A short region of the template cannot, therefore be copied.

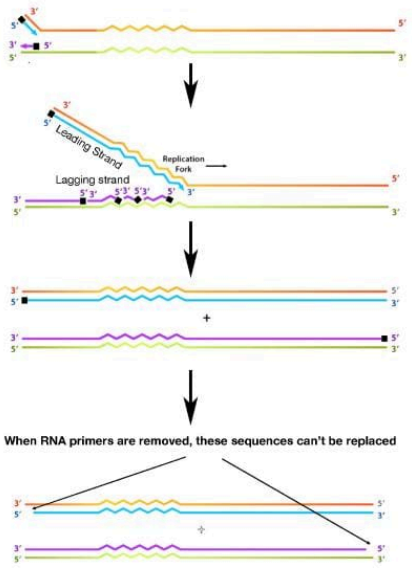


Figure 7.29 - Replication of a linear chromosome results in loss of sequences at the very ends with each round of replication

Telomerase

How can this problem be solved? It can be seen from Figure 7.29 that the end of the original template strand has a short 3' overhang resulting from the removal of the RNA primer across from it. In order to fill in this region, another primer would be needed, situated past the end of the template strand. But in order to build such a primer, it would be necessary for the template overhang to be longer. If it were possible to make the template strand longer, then another primer could be placed across from its end and the end of the strand could be copied. Such an extension of the template strand is exactly what happens in our reproductive cells. The parental template strand is extended by the enzyme telomerase, which adds telomere repeats and lengthens the template. We will see shortly how it accomplishes that feat.

RNA template

Telomerase is an unusual enzyme, in that it is made up of two components, an RNA and a reverse transcriptase. A reverse transcriptase is an RNA-dependent DNA polymerase, an enzyme that copies an RNA template to make DNA. The RNA component of the human telomerase, called hTERC, has a sequence that is complementary to the telomere repeat, TAGGG. As seen in Figure 7.31, this RNA can base-pair with the last telomere repeat on the parental DNA strand, while a portion of the RNA remains unpaired.



Figure 7.30 - Telomerase holding its RNA template (in center) - Wikipedia

Template for extension

The function of the unpaired region of the RNA is to serve as a template that can be used to extend the overhanging 3' end of the original DNA molecule. The protein component of telomerase has reverse transcriptase activity and can copy the RNA sequence into DNA. In human telomerase, the protein component is known as hTERT (telomerase reverse transcriptase). As seen in Figure 7.31 and 7.32, the reverse transcriptase extends the original 3' overhang using the RNA component as its template. The telomerase can then dissociate, and repeat the process multiple times to add many repeats of the telomere sequence.

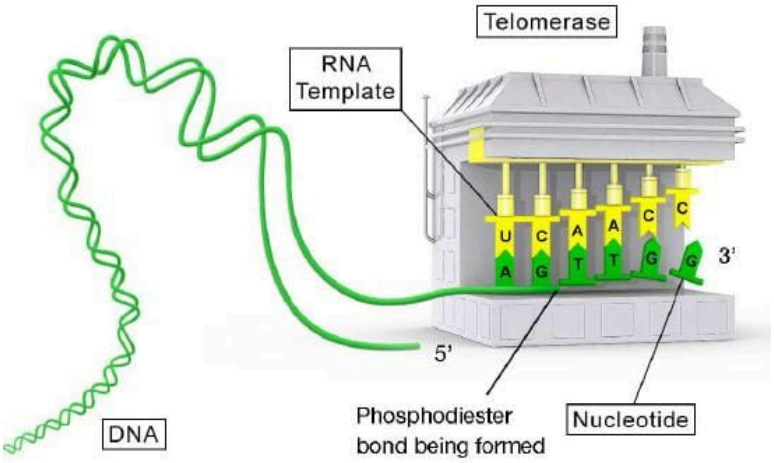


Figure 7.31 - Telomerase extends the template strand using its own RNA as template - Wikipedia

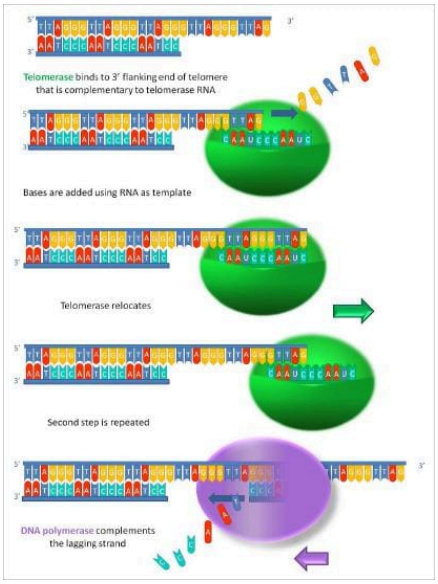


Figure 7.32 - Growing telomeres with telomerase and DNA polymerase. First, the 3' end of the template is extended by telomerase, then DNA polymerase completes synthesis of the lagging strand

Once the overhang has been extended by the addition of at least several telomere repeats, there is now room for the synthesis of an RNA primer complementary to the newly extended overhang (pointing back towards the rest of the chromosome). This primer can then be extended to complete synthesis of the lagging strand all the way to the end of the original parental DNA strand. Thus, the addition of telomere repeats on the parental DNA strands keeps the newly made DNA strands from becoming shorter with each cycle of replication. The fact that this happens in germ cells (reproductive cells) explains why each generation does not have shorter chromosomes than the parental generation.

The proofreading function of DNA polymerases monitors the accuracy of DNA replication while the enzyme telomerase keeps chromosomes that will be passed on to offspring from shortening. Between them, these two activities ensure that the genetic information is copied accurately, and that succeeding generations receive a full complement of the genetic information

Disassembly and reassembly of nucleosomes

The events of replication have an additional twist in eukaryotes. Recall that DNA is found in eukaryotic cells as chromatin, a complex of the DNA with proteins. At its least condensed, chromatin looks like a string of beads, consisting of the DNA wrapped around histone cores to make nucleosomes. The nucleosome structure must be disrupted to make DNA available for replication and restored after replication is completed (Figure 7.33).

Ahead of the replication fork, chromatin structure is disassembled by ATP-dependent chromatin remodeling complexes, allowing access to the DNA template. Once the new strands of DNA have been synthesized, both the original nucleosomes and new nucleosomes must be reassembled behind the replication fork. Since replication gives rise to two DNA molecules where there was one, twice the amount of histones is needed to package the DNA. Preparation for DNA replication, therefore, involves the synthesis of large amounts of histones to supply the need. Interestingly, it appears that newly synthesized DNA is packaged into nucleosomes using the original histones that were displaced to allow the replication fork to pass, as well as newly synthesized histones.

We also know that post-translational modifications like acetylation, methylation or phosphorylation of the histones can regulate the degree to which a given region of the genome is accessible for use. One question that remains the subject of intense research is how these modifications are accurately passed on to the new nucleosomes.

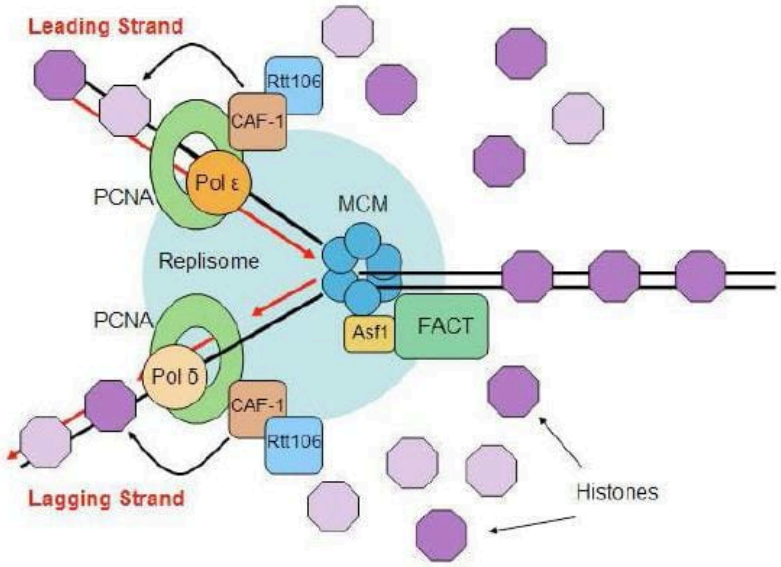


Figure 7.33 - Disruption of nucleosomes during DNA replication - Wikipedia

This page titled 7.3: DNA Replication is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.4: DNA Repair

Source: BiochemFFA_7_3.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy



Safeguarding the genome

In the last section we considered the ways in which cells deal with the challenges associated with replicating their DNA, a vital process for all cells. It is evident that if DNA is the master copy of instructions for an organism, then it is important not to make mistakes when copying the DNA to pass on to new cells. Although proofreading by DNA polymerases greatly increases the accuracy of replication, there are additional mechanisms in cells to further ensure that newly replicated DNA is a faithful copy of the original, and also to repair damage to DNA during the normal life of a cell.

DNA damage

All DNA suffers damage over time, from exposure to ultraviolet and other radiation, as well as from various chemicals in the environment (Figures 7.34 & 7.35). 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.

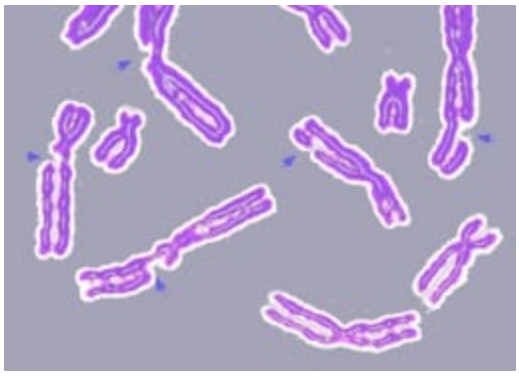


Figure 7.34 - Chromosome breaks - Wikipedia

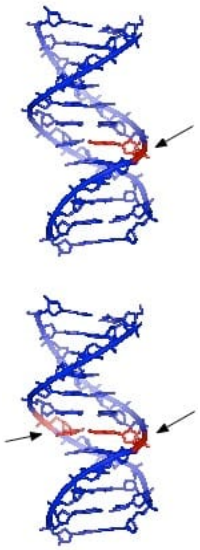


Figure 7.35 - Single- (top) and double-strand (bottom) damaged DNA - Wikipedia

Post-replicative mismatch repair

We earlier discussed proofreading by DNA polymerases during replication. While proofreading 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 polymerases? 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 10⁷ nucleotides in the absence of mismatch repair, this is further reduced a hundred-fold to one in 10⁹ 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 remove the region of the mismatch.
- Correctly fill in the gap created by the excision of the mismatch region.

Distinguishing strands

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.

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, DNA adenine methylase (Dam) that adds methyl groups on the to adenines in GATC sequences in DNA (Figure 7.36). Newly replicated DNA has not yet undergone methylation and thus, can be distinguished from the template strand, which is methylated.

The mismatch repair proteins selectively replace the strand lacking methylation, 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.

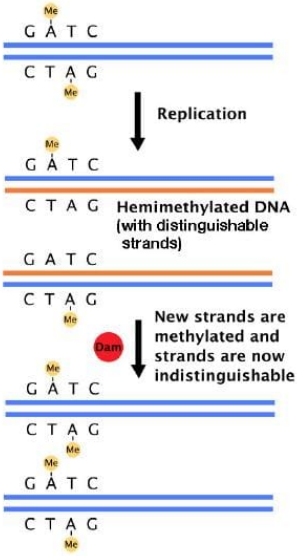


Figure 7.36- Dam methylase adds methyl groups at GATC sequences

Mismatch repair genes

Mismatch repair has been well studied in bacteria, and the proteins involved have been identified. In E.coli, mismatch repair proteins are encoded by a group of genes collectively known as the mut genes. Important components of the mismatch repair machinery are the proteins MutS, MutL and MutH (Figure 7.37).

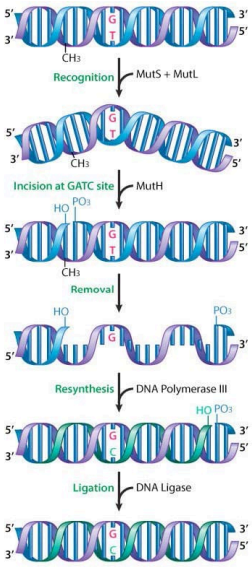


Figure 7.37 - Mismatch repair in E. coli - Image by Aleia Kim

MutS acts to recognize the mismatch, while MutL and MutH are recruited to the mismatch site by the binding of MutS. MutH is

an endonuclease that cuts the newly synthesized and, as yet, unmethylated DNA strand at a GATC. This activates a DNA helicase and an exonuclease that help unwind and remove the region containing the mismatch. DNA polymerase III fills in the gap, using the opposite strand as the template, and ligase joins the ends, to restore a continuous strand.

Eukaryotes also have a mismatch repair system that repairs not only single base mismatches but also insertions and deletions. Homologs to the E. coli MutS and MutL have been identified in other organisms, including humans: hMSH1 and hMSH2 (human MutS homolog 1 and 2) are homologous to MutS, while hMLH 1 is homologous to MutL. These, together with additional proteins, carry out mismatch repair in eukaryotic cells.

DNA methylation is not used by eukaryotic cells as a way to distinguish the new strand from the template, and it is not yet completely understood how the mismatch repair system in eukaryotes "knows" which strand to repair. There is evidence that the newly made DNA may be recognized by the fact that it is nicked, or discontinuous. This suggests that discontinuity resulting from Okazaki fragments that have not yet been joined together may permit the new strand to be distinguished from the old, continuous template strand.

Repairing damage to DNA

In the preceding section we looked at 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:
a. Radiation (e.g., UV rays in sunlight and in tanning booths, or ionizing radiation)
b. Exposure to damaging chemicals, such as nitrosamines or polycyclic aromatic hydrocarbons, in the environment (see Figure 7.38)
c. Chemical reactions within the cell (such as the deamination of cytosine to give uracil, or the methylation of guanine to produce methylguanine).

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 use tanning beds, or smoke, regularly.)

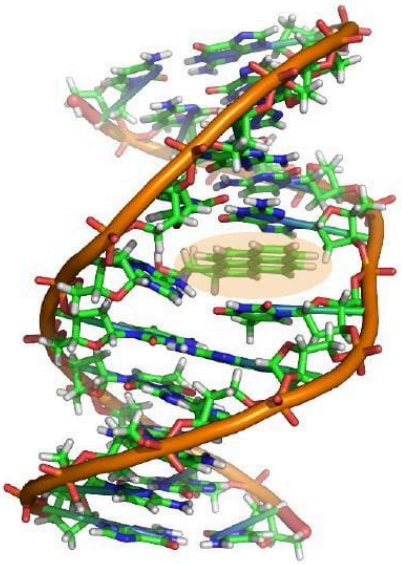


Figure 7.38 - DNA adducted to benz-o-pyrene (highlighted in orange)

Types of damage

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 cyclobutane pyrimidine dimers or CPDs (see Figure 7.39). Note that these are two neighboring pyrimidine bases on the same strand of DNA. UV exposure can also lead to the formation of another type of lesion, known as a (6-4) photoproduct or 6-4PP (Figure 7.39). Ionizing radiation can cause breaks in the DNA backbone, in one or both strands.

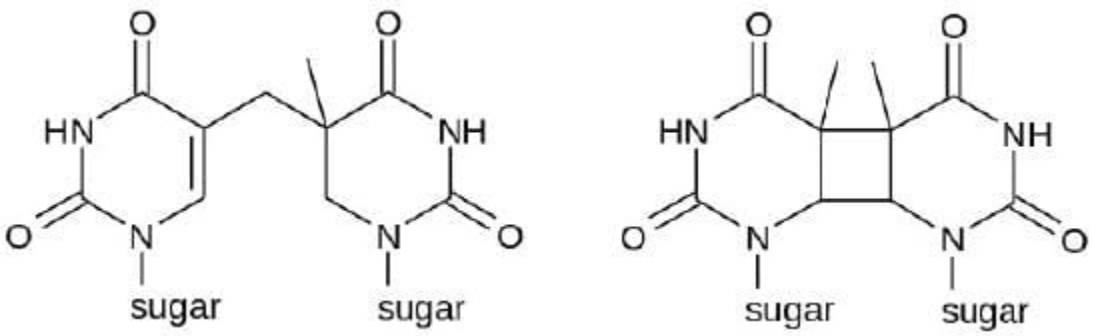


Figure 7.39 - Possible chemical structures of a pyrimidine dimer - 6-4PP (left) and CPD (right) - Wikipedia

Molecules like benzopyrene, found in automobile exhaust, can attach themselves to bases, forming bulky DNA adducts in which large chemical groups are linked to bases in the DNA. Damage like pyrimidine dimers, 6-4PPs or chemical adducts can physically distort the DNA helix, causing DNA and RNA polymerase to stall when they attempt to copy those regions of DNA (Figure 7.40).

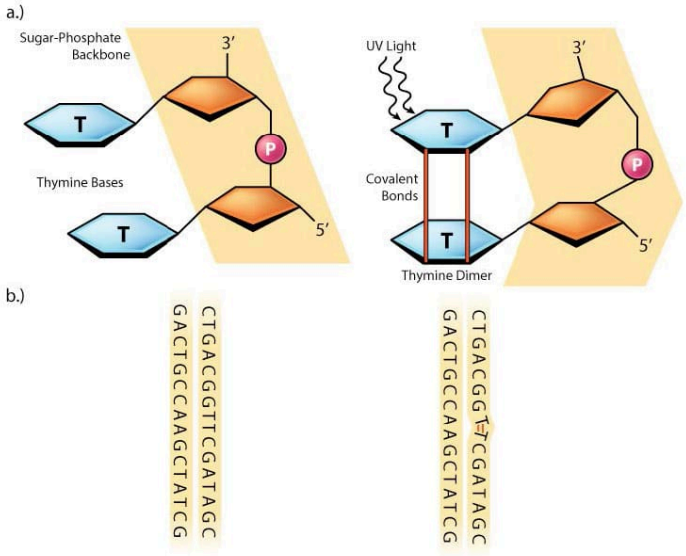


Figure 7.40- Thymine dimers distort the DNA helix -Aleia Kim

Chemical reactions occurring within cells can cause cytosines in DNA to be deaminated to uracil. Other sorts of damage in this category include the formation of oxidized bases like 8-oxo-guanine or alkylated bases like O6-methylguanine. 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. O6-methylguanine similarly can form base pairs with thymine instead of cytosine.

Removing damage

Cells have several ways to remove the sorts of damage described above. The first of these is described as **direct reversal**. Many organisms (though, unfortunately for us, not humans) can repair UV damage like CPDs and 6-4PPs because they possess enzymes called photolyases (photo=light; lyase=breakdown enzyme - Figure 7.41). Photolyases work through a process called photoreactivation, and use blue light energy to catalyze a photochemical reaction that breaks the aberrant bonds in the damaged DNA and returns the DNA to its original state.

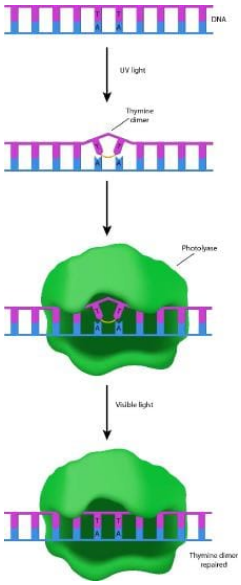


Figure 7.41- Direct repair of thymine dimers by photolyase - Pehr Jacobsen

Suicide enzyme

O6-methylguanine in DNA can also be removed by direct reversal, with the help of the enzyme O6-methylguanine methyltransferase. This is a very unusual enzyme that removes the methyl group from the guanine and transfers it onto a cysteine residue in the enzyme. The addition of the methyl group to the cysteine renders the enzyme non-functional.

As you know, most enzymes are catalysts that remain unchanged over the course of the reaction, permitting a single enzyme molecule to repeatedly catalyze a reaction. Because the O6-methylguanine methyltransferase does not fit this description, it is sometimes not regarded as a true enzyme. It has also been called a suicide enzyme, because the enzyme “dies” as a result of its own activity.

Excision repair

Excision repair is another common strategy. Excision repair is a general term for the cutting out and re-synthesizing of the damaged region of a DNA. There are several different kinds of excision repair, but they all involve excising the portion of the DNA that is damaged, followed by repair synthesis using the other strand as template, and finally, ligation to restore continuity to the repaired strand. Cells possess several different kinds of excision repair, each geared to specific kinds of DNA damage. Between them, these repair systems deal with the wide variety of insults to the genome.

Nucleotide excision repair

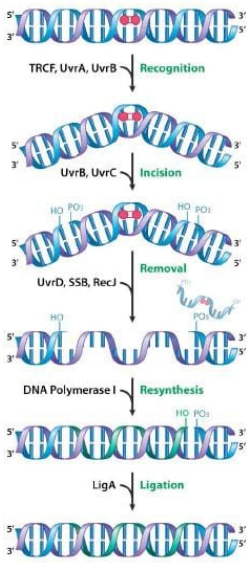


Figure 7.42 - Nucleotide excision repair - Aleia Kim

Nucleotide excision repair (NER) fixes damage such as the formation of chemical adducts, as well as UV damage. Both chemical adducts and the formation of CPDs or 6,4 photoproducts can cause significant distortion of the DNA helix. NER proteins act to cut the damaged strand on either side of the lesion. 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. Nucleotide excision repair has been extensively studied in bacteria.

In E. coli, recognition and excision of the damage is carried out by a group of proteins encoded by the uvrABC and uvrD genes. The protein products of the uvrA, uvrB and uvrC genes function together as the so-called UvrABC excinuclease. The damage is initially recognized and bound by a complex of the UvrA and UvrB proteins. Once the complex is bound, the UvrA dissociates, leaving the UvrB attached to the DNA, where it is then joined by the UvrC protein.

Strand nicking

It is the complex consisting of UvrB and C that acts to cut the phosphodiester backbone on either side of the damage, creating nicks in the strand about 12-13 nucleotides apart. A helicase encoded by uvrD then unwinds the region containing the damage, displacing it from the double helix together with UvrBC. The gap in the DNA is filled in by DNA polymerase, which copies the undamaged strand, and the nick is sealed with the help of DNA ligase.

Nucleotide excision repair is also an important pathway in eukaryotes. It is particularly important in the removal of UV damage in humans, given that we lack photolyases. A number of proteins have been identified that function in ways similar to the Uvr proteins.

The importance of these proteins is evident from the fact that mutations in the genes that encode them can lead to a number of genetic diseases, like Xeroderma pigmentosum, or XP. People with XP are extremely sensitive to UV exposure, because the damage caused by it cannot be repaired, leaving them at a much higher risk of developing skin cancer.

Two repair modes

Nucleotide excision repair operates in two modes, one known as global genomic repair and the other as transcription-coupled repair. While the function of both is to remove helix-destabilizing damage like cyclobutane pyrimidine dimers or chemical adducts, the way in which the lesions are detected differs.

In global genomic repair, damage is identified by surveillance of the entire genome for helix distorting lesions. In the case of transcription-coupled repair, the stalling of the RNA polymerase at a site of DNA damage is the indicator that activates this mode of nucleotide excision repair.

Base excision repair

Base excision repair (BER) is a repair mechanism that deals with situations like the deamination of cytosine to uracil (Figure 7.43) or the methylation of a purine base. These changes do not typically distort the structure of the DNA helix, unlike chemical adducts or UV damage.

In base excision repair a single damaged 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.

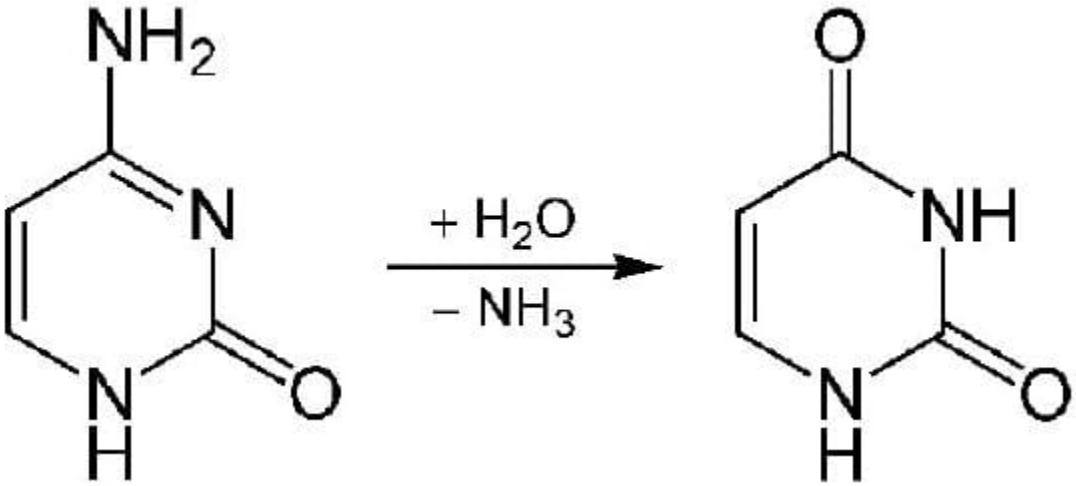


Figure 7.43 - Spontaneous hydrolysis of cytosine to form uracil

Uracil-DNA glycosylase

The removal of uracil from DNA is accomplished by the enzyme uracil-DNA glycosylase that can recognize uracil in DNA and break the glycosidic bond between the uracil and the sugar in the nucleotide (Figure 7.44). The removal of the base leaves a gap called an apyrimidinic site (AP site) because, in this case, uracil, a pyrimidine was removed. It is important to remember that at this point the backbone of the DNA is still intact, and the removal of a single base simply creates a gap like a tooth that has been knocked out.

The formation of the AP site triggers the activity of an enzyme known as an AP endonuclease that cuts the DNA backbone 5' to the AP site. In the remaining steps, a DNA polymerase binds to the nick, then using its exonuclease and polymerase activities, replaces the sequence in this region. Depending on the situation, a single nucleotide may be replaced (short patch BER) or a stretch of several nucleotides may be removed and replaced (long patch BER). Finally, as always, DNA ligase acts to seal the nick in the DNA.

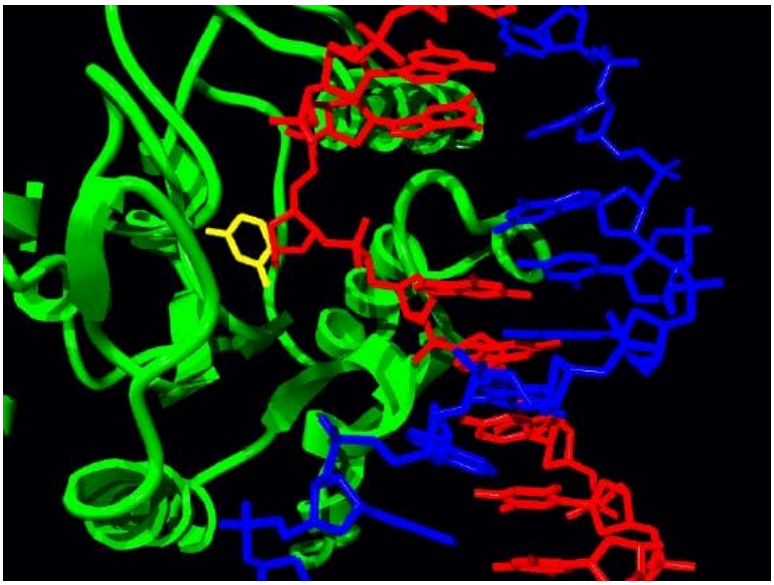


Figure 7.44 - Base excision repair by uracil-DNA glycosylase. The yellow uracil base has been “flipped” out of the DNA helix for excision

Repair of double-strand breaks

While all the repair mechanisms discussed so far fixed damage on one strand of DNA using the other, undamaged strand as a template, these mechanisms cannot repair damage to both strands. What happens if both strands are damaged? Ionizing radiation, exposure to certain chemicals, or reactive oxygen species generated in the cell can lead to double-strand breaks (DSBs) in DNA.

DSBs are a potentially lethal form of damage that, in addition to blocking replication and transcription, can also lead to chromosomal translocations, where part of one chromosome gets attached to a piece of another chromosome. Two different cellular mechanisms exist that help repair DSBs (Figure 7.45), homologous recombination (HR) and non-homologous end joining (NHEJ).

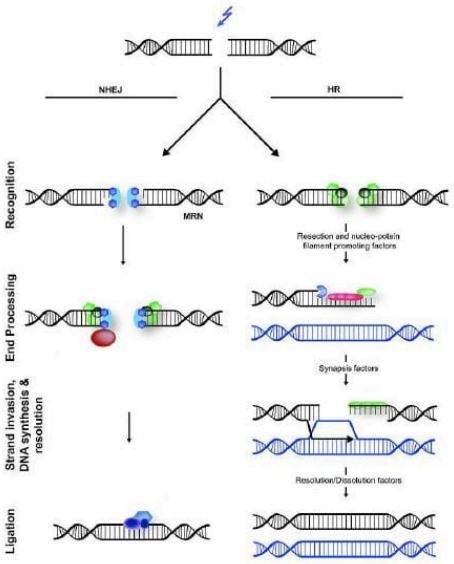


Figure 7.45 - Non-homologous end joining (left) versus homologous recombination (right) - Wikipedia

Homologous recombination repair commonly occurs in the late S and G2 phases of the cell, when each chromosome has been replicated and information from a sister chromatid can be used as a template to achieve error-free repair. Note that in contrast to excision repair, where the damaged strand was removed and the undamaged sister strand served as the template for filling in

the damaged region, HR must use the information from another DNA molecule, because both strands of the DNA are damaged in DSBs.

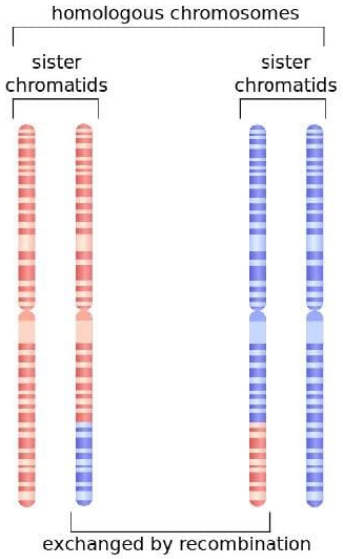


Figure 7.46 - Result of homologous recombination - Wikipedia

Nuclease action

Detection of the double-strand break triggers nuclease activity that chews back one strand on each end of the break. This results in the production of single-stranded 3' overhangs on each end. These single-stranded ends are bound by several proteins, creating a nucleoprotein filament that can then "search" for homologous (matching) sequences on a sister chromatid.

When such sequences are found, the nucleoprotein filament invades the undamaged sister chromatid, forming a crossover. This creates heteroduplexes made up of DNA strands from different chromatids. Strand invasion (Figure 7.47) is followed by branch migration, during which the Holliday junction moves along the DNA, extending the heteroduplex away from the original site of the crossover (Figure 7.48). In E. coli, branch migration depends on the activity of two proteins, RuvA and RuvB. The resulting recombination intermediate can be resolved, with the help of RuvC to give complete, error-free strands.

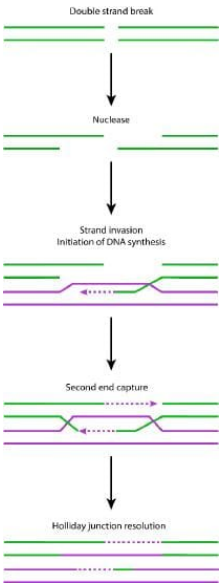


Figure 7.47 - Repair of DSBs by homologous recombination - Image by Pehr Jacobson

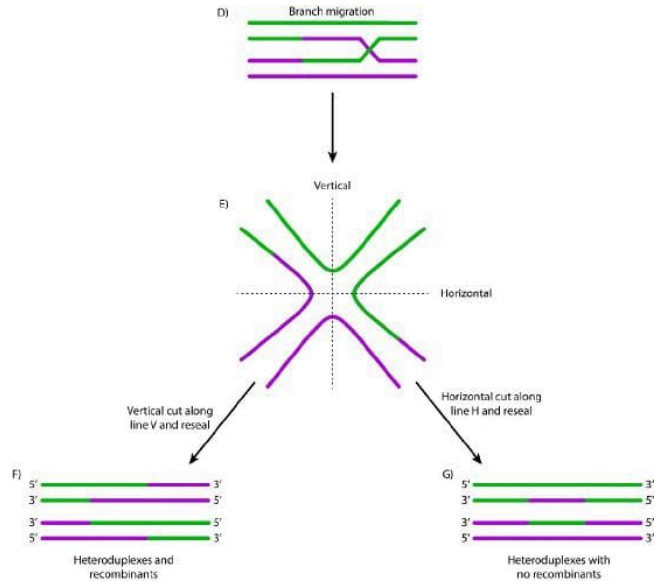


Figure 7.48 - Resolution of Holliday Junctions - Image by Pehr Jacobsen

Non-homologous end joining

In contrast to homologous recombination, Non-homologous end joining (NHEJ) is error-prone. It does not use or require a homologous template to copy, and works by simply chewing back the broken ends of DSBs and ligating them together. Not surprisingly, NHEJ introduces deletions in the DNA as a result.

Translesion DNA synthesis

As we have seen, cells have a variety of mechanisms to help safeguard the integrity of the information in DNA. One measure of last resort is translesion DNA synthesis, also known as bypass synthesis. Translesion synthesis occurs when a DNA polymerase encounters DNA damage on the template strand, but instead of stalling or skipping past the damage, replication switches to an error-prone mode, ignoring the template and incorporating random nucleotides into the new strand. In E.coli, translesion synthesis is dependent on the activities of proteins encoded by the umuC and umuD genes. Under the appropriate conditions (see SOS response, below) UmuC and UmuD are activated to begin bypass synthesis. Being error-prone, translesion synthesis gives rise to many mutations.

The SOS response

Named for standard SOS distress signals, the term “SOS repair” refers to a cellular response to UV damage. When bacterial cells suffer extensive damage to their DNA as a result of UV exposure, they turn on the coordinated expression of a large number of genes that are necessary for DNA repair. These include the uvr genes needed for nucleotide excision repair and recA, which is involved in homologous recombination. In addition to these mechanisms, which can carry out error-free repair, the SOS response can also induce the expression of translesion polymerases encoded by the dinA, dinB and umuCD genes.

How are all these genes induced in a coordinated way following UV damage? All of the genes induced in the SOS response are regulated by two components. The first is the presence of a short DNA sequence upstream of their coding region, called the SOS box. The second is a protein, the LexA repressor (Figure 7.49), that binds to the SOS box and prevents transcription of the downstream genes. Expression of the genes requires the removal of LexA from its binding site. How is this achieved?

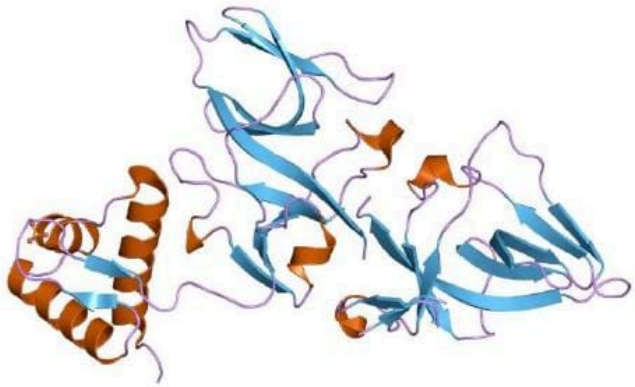


Figure 7.49 -Lex A repressor, N-terminal domain

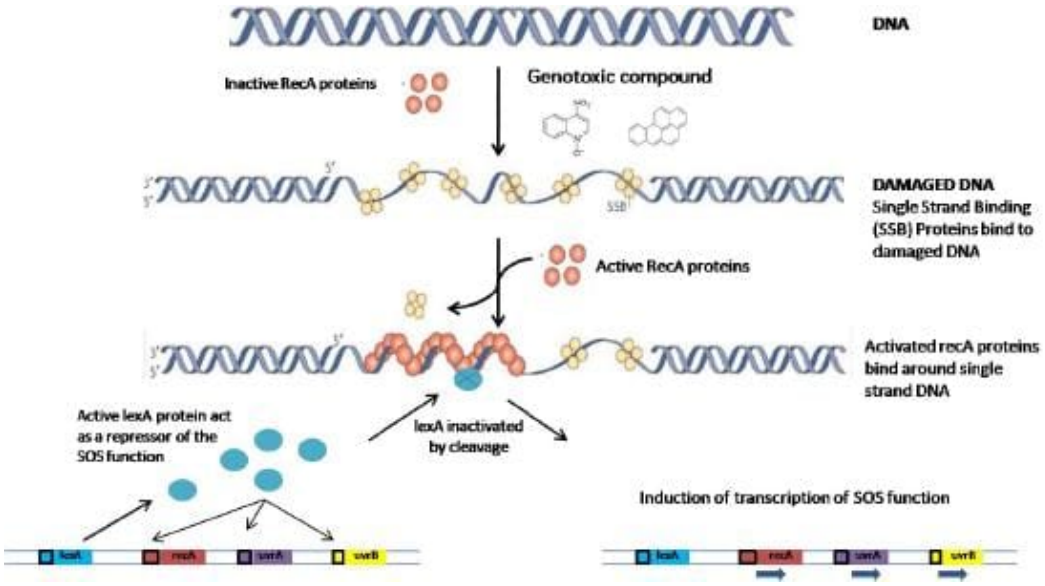


Figure 7.50 - Activation of SOS genes by lexA in E. coli

When exposure to radiation results in DNA breaks, the presence of single-stranded regions triggers the activation and binding of RecA proteins to the single-stranded region, creating a nucleoprotein filament. The interaction of the RecA with the LexA repressor leads to autocleavage of the repressor, allowing the downstream gene(s) to be expressed (Figure 7.50).

The genes controlled by the LexA repressor, as mentioned earlier, encode proteins that are necessary for accurate DNA repair as well as error-prone translesion synthesis. The various genes involved in DNA repair are induced in a specific order. In the initial stages, the repair genes that are derepressed are for nucleotide excision repair, followed by homologous recombination, both error-free mechanisms for repair. If the damage is too extensive to be repaired by these systems, error-prone repair mechanisms may be brought into play as a last resort.

SOS response and antibiotic resistance

The increased mutation rate in the SOS response may play a role in the acquisition of antibiotic resistance in bacteria (Figure 7.51).

An example is the development of resistance to topoisomerase poisons like the fluoroquinolone family of drugs. Fluoroquinolones inhibit the ability of topoisomerases to religate the ends of their substrates after nicking them to allow overwound DNA to relax. This results in accumulation of strand breaks that can trigger the SOS response. As a consequence of error-prone DNA synthesis

by low fidelity polymerases during the SOS response, there is a large increase in the number of mutations. While some mutations may be lethal to the bacteria, others can contribute to the rapid development of drug resistance in the population.

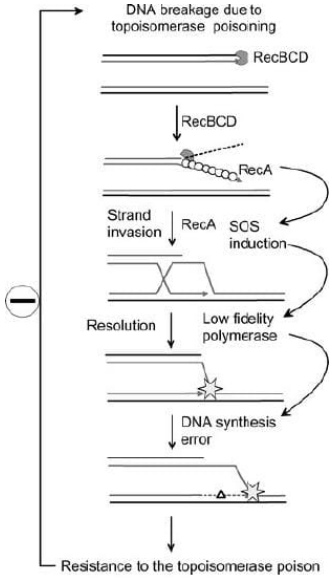


Figure 7.51 - A proposed mechanism of evolved bacterial antibiotic resistance - Wikipedia

This page titled 7.4: DNA Repair is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.5: Transcription

Source: BiochemFFA_7_4.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

In the preceding sections, we have discussed the replication of the cell's DNA and the mechanisms by which the integrity of the genetic information is carefully maintained. What do cells do with this information? How does the sequence in DNA control what happens in a cell? If DNA is a giant instruction book containing all of the cell's "knowledge" that is copied and passed down from generation to generation, what are the instructions for? And how do cells use these instructions to make what they need?

Information flow

You have learned in introductory biology courses that genes, which are instructions for making proteins, are made of DNA. You also know that information in genes is copied into temporary instructions called messenger RNAs that direct the synthesis of specific proteins. This description of flow of information from DNA to RNA to protein is often called the central dogma of molecular biology and is a good starting point for an examination of how cells use the information in DNA.

Consider that all of the cells in a multicellular organism have arisen by division from a single fertilized egg and therefore, all have the same DNA. Division of that original fertilized egg produces, in the case of humans, over a trillion cells, by the time a baby is produced from that egg (that's a lot of DNA replication!). Yet, we also know that a baby is not a giant ball of a trillion identical cells, but has the many different kinds of cells that make up tissues like skin and muscle and bone and nerves. How did cells that have identical DNA turn out so different?

The answer lies in gene expression, which is the process by which the information in DNA is used. Although all the cells in a baby have the same DNA, each different cell type uses a different subset of the genes in that DNA to direct the synthesis of a distinctive set of RNAs and proteins. The first step in gene expression is transcription, which we will examine next (Figure 7.52).

Transcription

Transcription is the process of copying information from DNA sequences into RNA sequences. This process is also known as DNA-dependent RNA synthesis. When a sequence of DNA is transcribed, only one of the two DNA strands is copied into RNA. We will consider what determines which strand of DNA is copied into RNA, later on.

But, apart from copying one, rather than both strands of DNA, how is transcription different from replication of DNA? DNA replication serves to copy all the genetic material of the cell and occurs before a cell divides, so that a full copy of the cell's genetic information can be passed on to the daughter cell. Transcription, by contrast, copies short stretches of the coding regions of DNA to make RNA. Different genes may be copied into RNA at different times in the cell's life cycle. RNAs are, essentially, temporary copies of the information in DNA and different sets of instructions are copied for use at different times and in different cell types.

Cells make several different kinds of RNA:

- mRNAs that code for proteins
- rRNAS that form part of ribosomes
- tRNAs that serve as adaptors between mRNA and amino acids during translation
- Small RNAs that regulate gene expression, including miRNAs and siRNAs
- Other small RNAs that have a variety of functions, including the small nuclear RNAs that are part of the splicing machinery.
- Long noncoding RNAs (lnc RNAs - Figure 7.53)

Building an RNA strand is very similar to building a DNA strand. This is not surprising, knowing that DNA and RNA are very similar molecules. Transcription is catalyzed by the enzyme RNA Polymerase. "RNA polymerase" is a general term for an enzyme that makes RNA. There are several different kinds of RNA polymerases in eukaryotic cells, while in prokaryotes, a single type of RNA polymerase is responsible for all transcription.

RNA synthesis

Like DNA polymerases, RNA polymerases synthesize new strands only in the 5' to 3' direction, but because they are making RNA, they use ribonucleotides (i.e., RNA nucleotides - Figure 7.54) rather than deoxyribonucleotides. Ribonucleotides are joined in exactly the same way as deoxyribonucleotides, i.e., the 3'OH of the last nucleotide on the growing chain is joined to the 5' phosphate on the incoming nucleotide to make a phosphodiester bond.

One important difference between DNA polymerases and RNA polymerases is that the latter do not require a primer to start making RNA. Once RNA polymerases are in the right place to start copying DNA, they just begin making RNA by joining together RNA nucleotides complementary to the DNA template.

Starting points

This, of course, brings us to an obvious question- how do RNA polymerases "know" where to start copying on the DNA?

Unlike the situation in replication, where every nucleotide of the parental DNA must eventually be copied, transcription, as we have already noted, only copies selected portions of the DNA into RNA at any given time. Consider the challenge here: in a human cell, there are approximately 6 billion base-pairs of DNA. Much of this is non-coding DNA, meaning that it will not need to be transcribed. The small percentage of the genome that is made up of coding sequences still amounts to between 20,000 and 30,000 genes in each cell. Of these genes, only a small number will need to be expressed at any given time.What indicates to an RNA polymerase where to start copying DNA to make a transcript?

Promoters

It turns out that patterns in the DNA sequence indicate where RNA polymerase should start and end transcription. These sequences are recognized by the RNA polymerase or by proteins that help RNA polymerase determine where it should bind the DNA to start transcription. A DNA sequence at which the RNA polymerase binds to start transcription is called a promoter. The DNA sequence that indicates the endpoint of transcription, where the RNA polymerase should stop adding nucleotides and dissociate from the template is known as a terminator sequence. The promoter and terminator, thus, bracket the region of the DNA that is to be transcribed.

A promoter is described as being situated upstream of the gene that it controls (Figure 7.57). What this means is that on the DNA strand that the gene is on, the promoter sequence is "before" the gene, or to put it differently, it is on the side of the gene opposite to the direction of transcription. Also notice that the promoter is said to "control" the gene it is associated with. This is because expression of the gene is dependent on the binding of RNA polymerase to the promoter sequence to begin transcription. If the RNA polymerase and its helper proteins do not bind at the promoter, the gene cannot be transcribed and it will therefore, not be expressed.

What is special about a promoter sequence? In an effort to answer this question, scientists examined many genes and their surrounding sequences (Figure 7.57). Because the same RNA polymerase has to bind to many different promoters, it would be predicted that promoters would have some similarities in their sequences. As expected, common sequence patterns were seen to be present in many promoters.

We will first take a look at prokaryotic promoters.

When prokaryotic genes were examined, the following features commonly emerged:

- A transcription start site (this the base in the DNA across from which the first RNA nucleotide is paired), which, by convention, is denoted as +1.
- A -10 sequence: this is a 6 bp region centered about 10 bp upstream of the start site. The consensus sequence at this position is TATAAT. In other words, if you count back from the transcription start site, the sequence found at roughly -10 in the majority of promoters studied is TATAAT.
- A -35 sequence: this is a 6 bp sequence at about 35 basepairs upstream from the start of transcription. The consensus sequence at this position is TTGACA.

It is important to understand that each nucleotide in a consensus sequence is simply the one that appeared at that position in the majority of promoters examined, and does not mean that the entire consensus sequence is found in all promoters. In fact, few

promoters have -10 and -35 sequences that exactly match the consensus. The box at the left shows the -10 and -35 sequences by percentage of occurrence of each base in the promoter.

What is the significance of these sequences? It turns out that the sequences at -10 and -35 are necessary for recognition of the promoter region by RNA polymerase (Figure 7.58). The sequences at -10 and -35 may vary a little in individual promoters, as mentioned above, but the extent to which they are different is limited. It is only when the RNA polymerase has stably bound at the promoter that transcription can begin. The process by which the promoter is recognized and bound stably has been well studied for the RNA polymerase of E. coli.

Core polymerase and holoenzyme

The E. coli RNA polymerase is made up of a core enzyme of five subunits (α2ββ'and ω) and an additional subunit called the σ (sigma) subunit. Together, the σ subunit and core polymerase make up what is termed the RNA polymerase holoenzyme. The core polymerase is the part of the RNA polymerase that is responsible for the actual synthesis of the RNA, while the σ subunit is necessary for binding of the enzyme at promoters to initiate transcription.

Loose association

The core polymerase and σ subunit are not always associated with each other. For the most part, the core polymerase is loosely associated with DNA, although it does not discriminate between promoters and other sequences in DNA, and the DNA strands are not opened up to allow transcription in this state. The role of the σ subunit is to reduce the affinity of the core polymerase for non-specific DNA sequences and to help the enzyme specifically bind to promoter sequences.

Holoenzyme binding

It is when the σ subunit associates with the core polymerase that the holoenzyme is able to bind specifically to promoter sequences. The initial binding of the holoenzyme at the promoter results in what is called a “closed” complex, meaning that the DNA template is still double-stranded and has not opened up to allow transcription. This closed complex is then converted to an “open” complex by the separation of the DNA strands to create a transcription bubble about 12-14 base-pairs long (Figure 7.60). The conversion of the closed complex to the open complex also requires the presence of the σ subunit.

Open complex & initiation

Once the open complex has formed, the DNA template can begin to be copied, and the core polymerase adds nucleotides complementary to one strand of the DNA. At this stage, known as initiation, the polymerase adds several nucleotides while still bound to the promoter, and without moving along the DNA template. Initially, short pieces of RNA a few nucleotides long may be made and released, without the polymerase leaving the promoter. Eventually, the enzyme makes the transition to the next stage, elongation, when an RNA of 8-9 bases is made and the enzyme moves beyond the promoter region.

Elongation

Once elongation commences and the RNA polymerase is moving down the DNA template, the σ subunit is no longer necessary and may dissociate from the core enzyme. The core polymerase can move along the template, unwinding the DNA ahead of it to maintain a transcription bubble of 12-15 base-pairs and synthesizing RNA complementary to one of the strands of the DNA. As already mentioned, an RNA chain, complementary to the DNA template, is built by the RNA polymerase by the joining of the 5' phosphate of an incoming ribonucleotide to the 3'OH on the last nucleotide of the growing RNA strand. Behind the RNA polymerase, the DNA template is rewound, displacing the newly made RNA from its template strand.

Termination

As mentioned earlier, a sequence of nucleotides called the terminator is the signal to the RNA polymerase to stop transcription and dissociate from the template. Some terminator sequences, known as intrinsic terminators, allow termination by RNA polymerase without the help of any additional factors, while others, called rho-dependent terminators, require the assistance of a protein factor called rho (ρ).

How does the sequence of the terminator cause the RNA polymerase to stop adding nucleotides and release the transcript?

To understand this, it is useful to know that the terminator sequence precedes the last nucleotide of the transcript. In other words, the terminator is part of the end of the sequence that is transcribed (Figure 7.61).

Intrinsic terminators

In intrinsic terminators, this sequence in the RNA has self-complementary regions that can base-pair with each other to form a hairpin structure that contains a GC-rich run in the “stem” of the hairpin. This hairpin is followed by a single-stranded region that is rich in U’s (Figure 7.62). The secondary structure formed by the folding of the end of the RNA into the hairpin causes the RNA polymerase to pause. Meanwhile, the run of U’s at the end of the hairpin permits the RNA-DNA hybrid in this region to come apart, because the base-pairing between A’s in the DNA template and the U’s in the RNA is relatively weak. This allows the transcript to be released from the DNA template and from the RNA polymerase.

Rho-dependent termination

In the case of rho-dependent termination, an additional protein factor, rho, is necessary. Rho is a helicase that can separate the transcript from the template it is paired with. As in intrinsic termination, rho-dependent termination requires the formation of a hairpin structure in the RNA that causes pausing of the RNA polymerase. Meanwhile, rho binds to a region of the transcript called the rho utilization site (rut) and moves along the RNA till it reaches the paused RNA polymerase. It then acts on the RNA-DNA hybrid, releasing the transcript from the template.

Coupled transcription and translation

In prokaryotes, which lack a nucleus, the DNA is not separated from the rest of the cell in a separate compartment, so the mRNA is immediately available to the translation machinery, as the transcript is coming off the template DNA. Indeed, in prokaryotic cells, translation of the mRNA can begin before the entire gene has been transcribed. Ribosomes can assemble at the 5’ end of the transcript, as it is displaced from the template, while the 3’ end of the gene is still being copied. The lag time between transcription and translation is thus, very short in prokaryotes.

Transcription in eukaryotes

Although the process of RNA synthesis is the same in eukaryotes as in prokaryotes, there are some additional considerations in eukaryotes. One is that in eukaryotes, the DNA template exists as chromatin, where the DNA is tightly associated with histones and other proteins. The "packaging" of the DNA must therefore be opened up to allow the RNA polymerase access to the template in the region to be transcribed (Figure 7.63). The restructuring of chromatin to allow access to regions of DNA is thus an important factor in determining which genes are expressed.

Multiple RNA polymerases

A second difference is that eukaryotes have multiple RNA polymerases, not just one as in bacterial cells. The different eukaryotic polymerases transcribe different classes of genes. For example, RNA polymerase I transcribes the ribosomal RNA genes, while RNA polymerase III copies tRNA genes. The RNA polymerase we will focus on most is RNA polymerase II, which transcribes protein-coding genes to make mRNAs.

All three eukaryotic RNA polymerases need additional proteins to help them get transcription started. In prokaryotes, RNA polymerase by itself can initiate transcription (the σ subunit is a subunit of the RNA polymerase, not an entirely separate protein). The additional proteins needed by eukaryotic RNA polymerases are referred to as transcription factors. We will see below that there are various categories of transcription factors.

Transcription and translation are de-coupled

Finally, in eukaryotic cells, transcription is separated in space and time from translation. Transcription happens in the nucleus, and the RNAs produced are processed further before they are sent into the cytoplasm.
Protein synthesis (translation) happens in the cytoplasm. As noted earlier, in prokaryotic cells, mRNAs can be translated as they are coming off the DNA template, and because there is no nuclear envelope, transcription and protein synthesis occur in a single cellular compartment. A representative eukaryotic gene, depicted in Figure 7.64 shows that transcription starts some 25 bp downstream of the TATA box, and creates a transcript that begins with a 5’ untranslated region (5’UTR) followed by the coding region which may include multiple introns and ending in a 3’ untranslated region or 3‘UTR (Figure 7.64). As detailed below, the initial transcript is further processed before it is used.

Eukaryotic promoters

Like genes in prokaryotes, eukaryotic genes also have promoters that determine where transcription will begin. As with prokaryotes, there are specific sequences in the promoter regions that are recognized and bound by proteins involved in the

initiation of transcription. We will focus primarily on the genes encoding proteins that are transcribed by RNA polymerase II. Such promoters commonly have a TATA box, a sequence similar to the -10 sequence in prokaryotic promoters. The TATA box is a sequence about 25-35 basepairs upstream of the start of transcription (+1). (Some eukaryotic promoters lack TATA boxes, and have, instead, other recognition sequences, known as DPE, or downstream promoter elements.) Interestingly, the TATA box is not directly recognized and bound by RNA polymerase II. Instead, this sequence is bound by other proteins that, together with the RNA polymerase, form the transcription initiation complex.

Eukaryotic promoters also have, in addition, several other short stretches of sequences, that affect transcription, within about 100 to 200 base-pairs upstream of the transcription start site. These sequences, which are sometimes called upstream elements or promoter-proximal upstream elements, are bound by activator proteins that interact with the transcription complex that forms at the TATA box. Examples of such upstream elements are the CAAT box and the GC box (Figure 7.64).

Making transcripts in eukaryotes

We noted earlier that all eukaryotic RNA polymerases need additional proteins to bind promoters and start transcription. The proteins that help eukaryotic RNA polymerases find promoter sites and initiate RNA synthesis are termed general transcription factors. We will focus on the transcription factors that assist RNA polymerase II, the enzyme that transcribes protein-coding genes. These transcription factors are named TFIIA, TFIIB and so on (TF= transcription factor, II=RNA polymerase II, and the letters distinguish individual transcription factors).

Transcription by RNA polymerase II requires the general transcription factors and the RNA polymerase to form a complex, at the TATA box, called the basal transcription complex or transcription initiation complex (Figure 7.65).

This is the minimum requirement for any gene to be transcribed. The first step in the formation of this complex is the binding of the TATA box by a transcription factor, TFIID. TFIID is made up of several proteins, one of which is called the TATA Binding Protein or TBP. Binding of the TBP causes the DNA to bend at this spot and take on a structure that is suitable for the binding of additional transcription factors and RNA polymerase.

Interestingly, the binding of the TBP is a necessary step in forming a transcription initiation complex even when the promoter lacks a TATA box. The order of binding of additional proteins after binding of the TBP, as determined through in vitro experiments, appears to be TFIIB, followed by TFIIF and RNA polymerase II, then TFIIE. The final step in the assembly of the basal transcription complex is the binding of a general transcription factor called TFIIH. Some evidence suggests that following the binding of the TBP to DNA, the rest of the proteins in the initiation complex may assemble as a very large complex that then binds directly to the DNA. In any case, the presence of all of these general transcription factors and RNA Polymerase II bound at the promoter is necessary for the initiation of transcription.

As in prokaryotic transcription, once the RNA polymerase binds, it can begin to assemble a short stretch of RNA. This must be followed by promoter clearance, in order to move down the template and elongate the transcript. This requires the action of TFIIH. TFIIH is a multifunctional protein that has helicase activity (i.e., it is capable of opening up a DNA double helix) as well as kinase activity. The kinase activity of TFIIH adds phosphates onto the C-terminal domain (CTD) of the RNA polymerase II. This phosphorylation appears to be the signal that releases the RNA polymerase from the basal transcription complex and allows it to move forward on the template, building the new RNA as it goes (Figure 7.66).

Termination of transcription is not as well understood as it is in prokaryotes. Termination does not occur at a fixed distance from the 3' end of mature RNAs. Rather, it seems to occur hand in hand with the processing of the 3' end of the primary transcript. The polyadenylation signal in the 3' untranslated region of the transcript appears to play a role in RNA polymerase pausing, and subsequent release of the completed primary transcript. Recognition of the polyadenylation signal triggers the binding of proteins involved in 3'end processing and termination.

Information Processing: Transcription

748

YouTube Lectures

by Kevin

HERE & HERE

749

Figure 7.52 - Overview of eukaryotic transcription

Wikipedia

750

Figure 7.53 - Transcripts may code for protein or may be functional as RNAs

Wikipedia

751

Figure 7.55 - RNA (green) being synthesized from DNA template (blue strand) by T7 RNA polymerase (purple). The non-template DNA strand is in red.

Figure 7.54 - The four ribonucleotides for making RNA

752

Figure 7.56 - Central dogma - DNA to RNA to protein

Wikipedia

753

YouTube Lectures

by Kevin

HERE & HERE

Figure 7.57 - Sequences upstream of transcription start site in several prokaryotic genes

Image by Martha Baker

754

-10 Sequence

T A T A A T

77% 76% 60% 61% 56% 82%

-35 Sequence

T T G A C A

69% 70% 61% 56% 54% 54%

Figure 7.58 - RNA polymerase promoter binding

Wikipedia

Figure 7.59 - A bacterial RNA polymerase (α2ββ'and ω)

755

Figure 7.60 - Synthesis of RNA in the transcription bubble

756

Figure 7.61 - Promoter and Terminator sequences determine where transcription starts and ends.

757

Figure 7.62 Transcription termination by intrinsic (top) and rho-dependent (bottom) mechanisms

Wikipedia

758

Figure 7.63 - Eukaryotic DNA is complexed with proteins in chromatin

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

759

Figure 7.64 - Region surrounding the transcriptional start site in eukaryotic DNA

Figure 7.65 - Transcription pre-initiation complex in eukaryotes

Wikipedia

760

761

This page titled 7.5: Transcription is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.6: RNA Processing

Source: BiochemFFA_7_5.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

So far, we have looked at the mechanism by which the information in genes (DNA) is transcribed into RNA. The newly made RNA, also known as the primary transcript is further processed before it is functional. Both prokaryotes and eukaryotes process their ribosomal and transfer RNAs. The major difference in RNA processing, however, between prokaryotes and eukaryotes, is in the processing of messenger RNAs. We will focus on the processing of mRNAs in this section. You will recall that in bacterial cells, the mRNA is translated directly as it comes off the DNA template. In eukaryotic cells, RNA synthesis, which occurs in the nucleus, is separated from the protein synthesis machinery, which is in the cytoplasm. The initial product of transcription of an mRNA is sometimes referred to as the pre-mRNA. After it has been processed and is ready to be exported from the nucleus, it is called the mature mRNA. The three main processing steps for mRNAs are (Figure 7.67):

- Capping at the 5' end
- Splicing to remove introns
- Addition of a polyA tail at the 3' end.

Although this description suggests that these processing steps occur post-transcriptionally, after the entire gene has been transcribed, there is evidence that processing occurs co-transcriptionally. That is, the steps of processing are occurring as the mRNA is being made. Proteins involved in mRNA processing have been shown to be associated with the phosphorylated C-terminal domain (CTD) of RNA polymerase II.

Capping

As might be expected, the addition of an mRNA cap at the 5’ end is the first step in mRNA processing, since the 5’end of the RNA is the first to be made. Capping occurs once the first 20-30 nucleotides of the RNA have been synthesized. The addition of the cap involves removal of a phosphate from the first nucleotide in the RNA to generate a diphosphate. This is then joined to a guanosine monophosphate which is subsequently methylated at N7 of the guanine to form the 7mG cap structure (Figure 7.68). This cap is recognized and bound by a complex of proteins that remain associated with the cap till the mRNA has been transported into the cytoplasm. The cap protects the 5' end of the mRNA from degradation by nucleases and also helps to position the mRNA correctly on the ribosomes during protein synthesis.

Splicing

Eukaryotic genes have introns, noncoding regions that interrupt the gene. The mRNA copied from genes containing introns will also therefore have noncoding regions that interrupt the information in the gene. These noncoding regions must be removed (Figure 7.69) before the mRNA is sent out of the nucleus to be used to direct protein synthesis.

Intron removal

Introns are removed from the pre-mRNA by the activity of a complex called the spliceosome. The spliceosome is made up of proteins and small RNAs that are associated to form protein-RNA enzymes called small nuclear ribonucleoproteins or snRNPs (pronounced snurps).

Splice junctions

The splicing machinery must be able to recognize splice junctions (i.e., where each exon ends and its associated intron begins) in order to correctly cut out the introns and join the exons to make the mature, spliced mRNA. What signals indicate exon-intron boundaries? The junctions between exons and introns are indicated by specific base sequences. The consensus sequence at the 5’ exon-intron junction (also called the 5’ splice site) is AGGURAGU. In this sequence, the intron starts with the second G (R stands for any purine). The 3' splice junction has the consensus sequence YAGRNNN, where YAG is within the intron, and RNNN

is part of the exon (Y stands for any pyrimidine, and N for any nucleotide).

There is also a third important sequence within the intron, about a hundred nucleotides from the 3' splice site, called a branch point or branch site, that is important for splicing. This site is defined by the presence of an A followed by a string of pyrimidines. The importance of this site will be seen when we consider the steps of splicing.

Splicing mechanism

There are two main steps in splicing. The first step is the nucleophilic attack by the 2'OH of the branch point A on the 5' splice site (the junction of the 5' exon and the intron). As a result of a trans-esterification reaction, the 5' exon is released, and a lariat-shaped molecule composed of the 3' exon and the intron sequence is generated (Figure 7.70). In the second step, the 3' OH of the 5' exon attacks the 3' splice site, and the two exons are joined together, and the lariat-shaped intron is released .

Spliceosome

As mentioned earlier, splicing is carried out by a complex consisting of small RNAs and proteins. The five small RNAs crucial to this complex, U1, U2, U4, U5 and U6 are found associated with proteins, as snRNPs. These and many other proteins work together to facilitate splicing. Although many details remain to be worked out, it appears that components of the splicing machinery associate with the CTD of the RNA polymerase and that this association is important for efficient splicing. The assembly of the spliceosome requires the stepwise interaction of the various snRNPs and other splicing factors (Figure 7.71). The initial step in this process is the interaction of the U1 snRNP with the 5' splice site. Additional proteins such as U2AF (AF = associated factor) are also loaded onto the pre-mRNA near the branch site. This is followed by the binding of the U2 snRNA to the branch site.

Next, a complex of the U4/U6 and U5 snRNPs is recruited to the spliceosome to generate a pre-catalytic complex. This complex undergoes rearrangements that alter RNA-RNA and protein-RNA interactions, resulting in displacement of the U4 and U1 snRNPs and the formation of the catalytically active spliceosome. This complex then carries out the two splicing steps described earlier.

Alternative splicing

On average, human genes have about 9 exons each. However, the mature mRNAs from a gene containing nine exons may not include all of them. This is because the exons in a pre-mRNA can be spliced together in different combinations to generate different mature mRNAs. This is called alternative splicing, and allows the production of many different proteins using relatively few genes, since a single RNA with many exons can, by combining different exons during splicing, create many different protein coding messages. Because of alternative splicing, each gene in our DNA gives rise, on average, to three different proteins. Alternative splicing allows the information in a single gene to be used to specify different proteins in different cell types or at different developmental stages (Figure 7.72).

Polyadenylation

The 3' end of a processed eukaryotic mRNA typically has a "poly(A) tail" consisting of about 200 adenine-containing nucleotides. These residues are added by a template-independent enzyme, poly(A)polymerase, following cleavage of the RNA at a site near the 3' end of the new transcript. Components of the polyadenylation machinery have been shown to be associated with the CTD of the RNA polymerase, showing that all three steps of pre-mRNA processing are tightly linked to transcription. There is evidence that the polyA tail plays a role in efficient translation of the mRNA, as well as in the stability of the mRNA. Like alternative splice sites, genes can have alternative polyA sites as well (Figure 7.73).

The cap and the polyA tail on an mRNA are also indications that the mRNA is complete (i.e., not defective). Once protein-coding messages have been processed by capping, splicing and addition of a poly A tail, they are transported out of the nucleus to be translated in the cytoplasm. Mature mRNAs are sent into the cytoplasm bound to export proteins that interact with the nuclear pore complexes in the nuclear envelope (Figure 7.74). Once the mature mRNA has been translocated to the cytoplasm, it is ready to be translated.

RNA editing

In addition to undergoing the three processing steps outlined above, many RNAs undergo further modification called RNA editing. Editing has been observed in not only mRNAs but also in transfer RNAs and ribosomal RNAs. As the name suggests, RNA editing is a process during which the sequence of the transcript is altered post-transcriptionally. A well-studied example of RNA editing is the alteration of the sequence of the mRNA for apolipoprotein B (see also HERE). The editing results in the deamination of a cytosine in the transcript to form a uracil, at a specific location in the mRNA. This change converts the codon at this position, CAA, which encodes a glutamine, into UAA, a stop codon. The consequence of this is that a shorter version of the protein is made, when the edited transcript is translated. It is interesting that the editing of this transcript occurs in intestinal cells but not in liver cells. Thus, the protein product of the apolipoprotein B gene is longer in the liver than it is in the intestine.

Insertion/deletion

Another kind of RNA editing involves the insertion or deletion of one or more nucleotides. One example of this sort of editing is seen in the mitochondrial RNAs of trypanosomes. Small guide RNAs indicate the sites at which nucleotides are inserted or deleted to produce the mRNA that is eventually translated (Figure 7.75).

The effect of either of these kinds of editing on the mRNA is that the encoded protein product is different, providing another point at which the product of expression of a gene can be controlled.

tRNA synthesis & processing

tRNAs are synthesized by RNA polymerase III, which makes precursor molecules called pre-tRNA that then undergo processing to generate mature tRNAs. The initial transcripts contain additional RNA sequences at both the 5’ and 3’ ends. Some pre-tRNAs also contain introns. These additional sequences are removed from the transcript during processing.

The 5’ leader sequence of the pre-tRNA (the additional nucleotides at the 5’-end) is removed by an unusual endonuclease called ribonuclease P (RNase P - Figure 7.76). RNase is a ribonucleoprotein complex composed of a catalytic RNA and numerous proteins. The 3’ trailer sequence (extra nucleotides at the 3’ end of the pre-tRNA) is later removed by different nucleases. All tRNAs must have a 3’ CCA sequence that is necessary for the charging of the tRNAs with amino acids. In bacteria, this CCA sequence is encoded in the tRNA gene, but in eukaryotes, the CCA sequence is added post-transcriptionally by an enzyme called tRNA nucleotidyl transferase (tRNT).

Introns

As mentioned earlier, some tRNA precursors contain an intron located in the anticodon arm. In eukaryotes, this intron is typically found immediately 3’ to the anticodon. The introns is spliced out with the help of a tRNA splicing endonuclease and a ligase.

Base modifications

Mature tRNAs contain a high proportion of bases other than the usual adenine (A), guanine (G), cytidine (C) and uracil (U). These unusual bases are produced by modifying the bases in the tRNA to form variants, such as pseudouridine (Figure 7.77) or dihyrouridine. Modifications to the bases are introduced into the tRNA at the final processing step by a variety of specialized enzymes. Different tRNAs have different subsets of modifications at specific locations, often the first base of the anti-codon (the wobble position).

rRNA synthesis and processing

Cells contain many copies of rRNA genes (between 100 and 2000 copies are seen in mammalian cells). These genes are organized in transcription units separated by non-transcribed spacers. Each transcription unit contains sequences coding for 18S, 5.8S and 28S rRNA, and is transcribed by RNA polymerase I into a single long transcript (47S). The 5S rRNA is separately transcribed. The sizes of ribosomal RNAs are, by convention, indicated by their sedimentation coefficients, which is a measure of their rate of sedimentation during centrifugation. Sedimentation is expressed in Svedberg units (hence the S at the end of the number) with larger numbers indicating greater mass.

The initial transcript contains 5’ and 3’ external transcribed spacers (ETS) as well as internal transcribed sequences (ITS). The primary transcript is first trimmed at both ends by nucleases to give a 45S pre-rRNA. Further processing of the pre-rRNA through cleavages guided by RNA-protein complexes containing snoRNAs (small nucleolar RNAs), gives rise to the mature 18S, 5.8S and 28S rRNAs (Figure 7.79). Ribosomal RNAs are also modified both on the ribose sugars and on the bases. Interestingly, methylation of ribose sugars is the major modification in rRNA. The modified base pseudouridine is also common in rRNA. Other modifications include base methylation, and acetylation. These modifications are thought to be important in modulating ribosome function.

Information Processing: RNA Processing

767

YouTube Lectures

by Kevin

HERE & HERE

768

Figure 7.68 - 5’ capping of eukaryotic mRNAs

Wikipedia

Figure 7.67 - Steps in processing of pre-mRNA

769

Figure 7.69 - Removal of introns from the primary transcript

Interactive Learning

Module

HERE

770

Figure 7.70 - Splicing of introns

Wikipedia

771

Figure 7.71 - Assembly of the spliceosome complex

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

772

Figure 7.72 - Alternative splicing leads to different forms of a protein from the same gene sequence

Figure 7.73 - Alternative poly-adenylation sites for a gene

773

Figure 7.74 - Structure of a mature eukaryotic mRNA

Interactive Learning

Module

HERE

774

Figure 7.76 - Structure of the RNA component of ribonuclease P

Figure 7.75 - Template guided - one mechanism of RNA editing

775

Figure 7.78 - Sequence of a mature tRNA

Wikipedia

Figure 7.77 - Synthesis of pseudouridine from uridine

Wikipedia

776

Figure 7.79 - Processing of ribosomal RNA

YouTube Lectures

by Kevin

HERE & HERE

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

778

The Codon Song

To the tune of "When I'm Sixty Four"

Metabolic Melodies Website HERE

Building of proteins, you oughta know
Needs amino A's
Peptide bond catalysis in ribosomes
Triplet bases, three letter codes

Mixing and matching nucleotides
Who is keeping score?
Here is the low down
If you count codons
You'll get sixty four

Got - to - line - up - right
16-S R-N-A and
Shine Dalgarno site

You can make peptides, every size
With the proper code
Start codons positioned
In the P site place
Initiator t-RNAs

UGA stops and AUGs go

Who could ask for more?
You know the low down
Count up the codons
There are sixty four

Recording by Tim Karplus

Lyrics by Kevin Ahern
Recording by Tim Karplus Lyrics by Kevin Ahern

This page titled 7.6: RNA Processing is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.7: Translation

Source: BiochemFFA_7_6.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Translation is the process by which information in mRNAs is used to direct the synthesis of proteins. As you have learned in introductory biology, in eukaryotic cells, this process is carried out in the cytoplasm of the cell, by large RNA-protein machines called ribosomes. Ribosomes contain ribosomal RNA (rRNA) and proteins. The proteins and rRNA are organized into two subunits, a large and a small. Ribosomes function by binding to mRNAs and holding them in a way that allows the amino acids encoded by the RNA to be joined sequentially to form a polypeptide. Transfer RNAs are the carriers of the appropriate amino acids to the ribosome.

The Genetic Code

We speak of genes (i.e., DNA) coding for proteins and the central dogma, which states that DNA makes RNA makes protein. What does this actually mean? A code can be thought of as a system for storing or communicating information. A familiar example is the use of letters to represent the names of airports (e.g., PDX for Portland, Oregon and ORD for Chicago's O'Hare). When a tag on your luggage shows PDX as the destination, it conveys information that your bag should be sent to Portland, Oregon. To function well, such a set-up must have unique identifiers for each airport and people who can decode the identifiers correctly. That is, PDX must stand only for Portland, Oregon and no other airport. Also, luggage handlers must be able to correctly recognize what PDX stands for, so that your luggage doesn't land in Phoenix, instead.

How does this relate to genes and the proteins they encode?

Genes are first transcribed into mRNA, as we have already discussed. The sequence of an mRNA, copied from a gene, directly specifies the sequence of amino acids in the protein it encodes. Each amino acid in the protein is specified by a sequence of 3 bases called a codon in the mRNA (Figure 7.81). For example, the amino acid tryptophan is encoded by the sequence UGG on an mRNA. All of the twenty amino acids used to build proteins have, likewise, 3-base sequences that encode them.

Degeneracy

Given that there are 4 bases in RNA, the number of different 3-base combinations that are possible is 43, or 64. There are, however, only 20 amino acids that are used in building proteins in cells. This discrepancy in the number of possible codons and the actual number of amino acids they specify is explained by the fact that the same amino acid may be specified by more than one codon. In fact, with the exception of the amino acids methionine and tryptophan, all the other amino acids are encoded by multiple codons. Codons for the same amino acid are often related, with the first two bases the same and the third being variable. An example would be the codons for alanine: GCU, GCA, GCC and GCG all stand for alanine. This sort of redundancy in the genetic code is termed degeneracy.

Stop and start codons

Three of the 64 codons are what are known as termination or stop codons, and as their name suggests, indicate the end of a protein coding sequence. The codon for methionine, AUG, is used as the initiation or start codon for the majority of proteins. This ingenious system is used to direct the assembly of a protein in the same way that you might string together colored beads in a particular order using instructions that used symbols like UGG for a red bead, followed by UUU for a green bead, CAC for yellow, and so on, till you came to UGA, indicating that you should stop stringing beads.

Translating the code

While the ribosomes are literally the factories that join amino acids together using the instructions in mRNAs, another class of RNA molecules, the transfer RNAs (tRNAs) are also needed for translation (Figure 7.83 and Interactive 7.1). Transfer RNAs are small RNA molecules, about 75-90 nucleotides long, that function to 'interpret' the instructions in the mRNA during protein synthesis. Transfer RNAs are extensively modified post-transcriptionally and contain a large number of unusual bases. The sequences of tRNAs have several self complementary regions, where the single-stranded tRNA folds on itself and base-pairs to

form what is sometimes described as a clover leaf structure.

This structure is crucial to the function of the tRNA, providing both the sites for attachment of the appropriate amino acid and for recognition of codons in the mRNA. In terms of the bead analogy above, someone or something has to be able to bring a red bead in when the instructions indicate UGG, and a green bead when the instructions say UUU. This, then, is the function of the tRNAs. They must be able to bring the amino acid corresponding to the instructions to the ribosome.

t-RNA specificity

A given transfer RNA is specific for a particular amino acid. It is linked covalently at its 3' end to the appropriate amino acid by an enzyme called aminoacyl tRNA synthetase. For example, there is a transfer RNA that is specific to the amino acid tryptophan, and a corresponding aminoacyl tRNA synthetase, called a tryptophanyl tRNA synthetase, that can attach the tryptophan specifically to this tRNA. Likewise, there is an aminoacyl tRNA synthetase specific for each amino acid. A tRNA with an amino acid attached to it is said to be charged (Figure 7.84). A pool of charged tRNAs is necessary to carry out protein synthesis. How do these tRNAs, carrying specific amino acids assist the ribosome in stringing together the correct amino acids, as specified by the sequence of the mRNA?

Codon recognition

As we already know, the amino acid sequence of the protein is determined by the order of the codons in the mRNA. We also have charged tRNAs carrying the various amino acids present.

How are the amino acids attached to each other in the order indicated by the base sequence of the mRNA? This requires recognition of the codons on the mRNA by the appropriate charged tRNAs. The amino acid tryptophan, as we noted, is specified by the codon UGG in the mRNA. This codon must be recognized by a tRNA charged with tryptophan. Every tRNA has a a sequence of 3 bases, the anticodon, that is complementary to the codon for the amino acid it is carrying. When the tRNA encounters the codon for its amino acid on the messenger RNA, the anticodon will base-pair with the codon. For the tryptophan tRNA this is what it would look like:

Sequence of tryptophan codon in mRNA:

5’ UGG 3’

Anticodon sequence in tryptophan tRNA:

5’ CCA 3’

Note that the sequences are both written, by convention, in the 5’ to 3’ direction. To base pair, though, they must be oriented in opposite directions (anti-parallel). The codon-anticodon basepair in the antiparallel orientation then would be:

5’ UGG 3’
3’ ACC 5’

The base-pairing of the anticodon on a charged tRNA with the codon on the mRNA is what brings the correct amino acids in to the ribosome to be added on to the growing protein chain (Figure 7.85).

Making a Polypeptide

With an idea of the various components necessary for translation and how they work, we can now take a look at the process of protein synthesis. The main steps in the process are similar in prokaryotes and eukaryotes. As we already noted, ribosomes bind to mRNAs and facilitate the interaction between the codons in the mRNA and the anticodons on charged tRNAs.

Ribosomes have two sites (P-site and A-site) for binding and positioning charged tRNAs so each can form base pairs between their anticodon and a codon from the mRNA. The start codon (AUG) is positioned to base pair with the tRNA in the P-site (peptidyl site). Next, the charged tRNA complementary to the codon adjacent to the start codon binds and occupies the A-site (aminoacyl site) in the ribosome (Figure 7.86).

At this point, the ribosome joins the amino acids carried on each tRNA by making a peptide bond. The bond between the amino acid and the tRNA in the P-site is broken and the dipeptide is joined to the tRNA on the A-site.

The initiator tRNA without its amino acid is then released, moving into a site known as the Exit or E-site, while the second tRNA carrying the dipeptide (and the codon it is base paired to) moves into the P-site. The A-site now is ready with a new codon for the next incoming charged tRNA. This process is repeated, with the ribosome moving on the mRNA one codon at a time, until the stop codon reaches the A-site. At this point, a release factor binds at the A-site, and helps to free the completed polypeptide from the ribosome. The ribosome then dissociates into the small and large subunits, once more.

Three steps

Having considered the steps of translation in broader terms, we can now look at them in greater detail. We will consider the three steps of translation (bel0w) individually.

Initiation (binding of the ribosomal subunits to the transcript and initiator tRNA)
Elongation (repeated addition of amino acids to the growing polypeptide, based on the sequence of the mRNA - Figure 7.87)
Termination (release of the completed polypeptide and dissociation of the ribosome into its subunits).

We already know that processed mRNAs are sent from the nucleus to the cytoplasm in eukaryotic cells, while in prokaryotic cells, transcription and translation occur in a single cellular compartment. The small and large subunits of ribosomes, each composed of characteristic rRNAs and proteins, are found in the cytoplasm and assemble on mRNAs to form complete ribosomes that carry out translation. Both prokaryotic and eukaryotic ribosomal subunits are made up of one or more major rRNAs together with a large number of ribosomal proteins. The small subunits of prokaryotic cells are called the 30S ribosomal subunits, while their counterparts in eukaryotes are the 40S subunits. The large ribosomal subunits in prokaryotes are the 50S subunits, while those in eukaryotic cells are 60S. These differences reflect the larger mass of eukaryotic ribosomes. The rRNA components of ribosomes are important for the recognition of the 5’ end of the mRNA, and also play a catalytic role in the formation of peptide bonds.

Initiation

Messenger RNAs have non-coding sequences both at their 5' and 3' ends, with the actual protein-coding region sandwiched in between these untranslated regions (called the 5' UTR and 3' UTR, respectively). The ribosome must be able to recognize the 5' end of the mRNA and bind to it, then determine where the start codon is located. It is important to note that both in prokaryotes and eukaryotes, ribosomes assemble at the 5’ end of the transcript by the stepwise binding of the small and large subunits. The small subunit first binds to the mRNA at specific sequences in the 5’ UTR. The large subunit then binds to the complex of the mRNA and small subunit, to form the complete ribosome.

Initiator tRNA

Initiation also requires the binding of the first tRNA to the ribosome. As we have noted earlier, the initiation, or start codon is usually AUG, which codes for the amino acid methionine. Thus, the initiator tRNA is one that carries methionine and is designated as tRNAmet or methionyl tRNAmet. In bacteria, the methionine on the initiator tRNA is modified by the addition of a formyl group, and is designated tRNAfmet. The initiator tRNA carrying methionine to the AUG is different from the tRNAs that carry methionine intended for other positions in proteins. As such, the initiator tRNA is sometimes referred to as tRNAimet.

Prokaryotic initiation

In prokaryotes, the 5’ end of the mRNA is the only free end available, as transcription is tightly coupled to translation and the entire mRNA is not transcribed before translation begins. Nevertheless, the ribosome must be correctly positioned at the 5’ end of the messenger RNA in order to initiate translation. How does the ribosome “know” exactly where to bind in the 5’UTR of the mRNA?

Shine-Dalgarno sequence

Examination of the sequences upstream of the start codon in prokaryotic mRNAs reveals that there is a short purine-rich sequence ahead of the start codon that is crucial to recognition and binding by the small ribosomal subunit (Figure 7.89). This sequence, called the Shine-Dalgarno sequence, is complementary to a stretch of pyrimidines at the 3’ end of the 16S rRNA component of the small ribosomal subunit (Figure 7.90). Base-pairing between these complementary sequences positions the

small ribosomal subunit at the right spot on the mRNA, with the AUG start codon at the P-site.

Initiation factors

The binding of the small ribosomal subunit to the mRNA requires the assistance of three protein factors called Initiation Factors 1, 2 and 3 (IF1, IF2, IF3). These proteins, which are associated with the small ribosomal subunit, are necessary for its binding to mRNA, but dissociate from it when the 50S ribosomal subunit binds. Of these proteins, IF3 is important for the binding of the small subunit to the mRNA, while IF2 is involved in bringing the initiator tRNA to the partial P-site of the small ribosomal subunit. IF1 occupies the A-site, preventing the binding of the initiator tRNA at that site.

Once the small ribosomal subunit is bound to the mRNA and the initiator tRNA is positioned at the P-site, the large ribosomal subunit is recruited and the initiation complex is formed. Binding of the 50S ribosomal subunit is accompanied by the dissociation of all three initiation factors. The removal of IF1 from the A-site on the ribosome frees up the site for the binding of the charged tRNA corresponding to the second codon (Figure 7.91).

Eukaryotic initiation

In eukaryotes, initiation follows a similar pattern, although the order of events and the specific initiation factors are different.

Eukaryotes have a large number of IFs that are known as eIFs (eukaryotic initiation factors). These initiation factors are involved in the binding of the initiator tRNA to the small subunit, as well as in association of the small subunit with mRNA and subsequent attachment of the large subunit.

Ribosome assembly

The assembly of the translation machinery in eukaryotes begins with the binding of the initiator tRNA to the 40S (small) subunit. This step requires the assistance of eIF2 and eIF3. Next the small subunit with the initiator tRNA binds to the 7-methyl G cap on the 5'end of the mRNA. The 40S subunit then moves along the mRNA, scanning for a a start codon. Binding of the ribosomal subunit to the mRNA is dependent not just on finding an AUG, but on the sequences surrounding the codon.

Kozak sequences

Specific sequences surrounding the AUG, called Kozak sequences for the scientist who defined them, have been shown to be necessary for the binding of the 40S subunit, with the bases at -4 and +1 relative to the AUG being especially important (Figure 7.92). Once the small subunit is properly positioned, the large ribosomal subunit (60S) binds, forming the initiation complex.

Elongation

After the ribosome is assembled with the initiator tRNA positioned at the AUG in the P-site, the addition of further amino acids can begin. In both prokaryotes and eukaryotes, the elongation of the polypeptide chain requires the assistance of elongation factors. In bacteria, the binding of the second charged tRNA at the A-site requires the elongation factor EF-Tu complexed with GTP (Figure 7.93). When the charged tRNA has been loaded at the A-site, EF-Tu hydrolyzes the GTP to GDP and dissociates from the ribosome. The free EF-Tu can then work with another charged tRNA to help position it at the A-site (Figure 7.94), after exchanging its GDP for a new GTP.

The corresponding step in eukaryotic cells is dependent on the elongation factor eEF1α.GTP. Once both P-site and A-site are occupied, the methionine carried by the tRNA in the P-site is joined to the amino acid carried by tRNA in the A-site, forming a peptide bond.

The reaction that joins the amino acids occurs in the ribosomal peptidyl transferase center, which is part of the large ribosomal subunit (Figure 7.95).

Ribozyme

Interestingly, there is strong evidence that this reaction is catalyzed by rRNA components of the large subunit, making the formation of peptide bonds an example of the activity of RNA enzymes, or ribozymes. The result of the peptidyl transferase activity is that the tRNA in the A-site now has two amino acids attached to it, while the tRNA at the P-site has none. This “empty” or deacylated tRNA is moved to the E-site on the ribosome, from which it can exit. The tRNA in the A-site, then moves to occupy the vacated P-site, leaving the A-site open for the next incoming charged tRNA. Yet another elongation factor, EF-G complexed with GTP, is required for the translocation of the ribosome along the mRNA in bacteria, while in eukaryotes, this role is played

by eEF2.GTP. Repeated cycles of these steps result in the elongation of the polypeptide by one amino acid per cycle, until a termination, or stop codon is in the A-site.

Termination

When a stop codon is in the A-site, proteins called release factors (RFs) are needed to recognize the stop codon and cleave and release the newly made polypeptide. In bacteria, RF1 is a release factor that can recognize the stop codon UAG, while RF2 recognizes UGA. Both RF1 and RF2 can recognize UAA. A third release factor, RF3, works with RF1 and RF2 to hydrolyze the linkage between the polypeptide and the final tRNA, to release the newly synthesized protein. This is followed by the dissociation of the ribosomal subunits from the mRNA, ending the process of translation .

Polypeptide processing

What happens to the newly synthesized polypeptide after it is released from the ribosome? As you know, functional proteins are not simply strings of amino acids. The polypeptide must fold properly in order to perform its function in the cell. It may also undergo a variety of modifications such as the addition of phosphate groups, sugars, lipids, etc. Some proteins are produced as inactive precursors that must be cleaved by proteases to be functional.

An additional challenge in eukaryotic cells is the presence of internal, membrane-bounded compartments. Each compartment contains different proteins with different functions. But the vast majority of proteins in eukaryotic cells are made by ribosomes, free or membrane-bound, in the cytoplasm of the cell (the exceptions are a handful of proteins made within mitochondria and chloroplasts).

Delivery

Each of the thousands of proteins made in the cytoplasm must, therefore, be delivered to the appropriate cellular compartment in which it functions. Some proteins are delivered to their destinations in an unfolded state, and are folded within the compartment in which they function. Others are fully folded and may be post-translationally modified before they are sent to their cellular (or extracellular) destinations.

Some proteins are delivered as they are being synthesized (co-translationally - see Movie 7.3) while others are sorted to their compartments post-translationally. But, with the exception of cytosolic proteins, all proteins must cross membrane barriers, through membrane channels or other "gates", or by transport within membrane vesicles that fuse with the membrane of the target organelle to deliver their contents.

Folding and post-translational modifications

Proper folding of a protein into its 3-dimensional conformation is necessary for it to function effectively. As described in an earlier chapter (HERE), the folding of a protein is largely influenced by hydrophobic interactions that result in folding of the protein in such a way as to position hydrophobic residues in the interior, or core, of the protein, away from the aqueous environment of the cell.

Proper folding may also involve the interaction of regions of the polypeptide that are distant from each other, so that portions of the N-terminal region of the polypeptide may be in close proximity to parts of the C-terminus of the final folded molecule.

As a polypeptide emerges from the ribosome, however, the N-terminal region of the polypeptide may begin to fold on itself, with adjacent parts of the chain interacting in inappropriate ways, before the entire protein has been made. This can result in misfolding of the protein and consequent malfunction. To prevent misfolding, cells have protein chaperones, whose function is to bind to and shield regions of polypeptides as they emerge from the ribosome, and keep them from improperly interacting with one another or with other proteins in the vicinity, until they can fold into their correct final shape (Figure 7.98). In addition, there are classes of chaperones that are able sequester proteins in such a way as to permit unfolding and refolding of misfolded polypeptides. These proteins ensure that the vast majority of proteins in cells are folded into their correct, functional 3-dimensional shapes.

Protein sorting

The process by which proteins are identified as belonging to a particular compartment and then correctly delivered to that destination is known as protein sorting. How does a cell know where a particular protein should be sent?

Proteins have "address labels" or sorting signals that indicate which cellular compartment they are destined for. Characteristic

sorting signals are found on proteins that are sent to the nucleus, the ER (Figure 7.99), the mitochondria, etc.

Signal sequences

What do these sorting signals look like?
Most sorting signals (also called signal sequences) are short stretches of amino acid sequence (that is, they are part of the amino acid sequence of the protein). Different cellular compartments have different "address labels".

Signal sequences may be found at the N-terminal or C-terminal region of proteins, or they may be within the amino acid sequence of the proteins. The location of the signal sequence for any given protein is fixed, however. Signal sequences for proteins to be delivered to the endoplasmic reticulum (ER) are found at the N-terminus of the protein. Mitochondrial and chloroplast proteins encoded by nuclear genes also have N-terminal signal sequences. Signal sequences for nuclear proteins, by contrast, are internal to the polypeptide, and may consist of one or more stretches of amino acids that will be displayed on the surface of these proteins once they are folded.

Free and membrane-bound ribosomes

Proteins are synthesized by ribosomes in the cytoplasm or by those that associate with membranes temporarily (membrane-bound ribosomes). The free ribosomes make proteins that are destined for the nucleus, as well as those going to chloroplasts, mitochondria and peroxisomes. Nuclear proteins are delivered in their folded state, while chloroplast and mitochondrial proteins are threaded through translocation channels in the membranes of these organelles, to be folded at their destination.

Proteins that are destined for the ER, the Golgi apparatus, lysosomes as well as those that are to be secreted from the cell are first delivered to the ER by ribosomes that associate with the membrane of the rough ER and synthesize the protein directly into the ER. Proteins delivered by this manner into the lumen of the ER undergo folding and modification in the ER. All proteins delivered to the ER, regardless of their final destination, have an N-terminal ER signal sequence of 15-30 amino acids.

Protein delivery into the endoplasmic reticulum

The N-terminal part of a protein is the first part of a nascent polypeptide that emerges from the ribosome (Figure 7.100). The sequence of amino acids in this region, if it is an ER signal, will be recognized and bound by a ribonucleoprotein complex called the Signal Recognition Particle (SRP). Binding of the SRP to the N-terminal signal sequence causes translation to pause. The SRP, in turn, is bound by an SRP receptor in the ER membrane (Movie 7.3 & Figure 7.101), effectively anchoring the ribosome to the membrane.

The location of SRP receptors near membrane channels in the ER positions the ribosome over a translocation channel. Once the ribosome is docked over the channel, the SRP releases the signal sequence, which is threaded through the channel, with its hydrophobic residues interacting with the hydrophobic interior of the membrane. Translation resumes at this point and the rest of the protein is delivered into the lumen of the ER as it is made. The ribosome remains associated with the ER membrane till translation is completed, at which point it dissociates. The signal sequence, which is no longer needed once the protein has been delivered, is cleaved off by a membrane associated signal peptidase, releasing the completed protein into the ER lumen.

While soluble proteins are delivered into the ER, integral membrane proteins do not pass all the way through, but, instead, are anchored in the membrane of the ER by hydrophobic stop transfer sequences.

Folding and modification

Proteins in the lumen of the ER are folded with the help of numerous chaperones resident in the endoplasmic reticulum. The environment within the ER lumen is also more oxidizing than the cytosol, and permits the formation of disulfide bonds to stabilize the folded proteins. Protein disulfide isomerase, an enzyme active in the ER lumen both helps to make disulfide bonds and removes bonds that were incorrectly made during the folding process. In addition, proteins in the ER undergo modifications such as glycosylation and addition of glycolipids. Multimeric proteins are also assembled from their subunits in the ER.

Proteins that have been correctly folded and modified are transported from the ER, in membrane vesicles, to their final destinations. Improperly folded proteins are recognized by a surveillance mechanism in the ER and are sent back to the cytoplasm to be degraded in proteasomes.

Information Processing: Translation

779

YouTube Lectures

by Kevin

HERE & HERE

780

Figure 7.81 - The standard genetic code

Wikipedia

Figure 7.80 - The central dogma in a bacterial cell

Wikipedia

781

Figure 7.82 - Coding in DNA, transcribed to RNA, translated to protein

782

Figure 7.83 - tRNA - 3D projection (left) and 2D projection (inset)

Wikipedia

783

Interactive 7.1 - Phenylalanyl-tRNA

PDB
Interactive 7.1 - Phenylalanyl-tRNA PDB

Figure 7.84 - Charging of a tRNA by aminoacyl tRNA synthetase

Wikipedia

784

Figure 7.85 - Codons in mRNA pair with anticodons on tRNA to bring the appropriate amino acid to the ribosome for polypeptide assembly

YouTube Lectures

by Kevin

HERE & HERE

785

Figure 7.86 - The A, P, and E sites in a ribosome

Image by Martha Baker

Figure 7.87 - Overview of elongation

Wikipedia

Interactive Learning

Module

HERE

786

Movie 7.1 - 30S ribosomal subunit

Wikipedia
Movie 7.1 - 30S ribosomal subunit Wikipedia

Table 7.1 - Location and function of rRNAs.

rRNA Name

Prokaryotes

Eukaryotes

Function

5S

Large Subunit

Large Subunit

tRNA binding?

5.8S

Large Subunit

Translocation?

16S

Small Subunit

mRNA alignment

18S

Large Subunit

mRNA alignment

23S

Large Subunit

Peptide bond formation

28S

Large Subunit

Peptide bond formation

787

Figure 7.89 - Conserved sequences adjacent to start codons for various bacterial genes

Image by Martha Baker

Figure 7.88 - Structure of 5S rRNA

788

Figure 7.90 - Base pairing between the Shine-Dalgarno sequence in the mRNA and the 16S rRNA

Image by Martha Baker

Movie 7.2 - Large ribosomal subunit

Wikipedia
Movie 7.2 - Large ribosomal subunit Wikipedia

789

YouTube Lectures

by Kevin

HERE & HERE

Figure 7.91 - Initiation - assembly of the ribosomal translation complex

Image by Martha Baker

790

Figure 7.92 - Kozak sequence plot showing relative abundance of bases surrounding the AUG (ATG) start codon of human genes

Figure 7.93 - EF-Tu (blue) bound to tRNA (red) and GTP (yellow)

791

Figure 7.94 - The process of elongation

Image by Martha Baker

792

Figure 7.95 - 50S ribosomal subunit. RNA in brown. Protein in blue. Peptidyl transferase site in red.

Wikipedia

Interactive Learning

Module

HERE

793

Figure 7.96 - The process of translation

Wikipedia

Movie 7.3 Translation of a protein secreted into the endoplasmic reticulum. Small subunit in yellow. Large subunit in green. tRNAs in blue.

Wikipedia
Movie 7.3 Translation of a protein secreted into the endoplasmic reticulum. Small subunit in yellow. Large subunit in green. tRNAs in blue. Wikipedia

794

Figure 7.97 - Another perspective of translation. The 3' end of the mRNA is on the left and the ribosome is moving from right to left

795

Figure 7.98 - Action of chaperone to facilitate proper folding of a protein (orange)

Image by Aleia Kim

YouTube Lectures

by Kevin

HERE & HERE

796

Figure 7.99 - Rough (ribosome bound) and smooth endoplasmic reticulum

Wikipedia

797

Figure 7.100 - N-terminal signal sequence (green) emerging from the ribosome.

798

Figure 7.101 - Translation of a protein into the endoplasmic reticulum

Image by Aleia Kim

799

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

801

Translation

To the tune of "Maria" (from "West Side Story")

Metabolic Melodies Website HERE

Translation
The most intricate thing I ever saw

From five prime to three prime, translation, translation
The final step that we know about the central dog-ma

Amino, carboxyl, translation, translation. . . .
Translation, translation, translation . .

Translation!

I just learned the steps of translation
And all the things they say
About tRNA
Are true

Translation!
To form peptide bonds in translation

The ribosomal cleft
Must bind to an E-F
tee-you!

Translation!
A-U-G binds the f-met's cargo
16S lines up Shine and Dalgarno
Translation
I'll never stop needing translation

The most intricate thing I ever saw
Translationnnnnnnnnnnnnnnnn

Recording by Tim Karplus

Lyrics by Kevin Ahern
Recording by Tim Karplus Lyrics by Kevin Ahern

802

Good Protein Synthesis

To the tune of “Good King Wenceslaus”

Metabolic Melodies Website HERE

Amino acids cannot join

By themselves together

They require ribosomes

To create the tether

All the protein chains get made

‘Cording to instruction

Carried by m-R-N-A

In peptide bond construction

Small subunit starts it all

With initiation

Pairing up two RNAs

At the docking station

Shine Dalgarno’s complement

In the 16 esses

Lines the A-U-G up so

Synthesis commences

Elongation happens in

Ribosomic insides

Where rRNA creates

Bonds for polypeptides

These depart the ribosome

Passing right straight through it

In the tiny channels there

Of the large subunit

Finally when the sequence of

One of the stop codons

Parks itself in the A site

Synthesis can’t go on

P-site RNA lets go

Of what it was holding

So the polypeptide can

Get on with its folding

Recording by David Simmons

Lyrics by Kevin Ahern
Recording by David Simmons Lyrics by Kevin Ahern

This page titled 7.7: Translation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.8: Gene Expression

Source: BiochemFFA_7_7.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

The processes of transcription and translation described so far tell us what steps are involved in the copying of information from a gene (DNA) into RNA and the synthesis of a protein directed by the sequence of the transcript (Figure 7.102). These steps are required for gene expression, the process by which information in DNA directs the production of the proteins needed by the cell.

But what determines whether a gene is expressed at a given time? Cells do not, as we know, express all of their genes all of the time. Some genes are expressed in particular cell types but not others, while others may be expressed at specific stages of development. Cells must also be able alter their patterns of gene expression in response to internal and external cues, controlling the production of proteins as needed, to meet their changing needs. Regulating gene expression is, therefore, crucial. Given that there are multiple steps involved in gene expression, there are several different points at which the process could be regulated. Not surprisingly, many regulatory mechanisms are known, each acting at a different stage in the path from DNA to protein.

Regulation of Transcription

The first step in gene expression is transcription, so regulation of transcription is an obvious way to affect whether a gene is expressed and to what extent.

What are the molecular switches that turn transcription on or off? Although there are additional factors that affect transcription, such as the accessibility of a gene to the transcriptional machinery, the basic mechanism by which transcription is regulated depends on highly specific interactions between transcription regulating proteins and regulatory sequences on DNA.

What are these regulatory sequences and what proteins bind them? In addition to the promoter sequences required for transcription initiation, genes have additional cis 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.

Transcriptional regulation in prokaryotes

Let us first consider some examples from 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. Groups of genes that are coordinately regulated by a single promoter are referred to as operons. The entire set of genes in an operon can be controlled through the action of DNA binding proteins that act as either repressors (preventing transcription of the genes) or activators (increasing transcription of the genes). The binding of these proteins to their DNA targets is allosterically controlled by the binding of specific small molecules that signal the state of the cell.

Induction of the lac operon

The lac operon is one such group of coordinately regulated genes that encode proteins needed for the uptake and breakdown of the sugar lactose. E.coli cells preferentially 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. The default state of the lac operon is OFF.

Removing a repressor

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 (Figure 7.104). Recall that the promoter is where the RNA polymerase must bind to begin transcription. The location on the DNA where the lac repressor is bound is called the operator (Figure 7.105). When

the repressor is bound at this position, it physically blocks the RNA polymerase from transcribing the genes, just as a vehicle blocking your driveway would prevent you from pulling out. Obviously, if you want to leave, the vehicle that is blocking your path must be removed. Likewise, in order for transcription to occur, the repressor must be removed from the operator to clear the path for RNA polymerase (Figure 7.106).

How is the repressor removed? When the sugar lactose is present, a small amount of it is taken up by the cells and converted to an isomeric form, allolactose (Figure 7.107). Allolactose binds to the repressor, changing its conformation so that it no longer binds to the operator. When the repressor is no longer bound to the operator, the "road-block" in front of the RNA polymerase is removed, permitting the transcription of the genes of the lac operon

What makes this an especially effective control system is that the genes of the lac operon encode proteins that enable the break down of lactose. Turning on these genes requires lactose to be present. Once the lactose has been broken down, the lac 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.

Recruiting RNA polymerase

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 Catabolite Activator Protein (CAP - Figure 7.108). CAP (also sometimes called CBP or cAMP binding protein) binds to a site adjacent to the promoter and is necessary to recruit RNA polymerase to bind the lac promoter.

cAMP binding

CAP binds to its site only when glucose levels are low. Low glucose levels are linked to the activation of an enzyme, adenylate cyclase, that makes the molecule cyclic AMP (cAMP). The binding of cAMP to the CAP causes a conformational change in CAP that allows it to bind to the CAP-binding site. When CAP is bound at this site, it is able to recruit RNA polymerase to bind at the promoter, and begin transcription.

The combination of CAP binding and the lac repressor dissociating from the operator when lactose levels are high ensures transcription of the lac operon just when it is most needed. The binding of CAP may be thought of as a green light for the RNA polymerase, while the removal of lac repressor is like the lifting of a barricade in front of it. When both conditions are met, the RNA polymerase transcribes the downstream genes.

Control of the trp operon by repression

The lac operon we have just described is a set of genes that are expressed only under the specific conditions of glucose depletion and lactose availability. Other genes may be expressed unless a particular condition is met. For these genes, the default state is ON.

An example of this is the trp operon, which encodes enzymes necessary for the synthesis of the amino acid tryptophan. These genes are constitutively expressed (always on), except when tryptophan is available from the cell's surroundings, making its synthesis unnecessary. Under conditions where tryptophan is abundant in the environment, the trp genes can be turned off. This is achieved by a repressor protein that will bind to the operator only in the presence of tryptophan (Figure 7.110). Binding of tryptophan to the repressor causes binding of the repressor to the operator. Because it acts together with the repressor to turn off the trp genes, tryptophan is called a co-repressor.

Attenuation

Another mechanism that regulates the expression of the trp operon is attenuation. Attenuation is a process by which the expression of an operon is controlled by termination of transcription before the first gene of the operon (Figure 7.111).

In the trp operon, this functions as follows: Transcription begins some distance upstream of the first gene in the operon, producing what is termed a 5' leader sequence. This leader sequence contains an intrinsic terminator that can form a hairpin structure that stops transcription when high levels of tryptophan are available to the cells. It can also form a different structure that permits continued transcription of the genes in the operon when tryptophan levels are low. How does the level of tryptophan influence which of these two structures are formed?

Recall that the 5' end of the RNA is the first part of the transcript to be made and that in bacteria translation is linked to transcription, so the 5' end of the RNA begins to be translated before the entire transcript is made. It turns out that the 5'

leader sequence of the trp operon mRNA encodes a short peptide that contains two tryptophan codons. If there is plenty of tryptophan available, the leader sequence will be easily translated. Under these conditions, the leader sequence is able to form the termination hairpin, preventing the transcription of the downstream trp genes.

If, however, levels of tryptophan are low, then the ribosome stalls as it attempts to translate the leader sequence. Under these conditions, the leader sequence adopts a different conformation that permits continued transcription of the genes of the trp operon.

Riboswitches

Similar in concept to the attenuation of the trp operon described above, but not dependent on translation, is a control mechanism called a riboswitch (Figure 7.113). Riboswitches are typically found in the 5'UTR of messenger RNAs (i.e., they are part of the sequence of the RNA). These sequences can control transcription of the downstream genes based on the conformation they adopt. One conformation allows continued transcription, while the other terminates it. So, what determines which conformation they adopt?

Features

Riboswitches have two characteristic features that are important for their function. One is a region of the sequence called an aptamer, which folds into a three-dimensional shape that can bind a small effector molecule. The other is an adjacent region of the RNA, called the expression platform, that can fold into different conformations depending on whether or not the aptamer is bound to the effector.

An example of a riboswitch found in bacteria is the guanine riboswitch, which controls the expression of genes required for purine biosynthesis. The aptamer region of this riboswitch binds to the effector, guanine, when levels of the base are high. The binding of the guanine triggers a change in the folding of the downstream expression platform, causing it to adopt a conformation that terminates transcription of the genes needed for the synthesis of guanine. In the absence of guanine, the expression platform assumes a different conformation that allows transcription of the purine biosynthesis genes. Thus, levels of guanine can be sensed and the genes needed for its synthesis can be expressed as needed.

Regulation of transcription 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 transcription initiation complex) are necessary but not sufficient for high levels of transcription. Promoter-proximal DNA sequences like the CAAT box and GC box bind proteins that interact with the transcription initiation complex, influencing its formation (Figure 7.114).

Distant regulatory sequences

Additional regulatory sequences called enhancers and the proteins that bind to them are needed to achieve high levels of transcription. Enhancers are short DNA sequences that regulate the transcription of genes, but may be located at a distance from the gene they control (although they are on the same DNA molecule as the gene). Often enhancers are many kilobases away on the DNA, either upstream or downstream of the gene. 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 transcription?

Transcriptional activators

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 transcription initiation complex that is bound at a distant site by looping of the DNA (Figure 7.115). This allows the protein bound at the enhancer to make contact with the proteins in the basal transcription complex. The interaction of the activator with the transcription initiation complex may be direct, or it may be through a “middle-man”, a protein complex called mediator.

One effect of this interaction is to assist in recruiting proteins necessary for transcription, like the general transcription factors and RNA polymerase to the promoter, increasing the frequency and efficiency of formation of the transcription initiation complex. There is also evidence that at some promoters, following assembly of the transcription initiation complex, the RNA polymerase remains stalled at the promoter. In such cases, the interaction with the transcription initiation complex of an activator bound to an enhancer could play a role in facilitating the transition of the RNA polymerase to the elongation phase of transcription.

Chromatin remodeling proteins

Another mechanism by which activators bound at the enhancer can affect transcription is by recruiting to the promoter proteins that can modify the structure of that region of the chromosome. 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.

Silencers

In addition to enhancers, there are also negative regulatory sequences called silencers. Such regulatory sequences bind to transcriptional repressor proteins. Like the transcriptional activators, these repressors work by interacting with the transcription initiation complex. In the case of repressors, the effect they have on the transcription initiation complex is to reduce transcription.

DNA binding 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 transcription initiation complex (Figure 7.118). The DNA binding domain is the part of the protein that confers specificity for determining which gene(s) will be activated or repressed. The activation domain is the part of the protein that stimulates or represses transcription. The DNA binding domains of transcriptional activators form characteristic structures that recognize their target DNA sequences by making contacts with bases, usually in the major groove of the DNA helix. It is possible to engineer hybrid transcription factors that combine the DNA binding domain of one activator with the activation domain of another. Such proteins retain the specificity dictated by the DNA binding domain. Truncated transcription factors can also be generated that have their DNA binding domain but lack the activation domain. Such transcription factors can be useful tools in studying transcriptional regulation because their DNA binding domains can compete with the endogenous transcription factors for regulatory binding sites without increasing transcription from the target promoters.

Multiple factors

The description above may suggest that each gene in eukaryotes is controlled by the binding of a single transcriptional activator or repressor to a particular enhancer or silencer site. However, it turns out that the transcription of any given gene may be simultaneously regulated by a combination of proteins, both activators and repressors, bound at multiple regulatory sites on the DNA, all of which interact with the transcription initiation complex. The combinatorial nature of such regulation provides great versatility, with different combinations of regulatory elements and proteins working together in response to a wide variety of conditions and signals.

The mechanisms described so far have focused on the sequence elements in DNA that regulate transcription through the activator and repressor proteins bound to them. Following transcription, alternative splicing (see HERE) and editing of the transcripts can also modify the proteins that are produced by the cell. We will now examine some of the other ways in which gene expression is modulated in cells.

First, we will consider some so-called epigenetic mechanisms that affect gene expression. The term epigenetics derives from epi (above, or on top of) and genetic (of genes) and refers to the fact that these mechanisms act in addition to, or overlaid on, the information in the gene sequences. Two such epigenetic mechanisms are the covalent modifications of histones in chromatin and the methylation of DNA sequences.

Histone modification

As noted earlier, transcription in eukaryotes is complicated by the fact that the DNA is packaged with histones to make chromatin. This means that for a gene to be transcribed, the relevant regions of the chromatin must be opened up to allow access to the RNA polymerase and transcription factors. This provides another potential point of control of gene expression. Chromatin remodeling factors, mentioned earlier, assist in reorganizing the nucleosome structure at regions that need to be made accessible.

But what determines that a given region of the chromatin will be acted upon by the remodeling complexes? Transcriptional activator proteins bound at enhancers, sometimes work by recruiting histone modifying enzymes to the promoter region. An example of such a modifying enzyme is histone acetyl transferase (HAT) that works to acetylate specific amino acid residues in the tails of the histones forming the nucleosome core (Figures 7.119 & 7.120). Acetylation of histones is thought to be responsible for loosening the interaction between histones and the DNA in nucleosomes and helps to make the DNA more readily accessible for transcription. The opposite effect may be achieved if the enzymes recruited are histone deacetylases (HDAC) which remove acetyl groups from the tails of the histones in the nucleosome, and lead to tighter packing of the chromatin.

Writers, readers and erasers

In addition to the histone acetyl transferases and the deacetylases, other enzymes may add or remove methyl groups, phosphate groups, and other chemical moieties to specific amino acid side chains on the histone tails. The patterns of these covalent modifications, sometimes called the histone code, are established by the so-called "writers", or enzymes, such as histone methyltransferases, that add the chemical groups on to the histone tails. Yet other enzymes, like the histone demethylases, may act as "erasers," removing the chemical groups added by the "writers." The histone code is interpreted by "readers," proteins that bind to specific combinations of the modifications and assist in either silencing the expression of genes in the vicinity or making the region more transcriptionally active.

DNA methylation

Gene expression can also be regulated by methylation of the other component of chromatin - DNA. Enzymes called DNA methyltransferases (DNMTs) catalyze the covalent addition of a methyl group to C5 of cytosines in DNA. Patterns of cytosine methylation vary in different organisms, with methylation concentrated in some parts of the genome in some groups and scattered throughout the genome in others. In vertebrates, the cytosines that are methylated are generally next to a guanine (the CG dinucleotide is commonly abbreviated as CpG). Methylation of DNA seems to correlate with gene silencing while demethylation is associated with increased transcription (Figure 7.121).

How does methylation of the DNA at CpG sites regulate gene expression? Although the extent of DNA methylation near promoters has been observed to correlate with gene silencing, it is not clear how exactly methylation brings about this effect. It has been suggested that methylation could block the binding of proteins necessary for transcription. Methylation at enhancer sites might also prevent the binding of transcriptional activators to them.

Another interesting observation is that certain proteins that bind to methylated CpG sites also seem to interact with histone deacetylases. As noted above, histones deacetylases remove acetyl groups from histones, and promote tighter packing of chromatin and transcriptional silencing. Thus, methylation on DNA likely works in combination with histone modification to affect gene expression.

Regulatory RNAs

One of the most unexpected discoveries in the past few decades has been the role that RNAs play in regulating gene expression. The classic view that RNA either encoded proteins (mRNA) or assisted in their synthesis (rRNA and tRNA) is now known to be a vast underestimate of the various ways in which RNAs function in gene expression. It is now clear that regulatory RNAs have widespread and significant effects on gene expression, a realization that has revolutionized our understanding of gene regulation.

What are some of the ways in which regulatory RNAs function to modulate the expression of genes?

Small regulatory RNAs

MicroRNAs (miRNAs) and Short Interfering RNAs (siRNAs) are small, non-coding RNAs that act at the post-transcriptional level to regulate gene expression (Figure 7.123 & 7.124). These RNAs appear to silence genes by base-pairing with target mRNAs and marking them for degradation, or by blocking their translation. The functional forms of both miRNAs and siRNAs are from 20-30 nucleotides long and are derived by processing from longer primary transcripts. Mature miRNAs and siRNAs work in association with a class of proteins called Argonaute proteins to form a gene silencing complex.

MicroRNAs are transcribed from specific genes by RNA polymerase II. The primary transcript, known as a pri-miRNA folds on itself to form double-stranded hairpin structures that are cleaved by an RNase in the nucleus called Drosha. The products of Drosha cleavage, double-stranded RNAs of roughly 60-70 nucleotides known as pre-miRNAs, are exported to the cytoplasm, where they are further processed into the small 20-30 nucleotide lengths of mature double-stranded miRNAs by an enzyme known as Dicer. The RNA duplexes of miRNAs are not perfectly matched, and have loops and mismatches (Figure 7.124).

siRNAs also derive from double-stranded RNAs, but these may arise from either endogenous or exogenous sources (such as viruses). These double-stranded RNAs are processed in the cytoplasm by the same enzyme, Dicer, that generates the mature miRNAs, to produce the small, 20-30 nucleotide double-stranded RNAs.

In contrast to miRNAs, the mature siRNAs are perfectly base-paired along their lengths.

RISC assembly

Both miRNAs and siRNAs then are assembled with Argonaute proteins to form a silencing complex called RISC (RNA-induced silencing complex). Recall that both miRNAs and siRNAs are, at this point double-stranded. One strand of the RNA is referred to as the guide RNA, while the other is called the passenger RNA.

During the process of loading the RNA onto the Argonaute protein, the guide strand of the RNA remains associated with the protein, while the passenger strand is removed. The guide RNA associated with the Argonaute protein is the functional gene silencing complex (Figure 7.125).

Sequence specific base-pairing of the guide RNA with an mRNA leads to either the degradation of the mRNA by the Argonaute protein (in the case of the siRNAs) or in suppression of translation of the mRNA (for miRNAs). The extent to which these processes play a role in regulating gene expression is impressive. The expression of at least a third of all human genes has already been shown to be modulated by miRNAs, demonstrating clearly that these RNAs play a major role in gene regulation.

Long noncoding RNAs

Long noncoding RNAs (lncRNAs) are RNAs of greater than 200 nucleotides that do not code for proteins. Some of these RNAs are derived from intron sequences, while others, transcribed from intergenic regions form a subset of lncRNAs called lincRNAs (long intergenic noncoding RNAs). Yet other lncRNAs are produced as antisense transcripts of coding genes. An astounding 30,000 transcripts in humans are thought to be lncRNAs, but little is known of their function. From the few lncRNAs that have been intensively studied, it is evident that they do not all function in the same way. However, they appear to affect gene expression in a variety of ways including modification of chromatin structure, regulation of splicing, or serving as structural scaffolds for the assembly of nucleoprotein complexes. Additional mechanisms will doubtless be uncovered as these fascinating RNAs are investigated in years to come.

Regulation of translation

The synthesis of proteins is dependent on the availability of the mRNAs encoding them. If an mRNA is blocked at its 5' end, it cannot be translated. The rate of degradation of an mRNA will influence how long it is around to direct the synthesis of the protein it codes for. Gene expression can also, therefore, be regulated by mechanisms that alter the rate of mRNA degradation. Regulation of translation is used to control the production of many proteins. Two examples, ferritin and the transferrin receptor, are important for iron storage and transport in cells. Ferritin is an iron-binding protein that sequesters iron atoms in cells to keep them from reacting. When iron levels are high, there is a need for more ferritin than when iron levels are low. How are ferritin levels regulated? The 5'UTR of the ferritin mRNA contains a 28-nucleotide sequence called the Iron Response Element, or IRE (Figure 7.127). When iron levels are low, the IRE is bound by a protein. The presence of the IRE-binding protein at the 5'UTR blocks translation of the ferritin mRNA. However, if iron levels are high, the iron binds to the IRE-binding protein, which undergoes a conformational change and dissociates from the IRE. This frees up the 5' end of the ferritin mRNA for ribosome assembly and translation, producing more ferritin.

The other protein involved in iron transport, the transferrin receptor, is required for uptake of iron into cells, when intracellular iron levels are low. In the case of the transferrin receptor, it is when iron levels are low that more of it is needed. When iron levels are high, there is no need to make more transferrin receptor. The mRNA encoding the transferrin receptor also has IRE sequences, but in this case, the IRE is situated in the 3'UTR of the transcript (Figure 7.128). The IRE is, as in the case of ferritin, bound by the IRE-binding protein. When iron levels in the cell are high, the iron binds the IRE-binding protein, which dissociates from the IRE. This leaves the 3'UTR susceptible to attack by RNases, leading to degradation of the transferrin receptor mRNA. At times when iron levels are low, the IRE-binding protein remains bound to the 3' UTR of the mRNA, stabilizing it and permitting more transferrin receptor to be made by translation.

Gene expression is controlled at many steps

As can be seen from the examples in this section, regulation of gene expression in eukaryotic cells is a function of multiple mechanisms that act at different stages in the flow of information from DNA to protein, responding to the internal state of the cell as well as external conditions and signals.

Information Processing: Gene Expression

803

YouTube Lectures

by Kevin

HERE & HERE

804

Figure 7.102 - Multiple levels of control of gene expression

Wikipedia

805

Figure 7.103 - Prokaryotic genes organized in an operon

Wikipedia

Figure 7.104 - Protein binding sites in the lac regulatory region

Image by Martha Baker

Interactive Learning

Module

HERE

806

Figure 7.105 - Lac operon structure and products

Image by Martha Baker

Figure 7.106 - Lac operon in the absence (middle) and presence (bottom) of inducer

Image by Martha Baker

807

Figure 7.107 - Allolactose (top) and lactose (bottom)

Figure 7.108 - CAP (blue) bound to the DNA adjacent to the lac promoter (orange). cAMP shown in pink.

Wikipedia

Figure 7.109 - Lac operon in the presence (top) and absence (bottom) of glucose

Image by Martha Baker

808

Figure 7.110 - Structure and regulation of the trp operon

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

809

Figure 7.111 - Attenuation in regulation of the trp operon

Wikipedia

Figure 7.112 - Sequence of the leader region of the trp operon

XX AUG AAA GCA AUU UUC GUA CUG AAA GGU UGG UGG CGC ACU UCC UGA -XX

MET LYS ALA ILE PHE VAL LEU LYS GLY TRP TRP ARG THR SER STOP

810

Figure 7.113 - Riboswitch features

811

Figure 7.114 - Regulatory sequences for a eukaryotic gene

Wikipedia

812

Figure 7.115 - DNA looping allows contact between activator bound at a distant enhancer and the basal transcription complex

Image by Martha Baker

813

Figure 7.116 - Transcription factors in regulation of eukaryotic transcription

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

814

Figure 7.117 - Binding of c-myc protein to its target DNA sequence

Wikipedia

Figure 7.118 Activators bound at multiple sites can regulate transcription from a given promoter

OpenStax

815

Figure 7.119 - Transcriptional activation (right) and deactivation (left) by histone modification

Wikipedia

816

Figure 7.120 - Chromatin configuration affects transcription

Wikipedia

Interactive Learning

Module

HERE

817

Figure 7.121 - Inactivation of transcription by CpG methylation

Image by Indira Rajagopal

YouTube Lectures

by Kevin

HERE & HERE

818

Figure 7.122 - Epigenetic changes through histone and DNA modification

819

Figure 7.123 - miRNAs function in the regulation of gene expression

Wikipedia

Figure 7.124 Pre-miRNA hairpin structures with the mature guide miRNAs shown in red

Wikipedia

820

Figure 7.125 - Gene silencing by siRNA

Image by Pehr Jacobson

821

Figure 7.126 - Processed siRNA duplex with perfect base-pairing, 5’ phosphates and two bases overhanging at each 3’ end

822

Figure 7.127 -Regulation of ferritin mRNA translation

Image by Aleia Kim

YouTube Lectures

by Kevin

HERE & HERE

823

Figure 7.128 -Regulation of transferrin receptor mRNA translation

Image by Aleia Kim

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

God Bless These Complexes

To the tune of "God Bless America"

Metabolic Melodies Website HERE

All information in

Cells' DNA

Just increases

With pieces

Mixed and matched in the mRNAs

Linking exons

All together

Using snurps in

Complex-ES

God bless the spliceosomes

And trans-crip-tomes

(slow and loud) God bless the spliceosomes

And my ge-nome

Your blueprint info is

In DNA

Since you need it

Proofread it

Or you'll mutate the mRNA

You can translate

All the codons

With the cells' gen-

et-ic code

God bless the ribosomes

They translate code

(slow and loud) God bless the ribosomes

And proteomes

Recording by David Simmons

Lyrics by Kevin Ahern
Recording by David Simmons Lyrics by Kevin Ahern

The Book of Life

To the tune of "The Look of Love"

Metabolic Melodies Website HERE

The book of life - the stuff of dreams

Is everywhere, it seems

The book of life, is biochemistry and

Its words fill every day

Just what it says is written in the DNA

I just want to get to know it

How the info's coded

What are all the secrets?

Ribosomes can read it

Goodness knows it's needed

And so its alphabet's

In codon forms

For ribosome bookworms

They read it right

A protein's function to its sequence corresponds

It's not just randomly created peptide bonds

What a marvel of creation, how they do translation

Of m-R-N-A chains,

Using bits of glycine

Proline and some lysine

Translate the code

Instrumental

I just marvel at the knowledge

That I got in college

To learn all the secrets

Double helix spaces

Complementary bases

Pyrimidines

Paired to purines

The book of life

Recording by Carol Adriane Smith

Lyrics by Kevin Ahern

Recording by Carol Adriane Smith Lyrics by Kevin Ahern

This page titled 7.8: Gene Expression is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

7.9: Signaling

Source: BiochemFFA_7_8.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

Up to this point we have considered how cells carry out biochemical reactions and how they regulate the expression of the genes in response to their internal and external environments. It is intuitively obvious that even unicellular organisms must be able to sense features of their environment, such as the presence of nutrients, if they are to survive. In addition to being able to receive and respond to information from the environment, multicellular organisms must also find ways by which their cells can communicate among themselves.

Coordination

Since different cells take on specialized functions in a multicellular organism, they must be able to coordinate activities. Cells grow, divide, or differentiate in response to specific signals. They may change shape or migrate to another location. At the physiological level, cells in a multicellular organism, must respond to everything from a meal just eaten to injury, threat, or the availability of a mate. They must know when to divide, when to undergo apoptosis (programmed cell death), when to store food, and when to break it down. A variety of mechanisms have arisen to ensure that cell-cell communication is not only possible, but astonishingly swift, accurate and reliable.

How are signals sent between cells? Like pretty much everything that happens in cells, signaling is dependent on molecular recognition. The basic principle of cell-cell signaling is simple. A particular kind of molecule, sent by a signaling cell, is recognized and bound by a receptor protein in (or on the surface of) the target cell. The signal molecules are chemically varied- they may be proteins, short peptides, lipids, nucleotides or catecholamines, to name a few.

Signal properties

The chemical properties of the signal determine whether its receptors are on the cell surface or intracellular. If the signal is small and hydrophobic it can cross the cell membrane and bind a receptor inside the cell. If, on the other hand, the signal is charged, or very large, it would not be able to diffuse through the plasma membrane. Such signals need receptors on the cell surface, typically transmembrane proteins that have an extracellular portion that binds the signal and an intracellular part that passes on the message within the cell (Figure 7.130).

Receptors are specific for each type of signal, so each cell has many different kinds of receptors that can recognize and bind the many signals it receives. Because different cells have different sets of receptors, they respond to different signals or combinations of signals. The binding of a signal molecule to a receptor sets off a chain of events in the target cell. These events could cause change in various ways, including, but not limited to, alterations in metabolic pathways or gene expression in the target cell.

How the binding of a signal to a receptor brings about change in cells is the topic of this section. We will examine a few of the major receptor types and the consequences of signal binding to these receptors. Although the specific molecular components of the various signal transduction pathways differ, they all have some features in common (Figure 7.131):

- The binding of a signal to its receptor is usually followed by the generation of a new signal(s) within the cell. The process by which the original signal is converted to a different form and passed on within the cell to bring about change is called signal transduction.
- Most signaling pathways have multiple signal transduction steps by which the signal is relayed through a series of molecular messengers that can amplify and distribute the message to various parts of the cell.
- The last of these messengers usually interacts with a target protein(s) and changes its activity, often by phosphorylation.
- When a signal sets a particular pathway in motion, it is acting like an ON switch. This means that once the desired result has been obtained, the cell must have a mechanism that acts as an OFF switch.

Understanding this underlying similarity is helpful, because learning the details of the different pathways becomes merely a matter of identifying which molecular component performs a particular function in each individual case. We will consider several different signal transduction pathways, each mediated by a different kind of receptor.

Ligand-gated ion channel receptors

The simplest and fastest of signal pathways is seen in the case of signals whose receptors are gated ion channels (Figure 7.132). Gated ion channels are made up of multiple transmembrane proteins that create a pore, or channel, in the cell membrane. Depending upon its type, each ion channel is specific to the passage of a particular ionic species. The term "gated" refers to the fact that the ion channel is controlled by a "gate" which must be opened to allow the ions through. The gates are opened by the binding of an incoming signal (ligand) to the receptor, allowing the almost instantaneous passage of millions of ions from one side of the membrane to the other. Changes in the interior environment of the cell are thus brought about in microseconds and in a single step.

Swift response

This type of swift response is seen, for example, in neuromuscular junctions, where muscle cells respond to a message from the neighboring nerve cell (Figure 7.133). The nerve cell releases a neurotransmitter signal into the synaptic cleft, which is the space between the nerve cell and the muscle cell it is "talking to". An example of such a neurotransmitter signal is acetylcholine. When the acetylcholine molecules are released into the synaptic cleft, they diffuse rapidly till they reach their receptors on the membrane of the muscle cell. The binding of the acetylcholine to its receptor, an ion channel on the membrane of the muscle cell, causes the gate in the ion channel to open. The resulting ion flow through the channel can immediately change the membrane potential of the cell. This, in turn, can trigger other changes in the cell.

The speed with which changes are brought about in neurotransmitter signaling is evident when you think about how quickly you remove your hand from a hot surface. Sensory neurons carry information to the brain from your hand on the hot surface and motor neurons signal to your muscles to move the hand, in less time than it took you to read this sentence!

Nuclear hormone receptors

The receptors for signals like steroid hormones are part of a large group of proteins known as the nuclear hormone receptor superfamily. These receptors recognize and bind not only steroid hormones, but also retinoic acid, thyroid hormone, vitamin D and other signals. The subset of the nuclear hormone receptors that bind steroid hormones are intracellular proteins. Steroid hormones (Figure 7.135), as you are aware, are related to cholesterol, and as hydrophobic molecules, they are able to cross the cell membrane by themselves. This is unusual, as most signals coming to cells are incapable of crossing the plasma membrane, and thus, must have cell surface receptors.

Once within the cell, steroid hormones bind to their receptors, which may reside in the cytoplasm or in the nucleus. Steroid hormone receptors are proteins with a double life: they are actually dormant transcription regulators that are inactive till a steroid hormone binds and causes a conformational change in them. When this happens, the receptors, with the hormone bound, bind to regulatory sequences in the DNA and modulate gene expression. Because steroid hormone receptors act by modulating gene expression, the responses to steroid hormones are relatively slow. (There are also some effects of steroid hormones that do not involve transcriptional regulation, but the majority work through changing gene expression.) Like other transcriptional activators, steroid receptors have a DNA-binding domain (DBD) and an activation domain. They also have a ligand-binding domain (LBD) that binds the hormone.

Glucocorticoid receptor

Examples of such signaling pathways are those mediated by the glucocorticoid receptor (Figures 7.136 & 7.137). Glucocorticoids, sometimes described as stress hormones, are made and secreted by the adrenal cortex. Physiologically, they serve to maintain homeostasis in the face of stress and exhibit strong anti-inflammatory and immunosuppressive properties. Because of these effects, synthetic glucocorticoids are used in the treatment of a number of diseases from asthma and rheumatoid arthritis to multiple sclerosis. All of these effects are mediated through the signaling pathway which starts with the binding of a glucocorticoid hormone to its receptor. Recall that steroids can cross the plasma membrane, so glucocorticoids can diffuse into the cell and bind their receptors which are in the cytoplasm.

In the absence of the signal, glucocorticoid receptors are found bound to a protein chaperone, Hsp90 (Figure 7.137). This keeps the receptors from being transported to the nucleus. When a glucocorticoid molecule binds the receptor, the receptor undergoes a conformational change and dissociates from the Hsp90. The receptor, then, with the hormone bound, translocates into the nucleus. In the nucleus, it can increase the transcription of target genes by binding to specific regulatory sequences (labeled

HRE for hormone-response elements). The binding of the hormone-receptor complex to the regulatory elements of hormone-responsive genes modulates their expression. Many of these genes encode anti-inflammatory proteins, and their increased production accounts for the physiological effect of corticosteroid therapies.

The steroid receptor pathways are relatively simple and have only a couple of steps (Figure 7.138). Most other signaling pathways involve multiple steps in which the original signal is passed on and amplified through a number of intermediate steps, before the cell responds to the signal.

Cell surface receptors

We will now take a look at two signaling pathways, each mediated by a major class of cell surface receptor- the G-protein coupled receptors (GPCRs) and the receptor tyrosine kinases (RTKs). While the specific details of the signaling pathways that follow the binding of signals to each of these receptor types are different, it is easier to learn them when you can see what the pathways have in common, namely, interaction of the signal with a receptor, followed by relaying and amplification of the signal through a variable number of intermediate molecules, with the last of these molecules interacting with a target or target proteins and modifying their activity in the cell.

G-protein coupled receptors

G-protein coupled receptors (GPCRs) are involved in responses of cells to many different kinds of signals, from epinephrine, to odors, to light. In fact, a variety of physiological phenomena including vision, taste, smell, and the fight-or-flight response are mediated by GPCRs. What are G-protein coupled receptors?

G-protein coupled receptors are cell surface receptors that pass on the signals that they receive with the help of guanine nucleotide binding proteins (a.k.a. G-proteins). Before thinking any further about the signaling pathways downstream of GPCRs, it is necessary to know a few important facts about these receptors and the G-proteins that assist them.

Though there are hundreds of different G-protein coupled receptors, they all have the same basic structure (Figure 7.139):

They all consist of a single polypeptide chain that threads back and forth seven times through the lipid bilayer of the plasma membrane. For this reason, they are sometimes called seven-pass transmembrane (7TM) receptors. One end of the polypeptide forms the extracellular domain that binds the signal while the other end is in the cytosol of the cell.

When a ligand (signal) binds the extracellular domain of a GPCR, the receptor undergoes a conformational change, on its cytoplasmic side, that allows it to interact with a G-protein that will then pass the signal on to other intermediates in the signaling pathway.

G-proteins

What is a G-protein? As noted above, a G-protein is a guanine nucleotide-binding protein that can interact with a G-protein linked receptor. G-proteins are associated with the cytosolic side of the plasma membrane, where they are ideally situated to interact with the tail of the GPCR, when a signal binds to the GPCR. There are many different G-proteins, all of which share a characteristic structure- they are composed of three subunits called α, β and γ (Figure 7.140). Because of this, they are sometimes called heterotrimeric G proteins (hetero=different, trimeric= having three parts).

Ligand binding

The guanine nucleotide binding site is on the α subunit of the G-protein. This site can bind GDP or GTP. The α subunit also has a GTPase activity, i.e., it is capable of hydrolyzing a GTP molecule bound to it into GDP.

In the unstimulated state of the cell, that is, in the absence of a signal bound to the GPCR, the G-proteins are found in the trimeric form (α-β-γ bound together) and the α subunit has a GDP molecule bound to it. In this form, the α subunit is inactive. With this background on the structure and general properties of the GPCRs and the G-proteins, we can now look at what happens when a signal arrives at the cell surface and binds to a GPCR (Figure 7.141).

The signaling pathway

The binding of a signal molecule by the extracellular part of the G-protein linked receptor causes the cytosolic tail of the receptor to interact with, and alter the conformation of, a G-protein associated with the inner face of the plasma membrane.
This has two consequences. First, the α subunit of the G-protein loses its GDP and binds a GTP, instead. Second, the G-protein breaks up into the GTP-bound α part and the β-γ part.

The binding of GTP to the α subunit and its dissociation from the β-γ subunits activate the α subunit. The activated α subunit can diffuse freely along the cytosolic face of the plasma membrane and act upon its targets. (The β-γ unit is also capable of activating its own targets.)

What happens when G-proteins interact with their target proteins? That depends on what the target is. G-proteins interact with different kinds of target proteins, of which we will examine two major categories:

Ion channels

We have earlier seen that some gated ion channels can be opened or closed by the direct binding of neurotransmitters to a receptor that is an ion-channel protein. In other cases, ion channels are regulated by the binding of G-proteins. That is, instead of the signal directly binding to the ion channel, it binds to a GPCR, which activates a G-protein that then may cause opening of the ion channel, either directly, by binding to the channel, or indirectly, through activating other proteins that can bind to the channel. The change in the distribution of ions across the plasma membrane causes a change in the membrane potential.

Enzyme activation

The interaction of G-proteins with their target enzymes can regulate the activity of the enzyme, either increasing or decreasing its activity. The change in activity of the target enzyme, in turn, results in downstream changes in other proteins in the cell, and alters the metabolic state of the cell. This is best understood by examining the well-studied response of cells to epinephrine, mediated through the β-adrenergic receptor, a type of G-protein coupled receptor.

Epinephrine (Figure 7.142), also known as adrenaline, is a catecholamine that plays an important role in the body's 'fight or flight' response. In response to stressful stimuli, epinephrine is secreted into the blood, to be carried to target organs whose cells will respond to this signal. If you were walking down a dark alley in an iffy neighborhood, and you heard footsteps behind you, your brain would respond to potential danger by sending signals that ultimately cause the adrenal cortex to secrete epinephrine into the blood stream. The epinephrine circulating in your system has many effects, including increasing your heart rate, but among its prime targets are your muscle cells. The reason for this is that your muscle cells store energy in the form of glycogen, a polymer of glucose. If you need to run or fight off an assailant, your cells will need energy in the form of glucose.

But how does epinephrine get your cells to break down the glycogen into glucose? Binding of epinephrine to the β-adrenergic receptor on the surface of the cells causes the receptor to activate a G-protein associated with its cytoplasmic tail. As described above, this leads to the α subunit exchanging its GDP for GTP and dissociating from the β-γ subunits. The activated α subunit then interacts with the enzyme adenylate cyclase (also known as adenylyl cyclase) stimulating it to produce cyclic AMP (cAMP) from ATP. Cyclic AMP is often described as a "second messenger", in that it serves to spread the signal received by the cell. How does cAMP accomplish this?

cAMP molecules bind to, and activate an enzyme, protein kinase A (PKA - Figure 7.145). PKA is composed of two catalytic and two regulatory subunits that are bound tightly together. Upon binding of cAMP, the catalytic subunits are released from the regulatory subunits, allowing the enzyme to carry out its function, namely phosphorylating other proteins. Thus, cAMP can regulate the activity of PKA, which in turn, by phosphorylating other proteins can change their activity. In this case, the relevant protein that is activated is an enzyme, phosphorylase kinase. This enzyme can then phosphorylate and activate glycogen phosphorylase, the enzyme ultimately responsible for breaking glycogen down into glucose-1-phosphate - readily converted to glucose. The activation of glycogen phosphorylase supplies the cells with the glucose they need, allowing you to fight or flee, as you might see fit. Simultaneously, PKA also phosphorylates another enzyme, glycogen synthase. In the case of glycogen synthase, phosphorylation inactivates it, and prevents free glucose from being used up for glycogen synthesis, ensuring that your cells are amply supplied with glucose (Figure 7.146).

Common pattern

Although the steps described above seem complicated, they follow the simple pattern outlined at the beginning of this section:

• Binding of signal to receptor
• Several steps where the signal is passed on through intermediate molecules (G-proteins, adenylate cyclase, cAMP, and finally, PKA)
• Phosphorylation of target proteins by the kinase, leading to changes in the cell. The specific changes depend on the proteins that are phosphorylated by the PKA.

Why so many steps? If you need to activate glycogen phosphorylase to break down glucose in a hurry, why not have a system in which binding of a signal to the receptor directly activated the target enzyme?

The answer to this puzzle is simple: there is amplification of the signal at every step of the pathway. A single signal molecule binding to a receptor sets in motion a cascade of reactions, with the signal getting larger at each step, so that binding of one epinephrine molecule to its receptor results in the activation of a million glycogen phosphorylase enzyme molecules!

Turning signals off

If the signal binding to the receptor serves as a switch that sets these events in motion, there must be mechanisms to turn the pathway off. The first is at the level of the receptor itself. A kinase called G-protein receptor kinase (GRK) phosphorylates the cytoplasmic tail of the receptor. The phosphorylated tail is then bound by a protein called arrestin, preventing further interaction with a G-protein.

The next point of control is at the G-protein. Recall that the α subunit of the G-protein is in its free and activated state when it has GTP bound, and that it associates with the β-γ subunits and has a GDP bound when it is inactive. We also know that the α subunit has an activity that enables it to hydrolyze GTP to GDP. This GTP-hydrolyzing activity makes it possible for the α subunit, once it has completed its task, to return to its GDP bound state, re-associate with the β-γ part and become inactive again.

A third "off switch" is further down the signaling pathway, and controls the level of cAMP. We just noted that cAMP levels increase when adenylate cyclase is activated. When its job is done, cAMP is broken down by an enzyme called phosphodiesterase (Figure 7.147). When cAMP levels drop, PKA returns to its inactive state, putting a halt to the changes brought about by the activation of adenylate cyclase by an activated G-protein.

Yet another way that the effects of this pathway can be turned off is at the level of the phosphorylated target proteins. These proteins, which are activated by phosphorylation, can be returned to their inactive state by the removal of the phosphates by phosphatases.

Receptor tyrosine kinases

Another major class of cell surface receptors are the receptor tyrosine kinases or RTKs. Like the GPCRs, receptor tyrosine kinases bind a signal, then pass the message on through a series of intracellular molecules, the last of which acts on target proteins to change the state of the cell.

As the name suggests, a receptor tyrosine kinase is a cell surface receptor that also has a tyrosine kinase activity. The signal binding domain of the receptor tyrosine kinase is on the cell surface, while the tyrosine kinase enzymatic activity resides in the cytoplasmic part of the protein (Figure 7.148). A transmembrane α helix connects these two regions of the receptor.

What happens when signal molecules bind to receptor tyrosine kinases? Binding of signal molecules to the extracellular domains of receptor tyrosine kinase proteins causes two receptor molecules to dimerize (come together and associate - Figure 7.149). This brings the cytoplasmic tails of the receptors close to each other and causes the tyrosine kinase activity of these tails to be turned on. The activated tails then phosphorylate each other on several tyrosine residues (Figure 7.150). This is called autophosphorylation.

The phosphorylation of tyrosines on the receptor tails triggers the assembly of an intracellular signaling complex on the tails. The newly phosphorylated tyrosines serve as binding sites for a variety of signaling proteins that then pass the message on to yet other proteins to bring about changes in the cell. Receptor tyrosine kinases mediate responses to a large number of signals, including peptide hormones like insulin and growth factors like epidermal growth factor (EGF). We will examine how insulin and EGF act on cells by binding to receptor tyrosine kinases.

Insulin receptor

Insulin plays a central role in the uptake of glucose from the bloodstream. It increases glucose uptake by stimulating the movement of glucose receptor GLUT4 to the plasma membrane of cells.

How does insulin increase GLUT4 concentrations in the cell membrane? The binding of insulin to the insulin receptor (IR - Figure 7.151), results in dimerization of the receptor monomers and subsequent autophosphorylation of the cytosolic kinase domains. The activated tyrosine kinase domains also phosphorylate intracellular proteins called Insulin Receptor Substrates or IRS proteins. These proteins interact with, and activate another kinase called the PI3-kinase. PI3-kinase then catalyzes the formation of the lipid molecule PIP3, which serves to activate yet another kinase, PDK1, which in turn, activates the Akt group of kinases. It is this group of enzymes that appears to increase the translocation of the GLUT4 to the plasma membrane (Figure 7.152), as cells that lack functional Akts exhibit poor glucose uptake and insulin resistance.

EGFR pathway

Epidermal growth factor, EGF, is an important signaling molecule involved in growth, proliferation and differentiation in mammalian cells. EGF acts through the EGF receptor, EGFR, a receptor tyrosine kinase (Figure 7.153). Because of its role in stimulating cell proliferation and because overexpression of EGFR is associated with some kinds of cancers, EGFR is the target for many anti-cancer therapies. We can trace the signal transduction pathway from the binding of EGF to its receptor to the stimulation of cell division.

EGF binding to the EGFR is followed by receptor dimerization and stimulation of the tyrosine kinase activity of the cytosolic domains of the EGFR. Autophosphorylation of the receptor tails is followed by the assembly of a signaling complex nucleated by the binding of proteins that recognize phosphotyrosine residues. An important protein that is subsequently activated by the signaling complexes on the receptor tyrosine kinases is called Ras (Figure 7.154). The Ras protein is a monomeric guanine nucleotide binding protein that is associated with the cytosolic face of the plasma membrane
(in fact, it is a lot like the α subunit of trimeric G-proteins). Just like the α subunit of a G-protein, Ras is active when GTP is bound to it and inactive when GDP is bound to it. Also, like the α subunit, Ras can hydrolyze the GTP to GDP.

Ras activation

Activation of Ras accompanies the exchange of the GDP bound to the inactive Ras for a GTP. Activated Ras triggers a phosphorylation cascade of three protein kinases, which relay and distribute the signal. These protein kinases are members of a group called the MAP kinases (Mitogen Activated Protein Kinases). The final kinase in this cascade phosphorylates various target proteins, including enzymes and transcriptional activators that regulate gene expression.

The phosphorylation of various enzymes can alter their activities, and set off new chemical reactions in the cell, while the phosphorylation of transcriptional activators can change which genes are expressed. The combined effect of changes in gene expression and protein activity alter the cell's physiological state and promote cell division.

Once again, in following the path of signal transduction mediated by RTKs, it is possible to discern the same basic pattern of events: a signal is bound by the extracellular domains of receptor tyrosine kinases, resulting in receptor dimerization and autophosphorylation of the cytosolic tails, thus conveying the message to the interior of the cell.

The message is then passed on via a signaling complex to proteins that stimulate a series of kinases. The terminal kinase in the cascade acts on target proteins and brings about in changes in protein activities.

What is the OFF switch for RTKs? It turns out that RTKs with the signal bound can be endocytosed into the cell and broken down. That is, the region of the plasma membrane that the RTK is on can be internally pinched off into a vesicle containing the ligand-bound receptor which is then targeted for degradation.

Ras, which is activated by GTP binding, can also be deactivated by hydrolysis of the GTP to GDP. The importance of this mechanism for shutting down the pathway is evident in cells that have a mutant ras gene encoding a Ras protein with defective GTPase activity. Unable to shut off Ras, the cells continue to receive a signal to proliferate. The National Cancer Institute estimates that more than 30% of human cancers are driven by mutations in ras genes.

The descriptions above provide a very simple sketch of some of the major classes of receptors and deal primarily with the mechanistic details of the steps by which signals received by various types of receptors bring about changes in cells. A major take-home lesson is the essential similarity of the different pathways. Another point to keep in mind is that while we have looked at each individual pathway in isolation, a cell, at any given time receives multiple signals that set off a variety of different responses at once (Figure 7.155). The pathways described above show a considerable degree of "cross-talk" and the response to any given signal is affected by the other signals that the cell receives simultaneously. The multitude of different receptors, signals, and the combinations thereof are the means by which cells are able to respond to an enormous variety of different circumstances.

RTKs, cancer and cancer therapies

As described above, binding of EGF to its receptor triggers a signaling pathway that results in the activation of a series of Mitogen Activated Protein Kinases (MAP kinases). These kinases are so-called because they are activated by a mitogen, a molecule, like EGF and other growth factors, that stimulates mitosis or cell division. The final kinase in the MAP kinase cascade phosphorylates a number of target proteins, many of them transcription factors, that when activated, increase the expression of genes associated with cell proliferation.

Given that the EGF-receptor pathway normally functions to stimulate cell division, it is not surprising that malfunctions in the pathway could lead to uncontrolled cell proliferation, or cancer. Next, we will take a brief look at some examples of such defects.

HER2

The human EGF receptor (HER) family has four members, HER1, HER2, HER3 and HER4. These are all receptor tyrosine kinases, cell surface receptors that bind EGF (Figure 7.157) and stimulate cell proliferation.

A crucial step in the signal transduction pathway is the dimerization of the receptors following binding of the signal, EGF, to the receptor. While HER1, HER3 and HER4 must bind the signal to dimerize, the structure of the HER2 receptor can, apparently, allow the receptor monomers to dimerize independently of EGF binding.

This means that the downstream events of the signaling pathway can be triggered even in the absence of a growth signal. In normal cells, only a few HER2 receptors are expressed at the cell surface, so this property of HER2 plays a relatively minor role in stimulating cell division. However, in about a quarter of breast cancer patients, HER2 receptors are overexpressed, leading to increased dimerization and subsequent uncontrolled cell proliferation.

Breast cancers that are HER2-positive can be more aggressive with a greater tendency to metastasize (spread) so therapy that blocks HER2 signaling is key in successful treatment of such cancers. Herceptin, a monoclonal antibody against the HER2 receptor, has been shown to be an effective treatment against Her2-positive breast cancers. Herceptin works by binding specifically to the extracellular domain of the HER2 receptor (Figure 7.158). This prevents dimerization of the receptor and thus blocks downstream signaling. Additionally, the binding of the Herceptin antibody to the receptor signals the immune system to destroy the HER2-positive cells.

Bcr-abl

Another example of a cancer caused by defects in an RTK signaling pathway is chronic myeloid leukemia (CML). Patients with CML have an abnormal receptor tyrosine kinase that is the product of a hybrid gene called bcr-abl, formed by the breakage and rejoining of chromosomes 9 and 22. This abnormal tyrosine kinase is constitutively dimerized, even when no signal is bound. As a result, it continuously signals cells to divide, leading to the massive proliferation of a type of blood cells called granulocytes.

As with HER2, the problem in CML is a receptor tyrosine kinase that dimerizes in the absence of a growth signal. The approach in this case was to target the next step in the signaling pathway. As you know, dimerization of RTKs activates the tyrosine kinase domain of the receptor, which results in the autophosphorylation of the cytoplasmic domains of both monomers. The phosphorylated tyrosines serve to recruit a number of other signaling proteins that pass the signal on within the cell.

In the case of the bcr-abl RTK, the drug Gleevec (imatinib) was designed to bind near the ATP-binding site of the tyrosine kinase domain. This "locks" the site in a conformation that inhibits the enzymatic activity of the tyrosine kinase and thus blocks downstream signaling. With no "grow" signal passed on, cells stop proliferating.

Information Processing: Signaling

827

YouTube Lectures

by Kevin

HERE & HERE

828

Figure 7.130 - Schematic representation of a transmembrane receptor protein. E = extracellular; P = plasma membrane; I = intracellular

Wikipedia

Figure 7.129 - Some examples of signal molecules

829

Figure 7.132 - Ligand-gated ion channel receptor opening in response to a signal (ligand)

Wikipedia

Figure 7.131 -General features of signal transduction pathways

830

Figure 7.133 - Neuromuscular signaling - A = motor neuron axon; B = axon terminal; C = synaptic cleft; D = muscle cell; E = myofibril . Steps in the process - 1) action potential reaches the axon terminal; 2) voltage-dependent calcium gates open; (3) neurotransmitter vesicles fuse with the presynaptic membrane and acetylcholine (ACh) released into the synaptic cleft; (4) ACh binds to postsynaptic receptors on the sarcolemma; (5) ACh binding causes ion channels to open and allows sodium ions to flow across the membrane into the muscle cell; 6) flow of sodium ions across the membrane into the muscle cell generates action potential which travels to the myofibril and results in muscle contraction.

Wikipedia

831

Figure 7.134 - Nerve systems

Wikipedia

832

Figure 7.135 - Steroid hormones structures, with the names of their receptors

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

833

Figure 7.136 - Glucocorticoid receptor with its three domains - DNA binding (left), activator domain (top), and ligand binding domain (boxed).

Wikipedia

Figure 7.137 - Glucocorticoid signaling pathway

Wikipedia

834

Figure 7.138 - Steroid hormone signaling

Image by Aleia Kim

835

Figure 7.139 - Structure of a G-protein linked receptor

Wikipedia

836

Figure 7.140- A heterotrimeric G-protein: α subunit in blue, βγ subunits red and green

Interactive Learning

Module

HERE

837

Figure 7.141 - Cycle of G-protein activation - 1) binding of ligand; 2) change of receptor structure; 3) stimulation of α-subunit; 4) binding of GTP, release of GDP; 5) separation of α-subunit from β-γ; 6) hydrolysis of GTP by α-subunit and return to inactive state.

Wikipedia

YouTube Lectures

by Kevin

HERE & HERE

838

Figure 7.142 - β2-adrenergic receptor embedded in membrane (gray)

Wikipedia

Figure 7.143 - Epinephrine

Wikipedia

839

Figure 7.144 - G-protein coupled receptor. Signal starts with ligand binding (orange circle). Gs = G-protein; AC = adenylate cyclase.

Wikipedia

Figure 7.145 - Activation of Protein Kinase A by cAMP

Image by Martha Baker

840

Figure 7.146 - Simultaneous activation of glycogen breakdown and inhibition of glycogen synthesis by epinephrine's binding of b-adrenergic receptor. Red enzyme names = activated forms; black enzyme names = inactivated forms; GPb = glycogen phosphorylase b; GPa = glycogen phosphorylase a.

Image by Penelope Irving

841

β-Adrenergic Signaling Off Switches

1. GRK Phosphorylates Receptor Tail

Receptor Tail Bound by Arrestin

2. α Subunit G-protein Cleaves GTP to GDP

β-γ subunits Reassociate with α Subunit

3. cAMP Hydrolyzed by Phosphodiesterase

PKA Becomes Inactive

4. Dephosphorylation of Phosphorylated Proteins by Phosphoprotein Phosphatase

β-Adrenergic Signaling On Switches

1. Binding of Signal Molecule to Receptor

2. Passage of Signal Through Several Molecules (G-proteins, Adenylate Cyclase, cAMP, PKA)

3. Phosphorylation of Target Proteins

842

Figure 7.147 - Cyclic AMP is broken down by phosphodiesterase

Figure 7.148 - Structure of a receptor tyrosine kinase

YouTube Lectures

by Kevin

HERE & HERE

843

Figure 7.149 - Signal binding results in receptor dimerization and activation of tyrosine kinase activity

Figure 7.150 - Activated tyrosine kinases phosphorylate tyrosines on the receptor tails.

Figure 7.151 -The insulin receptor, a receptor tyrosine kinase

Wikipedia

Figure 7.152 - Effects of insulin binding to its receptor tyrosine kinase: 1) insulin binding; 2) activation of protein activation cascades. These include: 3) translocation of Glut-4 transporter to plasma membrane and influx of glucose; 4) glycogen synthesis; 5) glycolysis; and 6) fatty acid synthesis.

Wikipedia

844

Interactive Learning

Module

HERE

845

Figure 7.153 - EGFR signaling beginning at top with binding of EGF, dimerization of receptor, transmission of signal through

proteins, activation of kinases, phosphorylation of transcription factors and effects on transcription

Image by Aleia Kim

Figure 7.154 - Ras with GTP bound

Wikipedia

This page titled 7.9: Signaling is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8: BASIC TECHNIQUES

The environment of a cell is very complex, making it difficult to study individual reactions, enzymes, or pathways in situ. The traditional approach used by biochemists for the study of these things is to isolate molecules, enzymes, DNAs, RNAs, and other items of interest so they can be analyzed independently of the millions of other processes occurring simultaneously. Today, these approaches are used side by side with newer methods that allow us to understand events inside cells on a larger scale- for example, determining all the genes that are being expressed at a given time in specific cells. In this section we take a brief look at some commonly used methods used to study biological molecules and their interactions.

8.1: Cell Lysis

8.2: Fractionation and Chromatography Techniques

8.3: Electrophoresis

8.4: Detection, identification and quantitation of specific nucleic acids and proteins

8.5: Transcriptomics

8.6: Isolating Genes

8.7: Polymerase Chain Reaction (PCR)

8.8: Reverse Transcription

8.9: FRET

8.10: Genome Editing (CRISPR)

8.11: Protein Cleavage

8.12: Membrane Dynamics (FRAP)

This page titled 8: Basic Techniques is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.1: Cell Lysis

Source: BiochemFFA_8_1.pdf. The entire textbook is available for free from the authors at http://biochem.science.oregonstate.edu/content/biochemistry-free-and-easy

To separate compounds from cellular environments, one must first break open (lyse) the cells. Cells are broken open, in buffered solutions, to obtain a lysate. There are several ways of accomplishing this.

- **Osmotic shock and enzymes:** One way to lyse cells is by lowering the ionic strength of the medium the cells are in. This can cause cells to swell and burst. Mild surfactants may be used to disrupt membranes. Most bacteria, yeast, and plant tissues are resistant to osmotic shocks, because of the presence of cell walls, and stronger disruption techniques are usually 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), proteases, mannases, and others.
- **Mechanical disruption:** Mechanical agitation may be employed in the form of beads that are shaken with a mixture of cells. In this method, cells are bombarded with tiny, glass beads that break the cells open. Sonication (20-50 kHz sound waves) provides an alternative type of agitation that can be effective. The method is noisy, however, and generates heat that can be problematic for heat-sensitive compounds.
- **Pressure disruption:** 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) and then the pressure is rapidly released. The rapid pressure change causes dissolved gases in cells to be released as bubbles which, in turn, break open cells.
- **Cryopulverization:** Cryopulverization is often employed for samples having a tough extracellular matrix, such as connective tissue, seed, and cartilage. In this technique, tissues are frozen using liquid nitrogen and then impact pulverization (typically, grinding, using a mortar and pestle or a powerful electric grinder) is performed. The powder so obtained is then suspended in the appropriate buffer.



Figure 8.1 - A sonicator in use. Wikipedia

Whatever method is employed to create a lysate, crude fractions obtained from it must be further processed via fractionation.

This page titled 8.1: Cell Lysis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.2: Fractionation and Chromatography Techniques

Fractionation of samples, as the name suggests, is a process of separating out the components or fractions of the lysate. Fractionation typically begins with centrifugation of the lysate. Using low-speed centrifugation, one can remove cell debris, leaving a supernatant containing the contents of the cell. By using successively higher centrifugation speeds (and resulting g forces) it is possible to separate out different cellular components, like nuclei, mitochondria, etc., from the cytoplasm. These may then be separately lysed to release molecules that are specific to the particular cellular compartment. The soluble fraction of any lysate can, then, be further separated into its constituents using various methods.



Figure 8.2.1: A high-speed centrifuge can be used to obtain different cell fractions from a crude lysate. Wikipedia

Column Chromatography

One powerful method used for this purpose is chromatography. We will consider several chromatographic approaches. Chromatography is used to separate out the components of a mixture based on differences in their size, charge or other characteristics. During chromatography, the mobile phase (buffer or other solvent) moves through the stationary phase (usually a solid matrix) carrying the components of the mixture. Separation of the components is achieved, because the different components move at different rates, for reasons that vary, depending on the type of chromatography used. We will consider several different kinds of chromatography to illustrate this process.

- Ion exchange chromatography
- Gel exclusion chromatography
- Affinity chromatography
- HPLC

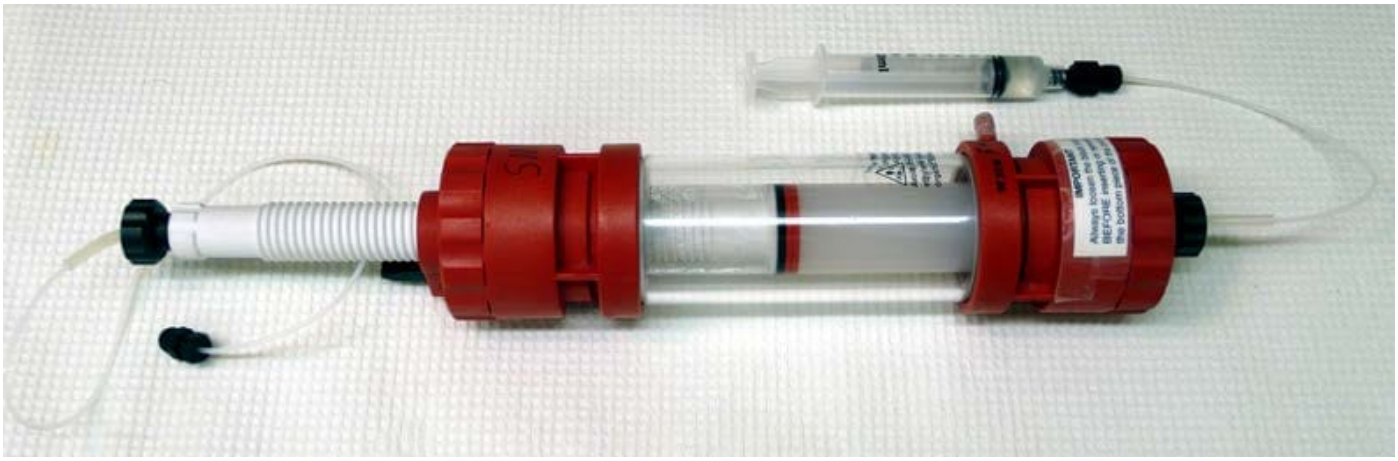


Figure 8.2.2: An ion exchange column apparatus.Wikipedia

These variations on chromatography are performed with the stationary phase held within so-called columns (Figure 8.2.1). These are tubes containing the stationary phase (also called the “support” or solid phase).

Supports are composed of tiny beads suspended in buffer (Figure 8.2.3) and are designed to exploit the chemistry or size differences of the components of the samples and thus provide a means of separation. Columns are “packed” or filled with the support, and a buffer or solvent carries the mixture of compounds to be separated through the support. Molecules in the sample interact differentially with the support and, consequently, travel through it at different speeds, thus enabling separation.



Figure 8.2.3: Ion exchange beads.Wikipedia

Ion exchange chromatography

In ion exchange chromatography, the support consists of tiny beads to which are attached chemicals possessing a charge. Before use, the beads are equilibrated in a solution containing an appropriate counter-ion to the charged molecule on the bead. Figure 8.2.5 shows the repeating unit of polystyrolsulfonate, a compound used as a cation exchange resin. As you can see, this molecule is negatively charged, and thus the beads would be equilibrated in a buffer containing a positively charged ion, say sodium. In the suspension, the negatively charged polystyrolsulfonate is unable to leave the beads, due to its covalent attachment, but the counter-ions (sodium) can be “exchanged” for molecules of the same charge.

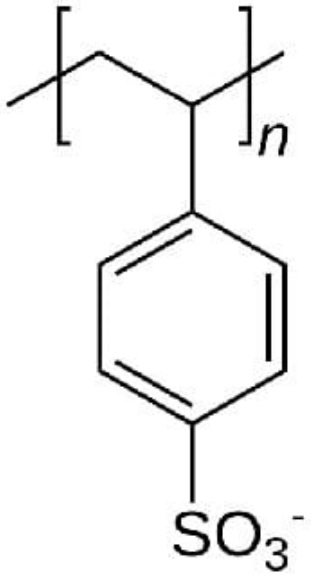


Figure 8.2.4: Polystyrolsulfonate, a cation exchange resin. Wikipedia

Exchanges

Thus, a cation exchange column will have positively charged counter-ions and negatively charged molecules covalently attached to the beads. Positively charged compounds from a cell lysate passed through the column will exchange with the counter-ions and "stick" to the negatively charged compounds covalently attached to the beads. Molecules in the sample that are neutral in charge or negatively charged will pass quickly through the column. At this point, only positively charged molecules from the original sample would be bound to the column. These may then be washed off, or eluted, by using buffers containing high concentrations of salt. Under these conditions, the interaction between the positively charged molecules and the polystyrosulfonate would be disrupted, allowing the molecules that were bound to the column to be recovered.

Anion exchange

On the other hand, in anion exchange chromatography, the chemicals attached to the beads are positively charged and the counterions are negatively charged (chloride, for example). Negatively charged molecules in the cell lysate 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 counter-ions to release them.

Uses

Ion exchange resins are useful for separating charged from uncharged, or oppositely charged, biomolecules in solution. The resins have a variety of other applications, including water purification and softening. Figure 8.2.5 shows use of a polystyrolsulfonate polymer in removing calcium for water softening.

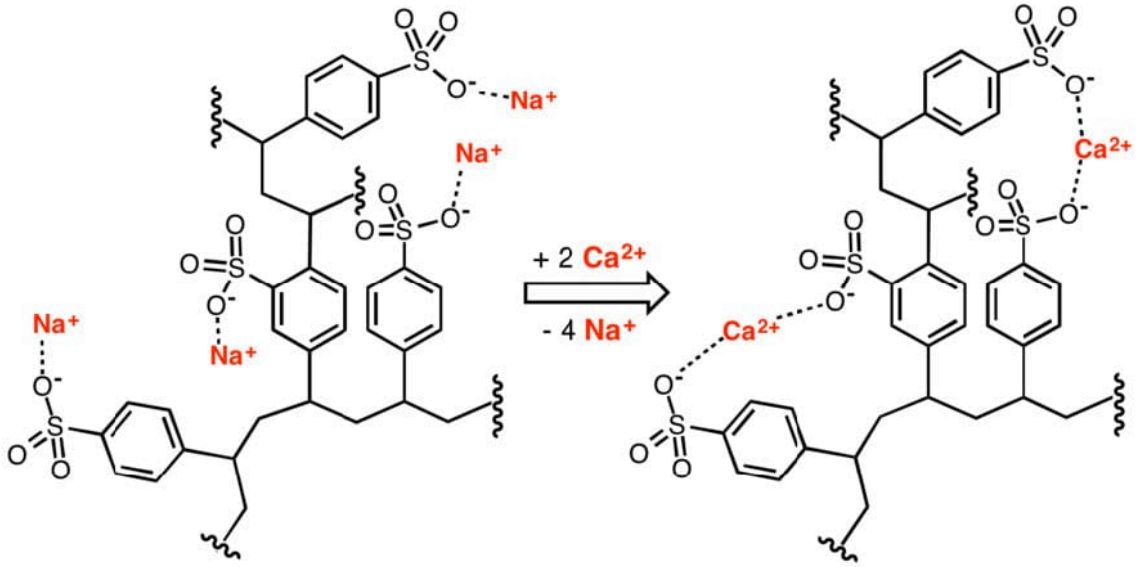


Figure 8.2.5*: Removal of calcium ions by an ion exchanger. Wikipedia*

Size Exclusion Chromatography

Size exclusion chromatography (also called molecular exclusion chromatography, gel exclusion chromatography, or gel filtration chromatography) is a low resolution separation method that employs beads with tiny “tunnels” in them that each have a precise opening. The size of the opening is referred to as an “exclusion limit,” which means that molecules above a certain molecular weight will not be able to pass through the tunnels. Molecules with physical sizes larger than the exclusion limit do not enter the tunnels and pass through the column relatively quickly, in the spaces outside 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 and elute last (Figure 8.2.6).

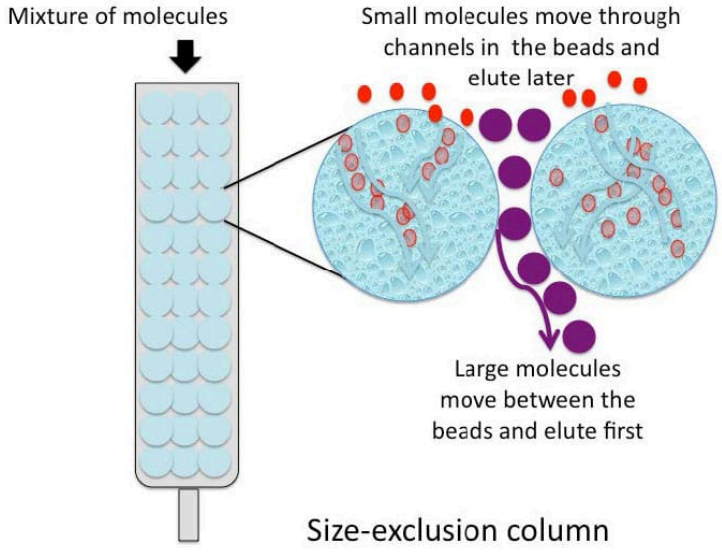


Figure 8.2.6: Separation of molecules by size in a size-exclusion (aka gel filtration) column

Figure 8.2.7 shows a profile of a group of proteins separated by size exclusion chromatography using beads with an exclusion limit of about 30,000 Daltons. Proteins 30,000 in molecular weight or larger elute in the void volume (left) while smaller proteins elute later (middle and right).

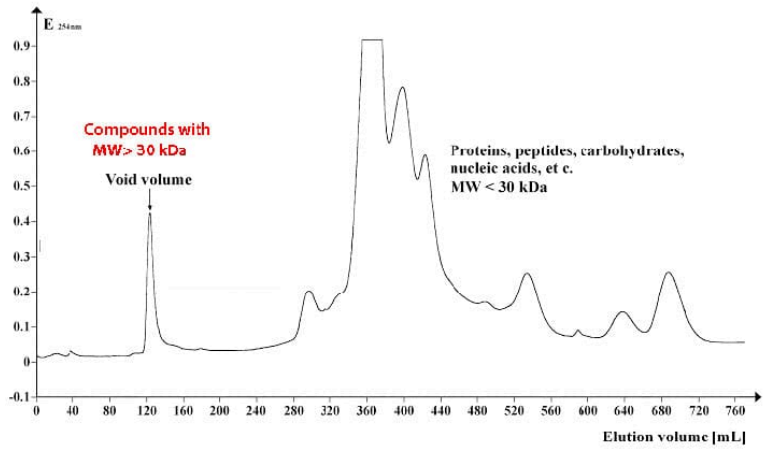


Figure 8.2.7: Result of a size exclusion separation. Wikipedia

Affinity chromatography

Affinity chromatography is a very powerful and selective technique that exploits the binding affinities of sample molecules (typically proteins) for molecules covalently linked to the support beads. In contrast to ion-exchange chromatography, where all molecules of a given charge would bind to the column, affinity chromatography exploits the specific binding of a protein or proteins to a ligand that is immobilized on the beads in the column.

For example, if one wanted to separate all of the proteins in a cell lysate that bind to ATP from proteins that do not bind ATP, one could use a column that has ATP attached to the support beads and pass the sample through the column. All proteins that bind ATP will “stick” to the column, whereas those that do not bind ATP will pass quickly through it. The bound proteins may then be released from the column by adding a solution of ATP that will displace the bound proteins by competing, for the proteins, with the ATP attached to the column matrix.

Histidine tagging

Histidine tagging (His-tagging) is a special kind of affinity chromatography and is a powerful tool for isolating a recombinant protein from a cell lysate. His-tagging relies on altering the DNA coding region for a protein to add a series of at least six histidine residues to the amino or carboxyl terminal of the encoded protein. This “His-Tag” is useful in purifying the tagged protein because histidine side chains can bind to nickel or cobalt ions. Separation of His-tagged proteins from a cell lysate is relatively easy (Figure 8.2.8).Passing the crude cell lysate through a column with nickel or cobalt attached to beads allows the His-tagged proteins to “stick,” while the remaining cell proteins all pass quickly through. The His-tagged proteins are then eluted by addition of imidazole to the column. Imidazole, which resembles the side chain of histidine, competes with the His-tagged proteins and displaces them from the column. Although non-tagged proteins in the lysate may also contain histidine as part of their sequence, they will not bind to the column as strongly as the His-tagged protein and will, thus, be displaced at lower imidazole concentrations than needed to elute the His-tagged protein. Surprisingly, many His-tagged proteins appear to function normally despite the added histidines, but if needed, the histidine tags may be cleaved from the purified protein by treatment with a protease that excises the added histidines, allowing the recovery of the desired protein with its native sequence.

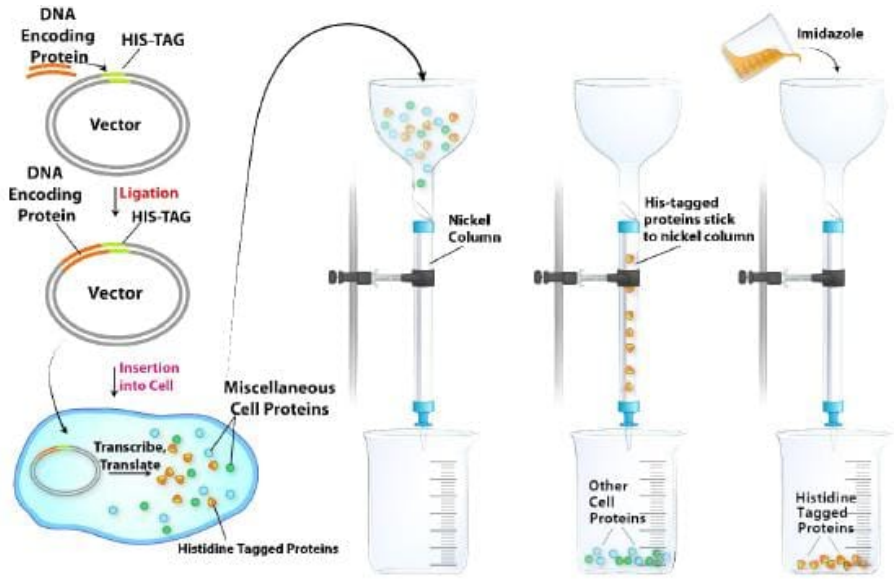


Figure 8.2.8*: Affinity chromatographic purification of a protein by histidine tagging.Image by Aleia Kim*

HPLC

High performance liquid chromatography (HPLC) is a powerful tool for separating a variety of molecules based on their differential polarities (Figure 8.2.9). A more efficient form of column chromatography, it employs columns with tightly packed supports and very tiny beads such that flow of solvents/buffers through the columns requires high pressures. The supports used may be polar (normal phase separation) or non-polar (reverse phase separation). In normal phase separations, non-polar molecules elute first followed by the more polar compounds. This order is switched in reverse phase chromatography. Of the two, reverse phase is much more commonly employed due to more reproducible chromatographic profiles (separations) that it typically produces.



Figure 8.2.9*: HPLC: Pumps on left/Column in center/Detector on the right. Wikipedia*

This page titled 8.2: Fractionation and Chromatography Techniques is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.3: Electrophoresis

Electrophoresis uses an electric field applied across a gel matrix to separate large molecules such as DNA, RNA, and proteins by charge and size. Samples are loaded into the wells of a gel matrix that can separate molecules by size and an electrical field is applied across the gel. This field causes negatively charged molecules to move towards the positive electrode. The gel matrix, itself, acts as a sieve, through which the smallest molecules pass rapidly, while longer molecules are slower-moving.

For DNA and RNA, sorting molecules by size in this way is trivial, because of the uniform negative charge on the phosphate backbone. For proteins, which vary in their charges, a clever trick must be employed to make them mimic nucleic acids - see polyacrylamide gel electrophoresis (PAGE) below. Different kinds of gels have different pore sizes. Like sieves with finer or coarser meshes, some gels do a better job of separating smaller molecules while others work better for larger ones. Gel electrophoresis may be used as a preparative technique (that is, when purifying proteins or nucleic acids), but most often it is used as an analytical tool.

Agarose Gel Electrophoresis

Agarose gel electrophoresis is a technique used to separate nucleic acids primarily by size. Agarose is a polysaccharide obtained from seaweeds (Figure 8.11). It can be dissolved in boiling buffer and poured into a tray, where it sets up as it cools (Figure 8.12) to form a slab. Agarose gels are poured with a comb in place to make wells into which DNA or RNA samples are placed after the gel has solidified. The gel is immersed in a buffer and a current is applied across the slab. Double-stranded DNA has a uniform negative charge that is independent of the sequence composition of the molecule. Therefore, if DNA fragments are placed in an electric field they will migrate from the cathode (-) towards the anode (+). The rate of migration is directly dependent on the ability of each DNA molecule to worm or wiggle its way through the sieving gel. The agarose matrix provides openings for macromolecules to move through. The largest macromolecules have the most difficult time navigating through the gel, whereas the smallest macromolecules slip through it the fastest.

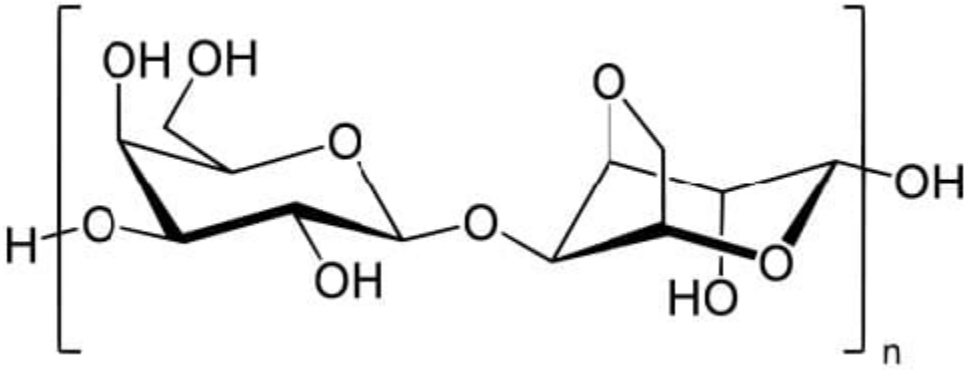


Figure 8.11 - Structure of the agarose polysaccharide. Wikipedia

Because electrophoresis uses an electric current as a force to drive the molecules through the matrix, the molecules being separated must be charged. Since the size to charge ratio for DNA and RNA is constant for all sizes of these nucleic acids, the molecules simply sort on the basis of their size - the smallest move fastest and the largest move slowest.

All fragments of a given size will migrate the same distance on the gel, forming the so-called “bands” on the gel. Visualization of the DNA fragments in the gel is made possible by addition of a dye, such as ethidium bromide, which intercalates between the bases and fluoresces when viewed under ultraviolet light (Figure 8.13) By running reference DNAs of known sizes alongside the samples, it is possible to determine the sizes of the DNA fragments in the sample. It is useful to note that, by convention, DNA fragments are not described by their molecular weights (unlike proteins), but by their length in base-pairs(bp) or kilobases (kb).

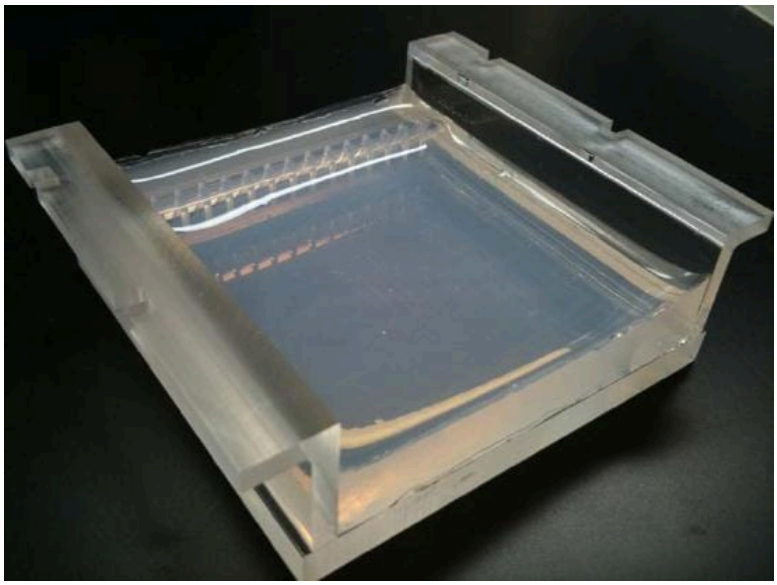


Figure 8.12 - Agarose gel electrophoresis separation of DNA - orange bands are DNA fragments. Wikipedia

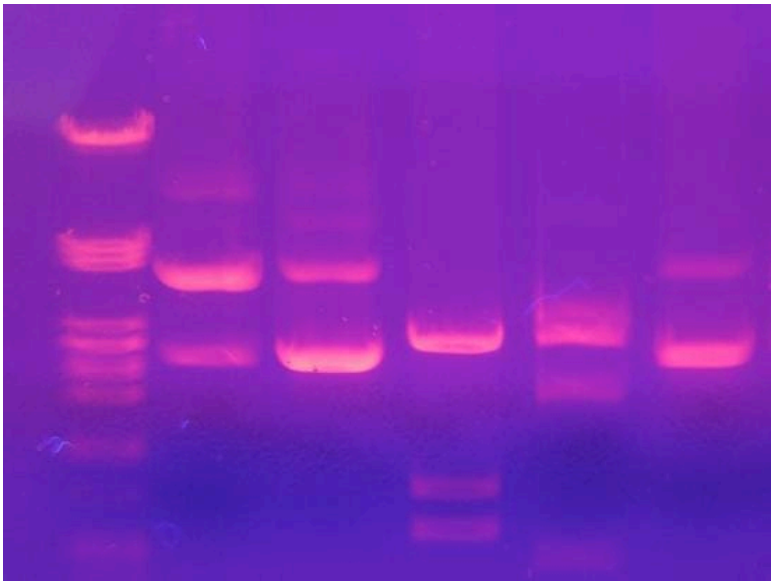


Figure 8.13 - DNA bands visualized with ethidium bromide staining. Wikipedia

Polyacrylamide gel electrophoresis (PAGE)

Like DNA and RNA, proteins are large macromolecules, but unlike nucleic acids, proteins are not necessarily negatively charged. The charge on each protein depends on its unique amino acid sequence. Thus, the proteins in a mixture will not necessarily all move towards the anode.

Additionally, whereas double-stranded DNA is rod-shaped, most proteins are globular (folded). Further, proteins are considerably smaller than nucleic acids, so the openings of the matrix of the agarose gel are simply too large to effectively provide separation. Consequently, unaltered (native) proteins are not very good prospects for electrophoresis on agarose gels. To separate proteins by mass using electrophoresis, one must make several modifications.

Gel matrix

First, a matrix made by polymerizing and cross-linking acrylamide units is employed. A monomeric acrylamide (Figure 8.14) is polymerized and the polymers are cross-linked using N,N'-Methylene-bisacrylamide (Figure 8.15) to create a mesh-like structure.

One can adjust the size of the openings of the matrix/mesh readily by changing the percentage of acrylamide in the reaction. Higher percentages of acrylamide give smaller openings and are more effective for separating smaller molecules, whereas lower percentages of acrylamide are used when resolving mixtures of larger molecules. (Note: polyacrylamide gels are also used to separate small nucleic acid fragments, with some acrylamide gels capable of separating pieces of DNA that differ in length by just one nucleotide.)

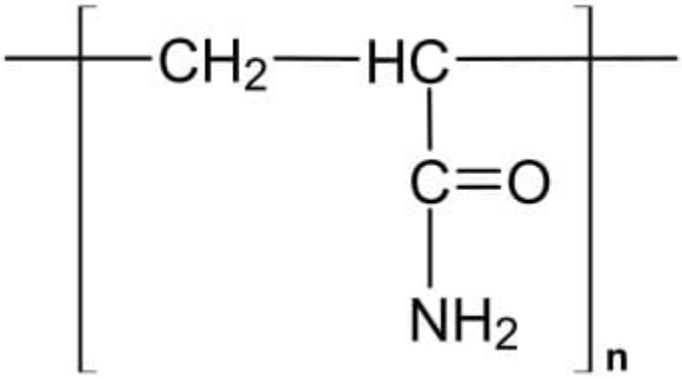


Figure 8.14 - Acrylamide monomer. Wikipedia

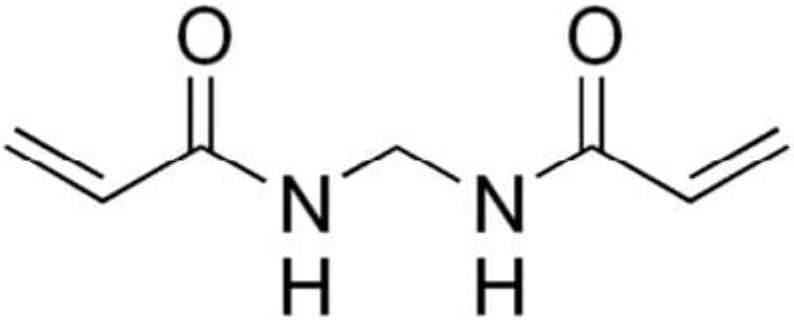


Figure 8.15 - N,N'-Methylenebisacrylamide - acrylamide crosslinking reagent. Wikipedia

Charge alteration by SDS

A second consideration is that proteins must be physically altered to "present" themselves to the matrix like the negatively charged rods of DNA. This is accomplished by treating the proteins with the anionic detergent, SDS (sodium dodecyl sulfate). SDS denatures the proteins so they assume a rod-like shape and the SDS molecules coat the proteins such that the exterior surface is loaded with negative charges, masking the original charges on the proteins and making the charge on the proteins more proportional to their mass, like the backbone of DNA.

Since proteins typically have disulfide bonds that prevent them from completely unfolding in detergent, samples are boiled with mercaptoethanol to break the disulfide bonds and ensure the proteins are as rod-like as possible in the SDS. Reagents like mercaptoethanol (and also dithiothreitol) are sulfhydryl-containing reagents that become oxidized as they reduce disulfide bonds in other molecules (see Figure 8.16)

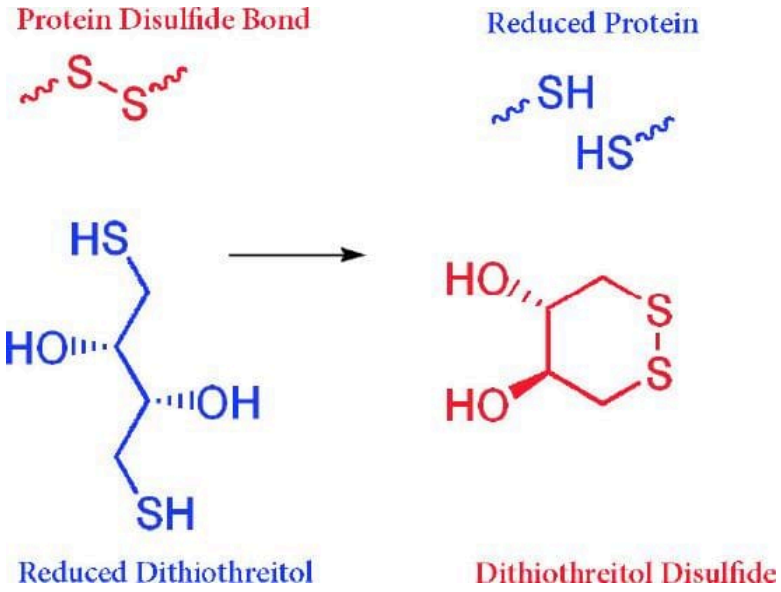


Figure 8.16 - Reduction of disulfide bonds by dithiothreitol. Wikipedia

Stacking Gel

A third consideration is that a “stacking gel” may be employed at the top of a polyacrylamide gel to provide a way of compressing the samples into a tight band before they enter the main polyacrylamide gel (called the resolving gel). Just like DNA fragments in agarose gel electrophoresis get sorted on the basis of size (largest move slowest and smallest move fastest), the proteins migrate through the gel matrix at velocities inversely related to their size. Upon completion of the electrophoresis, proteins may be visualized by staining with compounds that bind to proteins, like Coomassie Brilliant Blue (Figure 8.17) or silver nitrate.

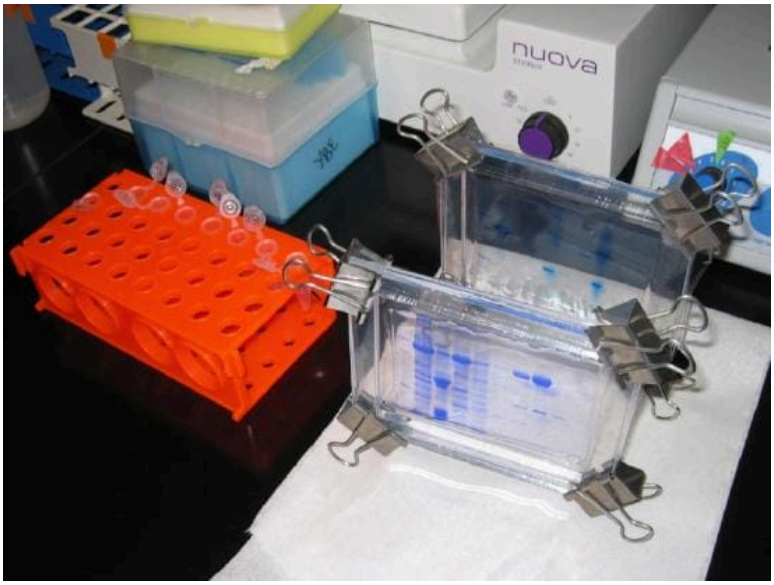


Figure 8.17 - Two SDS-PAGE gels - Proteins are the blue bands (stained with Coomassie Blue). Wikipedia

Non-denaturing gel electrophoresis

The SDS_PAGE technique described above is the commonest method used for electrophoretic separation of proteins. In some situations, however, proteins may be resolved on so-called “native” gels, in the absence of SDS. Under these conditions, the movement of proteins through the gel will be affected not simply by their mass, but by their charge at the pH of the gel, as well. Proteins complexed with other molecules may move as single entity, allowing the isolation of the binding partners of proteins of interest.

Isoelectric focusing

Proteins vary considerably in their charges and, consequently, in their pI values (pH at which their charge is zero). This can be exploited to separate proteins in a mixture. Separating proteins by isoelectric focusing requires establishment of a pH gradient in a tube containing an acrylamide gel matrix. The pore size of the gel is adjusted to be large, to reduce the effect of sieving based on size. Molecules to be separated are applied to the gel containing the pH gradient and an electric field is applied. Under these conditions, proteins will move according to their charge.

Positively charged molecules, for example, move towards the negative electrode, but since they are traveling through a pH gradient, as they pass through it, they reach a region where their charge is zero and, at that point, they stop moving. They are at that point attracted to neither the positive nor the negative electrode and are thus “focused” at their pI (Figure 8.18). Using isoelectric focusing, it is possible to separate proteins whose pI values differ by as little as 0.01 units.

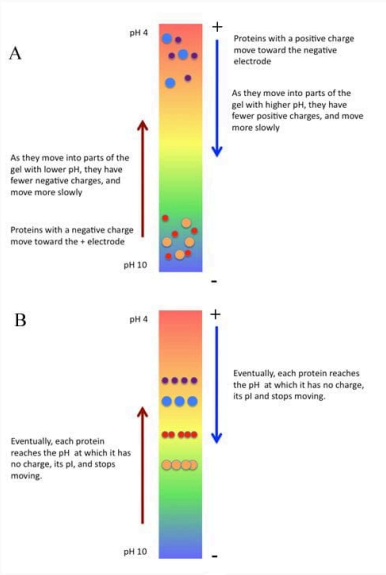


Figure 8.18 - Isoelectric focusing: A. At the start of the run; B. at the end of the run

2D gel electrophoresis

Both SDS-PAGE and isoelectric focusing are powerful techniques, but a clever combination of the two is a powerful tool of proteomics - the science of studying all of the proteins of a cell/tissue simultaneously. In 2-D gel electrophoresis, a lysate is first prepared from the cells of interest. The proteins in the lysate are separated first by their pI, through isoelectric focusing and then by size by SDS-PAGE.

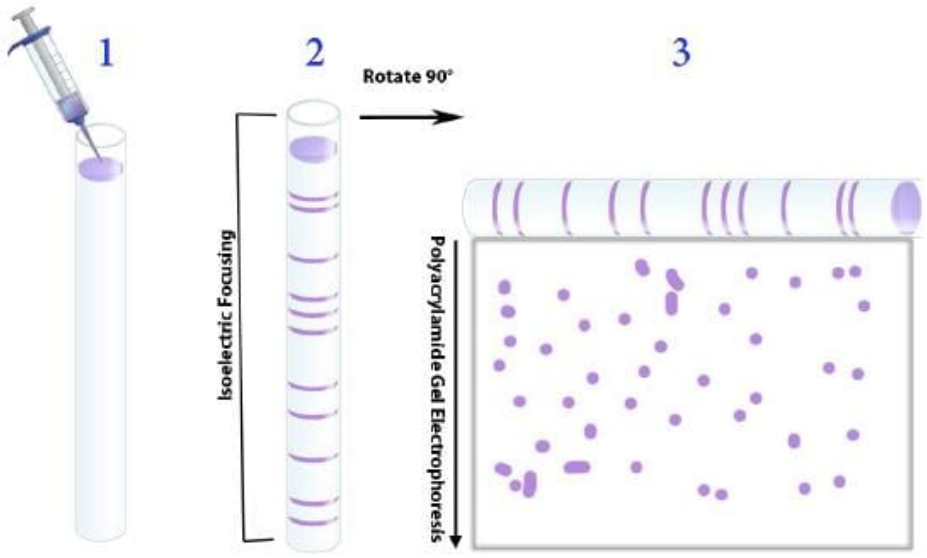


Figure 8.19 - Scheme for performing 2-D gel analysis. Image by Aleia Kim

The mixture of proteins is first applied to a tube or strip (Figure 8.19, Step 1) where isoelectric focusing is performed to separate the proteins by their pI values (Step 2). Next, as shown in the figure, the gel containing the proteins separated by their pIs is turned on its side and applied along the top of a polyacrylamide slab for SDS-PAGE to separate on the basis of size (Step 3). The proteins in the isoelectric focusing matrix are electrophoresed into the polyacrylamide gel and separated on the basis of size. The product of this analysis is a 2-D gel as shown in Figure 8.20.The power of 2-D gel electrophoresis is that virtually every protein in a cell can be separated and appear on the gel as a spot defined by its unique size and pI. In the figure, spots in the upper left correspond to large positively charged proteins, whereas those in the lower right are small negatively charged ones. Every spot on a 2-D gel can be eluted and identified by using high throughput mass spectrometry. This is particularly powerful when one compares protein profiles between different tissues or between control and treated samples of the same tissue.

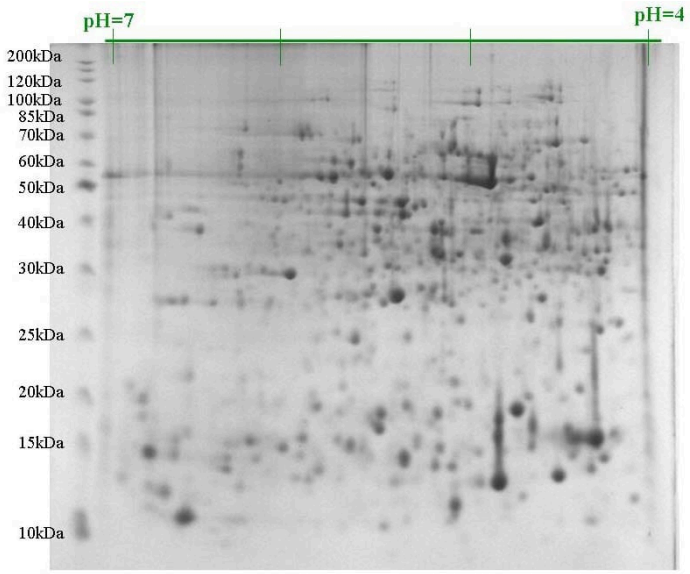


Figure 8.20 - Result of 2-D gel electrophoresis separation. Wikipedia

Protein profiles comparison

Comparison of 2-D gels of proteins from non-cancerous tissue and proteins from a cancerous tissue of the same type provides a quick identification of proteins whose level of expression differs between the two. Information such as this can be useful in designing treatments or in understanding the mechanism(s) by which the cancer develops.

This page titled 8.3: Electrophoresis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.4: Detection, identification and quantitation of specific nucleic acids and proteins

While gel electrophoresis can be used to resolve molecules in a mixture, by itself, the technique does not permit the detection and identification of specific nucleic acid sequences or proteins. For example, the 2-D gel shown above clearly separates a large number of proteins in a sample into individual spots. However, if we wanted to know whether a specific protein was present, we could not tell by simply looking at the gel. Likewise, in an agarose gel, while bands of DNA could be assigned a size, one could not distinguish between two DNAs of different sequence if they were both the same length in base-pairs. One way to detect the presence of a particular nucleic acid or protein is dependent on transferring the separated molecules from the gels onto a membrane made of nitrocellulose or nylon to create a "blot" and probing for the molecule(s) of interest using reagents that specifically bind to those molecules. The next section will discuss how this can be done for nucleic acids as well as for proteins.

Southern and Northern Blots

The Southern blot is named for its inventor, Oxford professor, Edwin Southern, who came up with a protocol for transferring DNA fragments from a gel onto a nitrocellulose sheet and detecting a specific DNA sequence among the bands on the blot. As shown in Figure 8.21, the method works as follows. A mixture of DNA molecules (often DNA that has been cut into smaller fragments using restriction endonucleases) is loaded on an agarose gel, as usual. After the gel run is complete, the DNA bands are transferred from the gel onto a membrane. This can be achieved by capillary transfer, where the gel is placed in contact with a piece of membrane and buffer is pulled through the gel by wicking it up into a stack of absorbent paper placed above the membrane. As the buffer moves, it carries with it the DNA fragments. The DNA binds to the membrane leaving a "print" of DNA fragments that exactly mirrors their positions in the gel. The blotting membrane may be treated with UV light, heat, or chemicals to firmly attach the DNA to the membrane.

Next, a probe, or visualizing agent specific for the molecule of interest is added to the membrane. In Figure 8.21, this is called a labeled probe. The probes in a Southern blot are pieces of DNA designed to be complementary to the desired target sequence. If the sequence of interest is present on the blot, the probe, which is complementary to it, can base-pair (hybridize) with it. The blot is then washed to remove all unbound probe. Probes are labeled with radioactivity or with other chemical reagents that allow them to be easily detected when bound to the blot, so it is possible to visually determine whether the probe has bound to any of the DNA bands on the blot. Given that the Southern blot relies on specific base-pairing between the probe and the target sequence, it is easy to adapt the technique to detect specific RNA molecules, as well. The modification of this method to detect RNAs was jokingly named a "northern" blot.

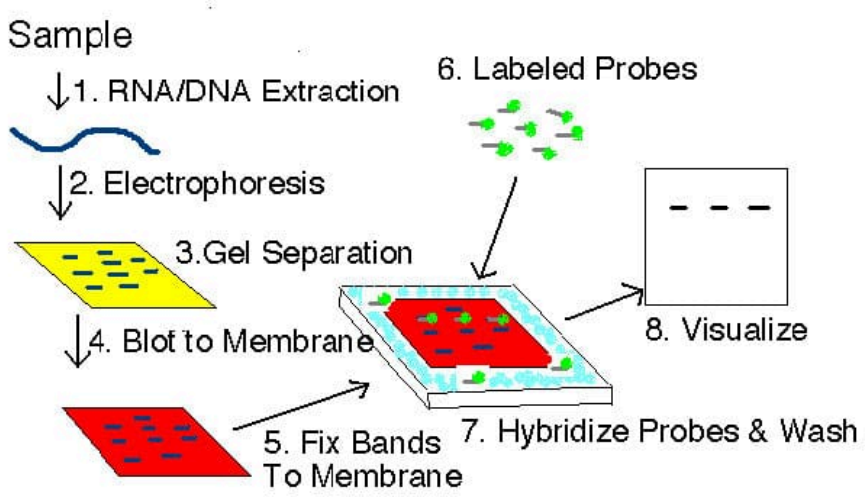


Figure 8.21 - Northern or Southern blot scheme. Southern blotting adds strand denaturation between steps 4 and 5. Wikipedia

Western Blots

Proteins cannot, for obvious reasons, be detected through base-pairing with a DNA probe, but protein blots, made by transferring proteins, separated on a gel, onto a membrane, can be probed using specific antibodies against a particular protein of interest. Protein detection usually employs two antibodies, the first of which is not labeled. The label is on the second antibody, which is designed to recognize only the first antibody in a piggyback fashion. The first antibody specifically binds to the protein of interest on the blot and the second antibody recognizes and binds the first antibody. The second antibody commonly carries an enzyme or reagent which can cause a reaction to produce a color upon further treatment. In the end, if the molecule of interest is in the original mixture, it will "light" up and reveal itself on the blot. This variation on the blotting theme was dubbed a western blot (Figure 8.22).

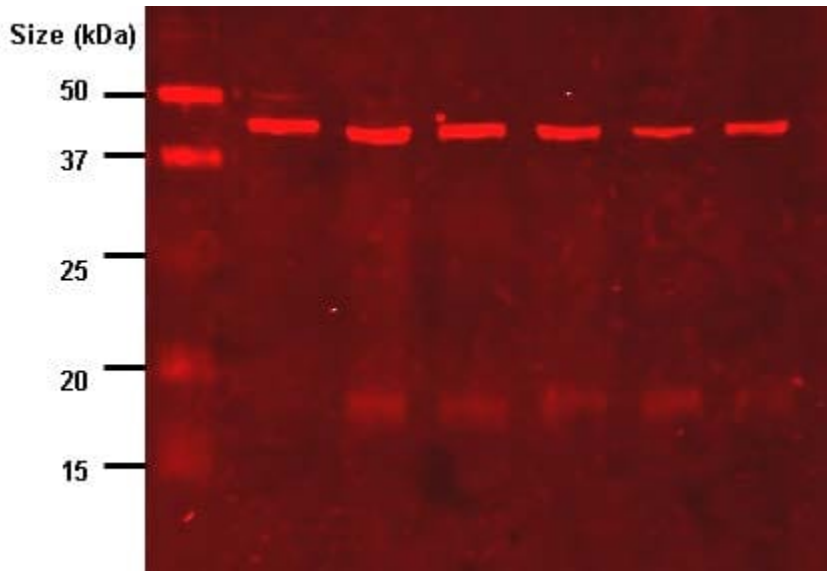


Figure 8.22 - Result of a western blot analysis. Wikipedia

In each of the blots described above, binding of the probe to the target molecule allows one to determine whether the target sequence or protein was in the sample. Although blots are designed to be used for detection, rather than for precise quantitation, it is possible to obtain estimates of the abundance of the target molecule from densitometry measurements of signal intensity.

Microarrays

2-D gels are a way of surveying a broad spectrum of protein molecules simultaneously. One approach to doing something similar with DNA or RNA involves what are called microarrays. Microarrays are especially useful for monitoring the expressions of thousands of genes, simultaneously. Where a northern blot would allow the identification of a single mRNA from a mixture of mRNAs, a microarray experiment can allow the simultaneous identification of thousands of mRNAs that may be made by a cell at a given time. It is also possible to perform quantitation much more reliably than with a blot.

Microarrays employ a glass slide, or chip, to which are attached short sequences of single-stranded DNA, arranged in a grid, or matrix (Figure 8.23) Each position in the grid corresponds to a unique gene. That is, the DNA sequence at this spot is part of the sequence of a specific gene. Each spot on the grid has multiple identical copies of the same sequence. The gene sequence immobilized at each position in the grid is recorded.

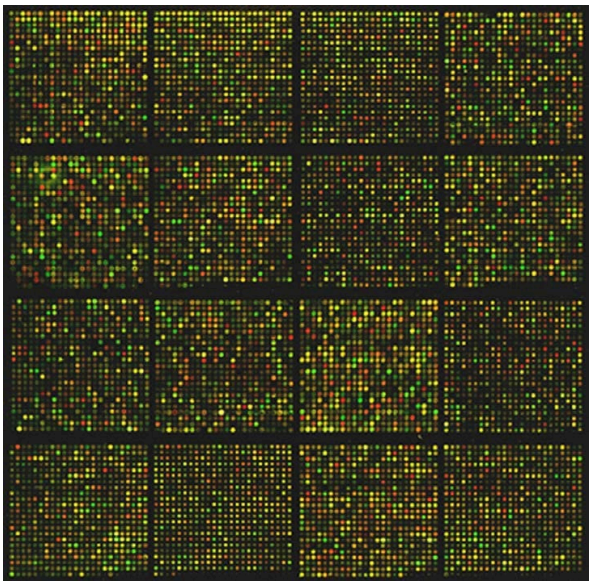


Figure 8.23 - Microarray design. Image by Taralyn Tan

To the slide are added a mixture of sample molecules, some of which will recognize and bind specifically to the sequences on the slide. Binding between the sample molecules and the sequences attached to the slide occurs by base pairing, in the case of DNA microarrays. The slide is then washed to remove sample molecules that are not specifically bound to the sequences in the grid.

Sample molecules are tagged with a fluorescent dye, allowing the spots where they bind to be identified. The grid is analyzed spot by spot for binding of the sample molecules to the immobilized sequences. The more sample molecules are bound at a spot, the greater the intensity of dye fluorescence that will be observed. Information from this analysis can give information about the presence/absence/abundance of molecules in the sample that bind to the sequences in the grid.

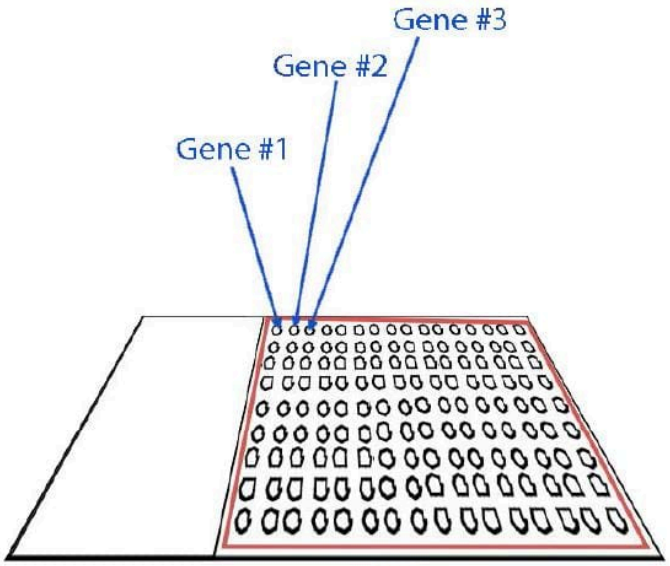


Figure 8.24 - Large scale microarray analysis of mouse transcriptome. Wikipedia

This page titled 8.4: Detection, identification and quantitation of specific nucleic acids and proteins is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.5: Transcriptomics

Consider a matrix containing all of the known gene sequences in a genome. To make such a matrix for analysis, one would need to make copies of every gene, either by chemical synthesis or by using the polymerase chain reaction. The strands of the resulting DNAs would then be separated to obtain single-stranded sequences that could be attached to the chip. Each box of the grid would contain sequence from one gene. With this grid, one could analyze the transcriptome - all of the mRNAs being made in selected cells at a given time. For a simple analysis, one could take a tissue (say liver) and extract all the mRNAs from it. This mRNA population represents all the genes that were being expressed in the liver cells at the time the RNA was extracted. These RNAs should be able to hybridize (base-pair) with their corresponding genes on the microarray. Genes that were not being expressed would have no mRNAs to bind to their corresponding genes on the grid.

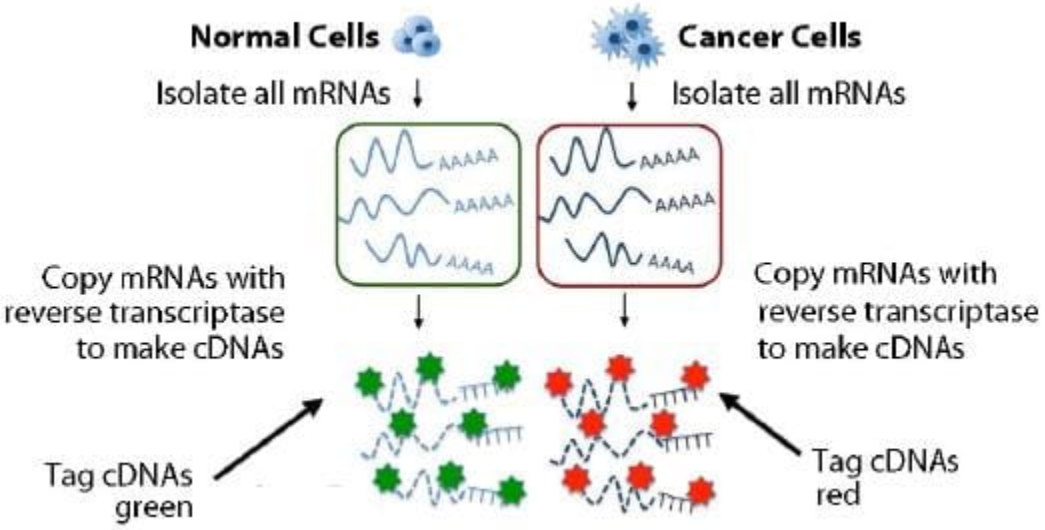


Figure 8.25 - Copying and labeling of transcriptome. Image by Taralyn Tan

In practice, the mRNAs are not used directly, but are copied into single-stranded DNA copies called cDNAs. The cDNAs are tagged with a fluorescent dye and added to the microarray under conditions that allow base pairing so that the cDNAs can find and base pair with complementary sequences on the matrix (Figure 8.26). The matrix is then washed to remove unhybridized cDNAs. The presence/absence/abundance of each mRNA is then readily determined by measuring the amount of dye at each box of the grid.

Figure 8.26 - Add labeled cDNAs to microarray plate. Image by Taralyn Tan

In Figure 8.27, a fluorescent cDNA has bound to the spot on the far right in the third row of the grid. This means that the sequence of the cDNA was complementary to the sequence of the gene sequence immobilized at that spot. Because the identity of the genes at each position on the grid is known, we then know that the sample contained mRNA that corresponded to that particular gene. In other words, that gene was being expressed in the cells from which the mRNAs were obtained.

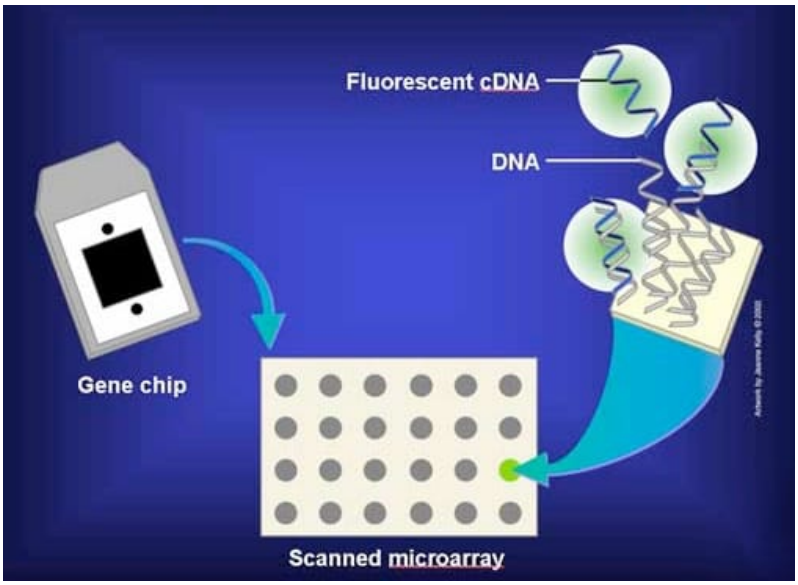


Figure 8.27 - Binding of a fluorescent cDNA copy of a specific mRNA to DNA immobilized on one spot in a microarray. Wikipedia

A more powerful analysis could be performed with two sets of mRNAs simultaneously. . One set of cDNAs could come from a cancerous tissue and the other from a non-cancerous tissue, for example. The cDNAs derived from each sample is marked with a different color (say green for normal and red for cancerous) (Figure 8.25). The cDNAs are mixed and then added to the matrix and complementary sequences are once again allowed to form duplexes (Figure 8.27).

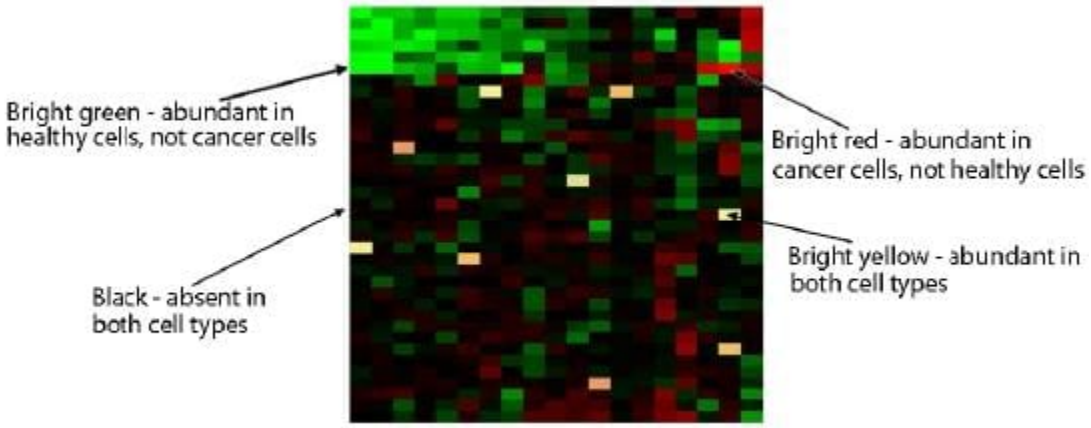


Figure 8.28 - Microarray analysis comparing gene expression in normal and cancer cells. Wikipedia

Unhybridized cDNAs are washed away and then the plate is analyzed. Red grid boxes correspond to an mRNA present in the cancerous tissue, but not in the non-cancerous tissue. Green grid boxes correspond to an mRNA present in the non-cancerous tissue, but not in the cancerous tissue. Yellow would correspond to mRNAs present in equal abundance in the two tissues (Figure 8.28). The intensity of each spot also gives information about the relative amounts of each mRNA in each tissue.

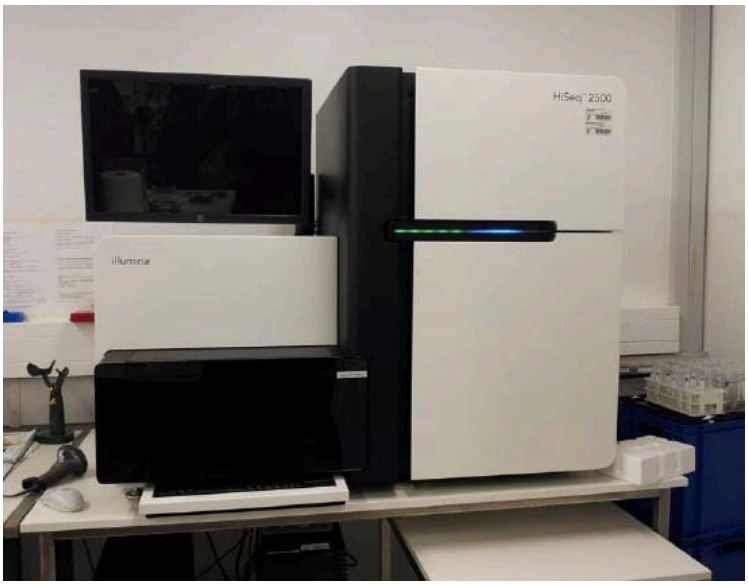


Figure 8.29 - Automated high throughput sequencer. Wikipedia

The same principle used for nucleic acid microarrays can be adapted for analyzing other molecules. For example, polypeptides could be bonded to the glass slide instead of DNA to create a protein chip. Protein chips are useful for studying the interactions of proteins with other molecules as well as for diagnostics.

RNA-Seq Technique

Like microarrays, a newer method called RNA-Seq, is a tool for simultaneously detecting and quantitating all of the transcripts in a given sample. This method relies on recently developed sequencing technologies called next-generation sequencing, or deep sequencing. These techniques allow for rapid, parallel sequencing of millions of DNA fragments and can, thus, be used not only for genomic DNA, but also to sequence all of the reverse-transcribed RNAs from a given sample.

To determine all the protein-coding genes that were being expressed in a particular set of cells under specific physiological conditions, all of the mRNA would first be extracted and reverse-transcribed into cDNA. This step is similar to the preparation of samples for microarrays. However, at this point, the cDNAs are fragmented into smaller pieces, and have small sequencing adapters attached at either end. The fragments are then subjected to high-throughput sequencing, to obtain short sequences from all of the fragments. These data are aligned against the genome sequence and used to measure the level of expression of different genes. RNA-Seq offers some advantages over microarrays. With microarrays, an RNA can only be detected if the gene sequence corresponding to it is present on the grid. In RNA-Seq every RNA present in the sample is sequenced, so detection of RNAs is not limited by the probes on a chip. RNA-Seq is more sensitive than microarrays and offers a much larger range over which gene expression can be measured accurately.

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

8.6: Isolating Genes

Earlier in this chapter, we discussed methods such as column chromatography that are used to purify proteins of interest. Using combinations of these methods, it is possible to isolate a protein to a high degree of purity, thus enabling us to study the protein's activity and properties. This problem is harder to solve for nucleic acids. Genomic DNA can be readily obtained from cells, but is too complex to be analyzed as a whole. Individual genes are the units of DNA that correspond to proteins, and thus, it makes more sense to isolate specific genes for study. Methods to isolate genes were not available till the 1970s, when the discovery of restriction enzymes and the invention of molecular cloning provided, for the first time, ways to obtain large quantities of specific DNA fragments, for study. Although, for purposes of obtaining large amounts of a specific DNA fragment, molecular cloning has been largely replaced by direct amplification using the polymerase chain reaction described later, cloned DNAs are still very useful for a variety of reasons. The development of molecular cloning was dependent on the discovery of restriction endonucleases, described below.

Restriction enzymes

Restriction enzymes, or restriction endonucleases, are enzymes made by bacteria. These enzymes protect bacteria by degrading foreign DNA molecules that are carried into their cells by, for example, an invading bacteriophage. Each restriction enzyme recognizes a specific sequence, usually of four or six nucleotides in the DNA. These sequences, when they occur in the bacterium's own DNA, are chemically modified by methylation, so that they are not recognized and degraded. Where these sequences occur in foreign DNA, they are cut by the restriction enzyme.

The utility and importance of restriction enzymes lies in their ability to recognize specific sequences in DNA and cut near or (usually) at the site they recognize. Over 3000 such enzymes are known. Sequences recognized by these enzymes are typically 4-8 base pairs long and the most commonly used enzymes recognize sequences described as palindromic.

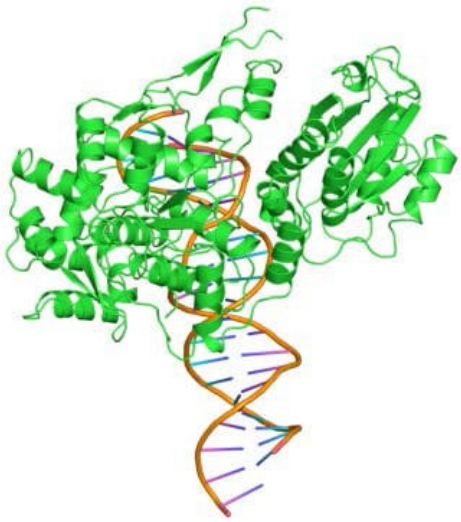


Figure 8.6.1***:*** *-A restriction enzyme bound to its recognition sequence on DNA. Wikipedia*

Palindrome

In molecular biology, the term palindrome means that the sequence of the recognition site when read in the 5‘ to 3‘ direction for the top strand is exactly the same as that of the bottom strand. Consider the sequence recognized by the restriction enzyme known as Hind III (pronounced hin-dee-three). It is

5’ -A-A-G-C-T-T-3’
3’ -T-T-C-G-A-A-5’

On the top strand, the recognition sequence is

5’ AAGCTT 3’

which is the same as the bottom strand (read in the same 5’ to 3’ direction).

While all restriction enzymes must recognize and bind to particular DNA sequences, the exact spot at which they cut the DNA varies. Some enzymes leave a staggered sequence after cutting that has an overhang at the 5’ end of one strand of the duplex; some leave a staggered sequence after cutting that has an overhang at the 3’ end; and some cut both strands in the same place, leaving no overhanging sequence - called blunt end cutters.

Consider cutting a DNA sequence that contains the Hind III recognition site, which is

5’ -A-A-G-C-T-T-3’
3’ -T-T-C-G-A-A-5’

Embedded within a DNA sequence, the Hind III sequence would look like this (Ns correspond to any base and represent all of the DNA around the recognition site).

5’ -N-N-N-A-A-G-C-T-T-N-N-N-3’
3’ -N-N-N-T-T-C-G-A-A-N-N-N-5’

After cutting with Hind III, it would look as follows:

5’ -N-N-N-A 3‘ 5’A-G-C-T-T-N-N-N-N-3’
3’ -N-N-N-T-T-C-G-A-5‘ 3’ A-N-N-N-N-5’

where gaps have been inserted to illustrate where cutting has occurred. Hind III cuts between the two ‘A’ containing nucleotides near the 5’ end of the recognition sequence and thus leaves 5’ overhangs (Figure 8.6.2).

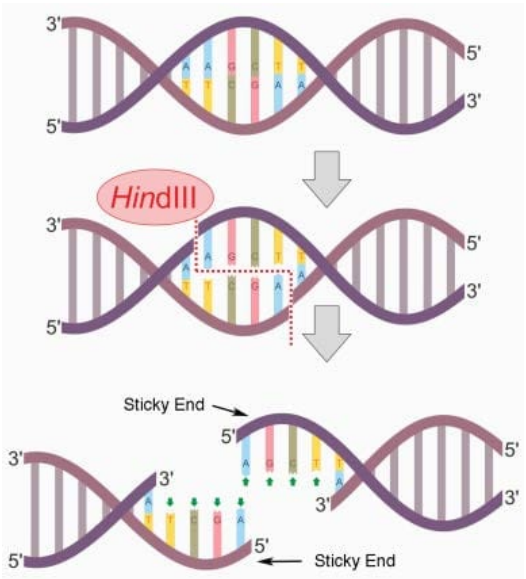


Figure 8.6.2***:*** *Result of cutting DNA with Hind III. Wikipedia*

The restriction enzyme Pst I, on the other hand, recognizes the following sequence

5’ -N-N-N-C-T-G-C-A-G-N-N-N-N-3’
3’ -N-N-N-G-A-C-G-T-C-N-N-N-N-5’

and cuts between the A and the G near the 3’ end of the recognition sequence.

5’ -N-N-N-C-T-G-C-A 3‘ 5’G-N-N-N-N 3’
3’ -N-N-N-G 5‘ 3’ A-C-G-T-C-N-N-N-N 5’

As you can see, cutting a DNA with Pst I leaves 3' overhangs of the recognition sequence. The ends left after cutting by a restriction enzyme that overhang either at the 5' end or the 3' end are referred to as being "sticky" because they can form proper base pairs and more readily be joined to a similarly "sticky end". This means that you can take two unrelated pieces of DNA, cut them with the same restriction enzyme so that they have compatible sticky ends, and then "paste" them together using DNA ligase to form a new hybrid molecule, or recombinant.

Making Recombinant DNAs

Joining together of DNA fragments from different sources creates recombinant DNA. The ability to cut and paste DNA might seem like purely a technical feat, but one key application that arose out of this is molecular cloning. In molecular cloning a gene of interest can be inserted into a vector, usually a plasmid, by cutting both the vector and the gene (called the insert) with the same enzyme to generate sticky ends and joining the two pieces together to generate a recombinant (Figure 8.6.3). A plasmid is a type of autonomously replicating, extrachromosomal DNA. It is quite simple to extract plasmids from the cells, engineer them to contain the gene of interest and re-introduce the recombinant plasmid into the bacteria. The idea was that when the plasmid DNA was replicated, the extra inserted gene would also be copied. Thus, by growing up a lot of the bacteria carrying the plasmid, many copies of the gene of interest could be obtained, to provide sufficient amounts of the gene to use in experiments. While we now have easier methods to accomplish this goal, cloned DNAs remain very useful. For example, it is possible to clone a gene that encodes a protein of interest so that it can be expressed at high levels in the cells into which the recombinant plasmid is introduced.

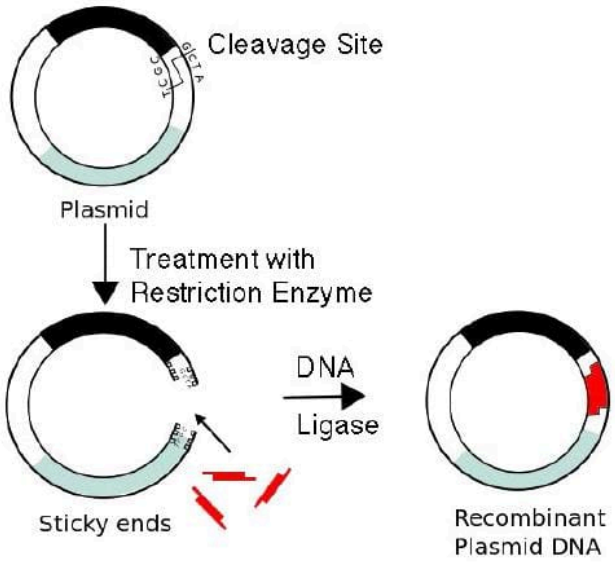


Figure 8.6.3*: Recombinant DNA construction. Wikipedia*

Whatever the purpose for which the recombinant plasmid is made, it typically carries an antibiotic resistance gene (or genes), called a selectable marker. Cells that take up the plasmid will be able to grow in the presence of the antibiotic. If bacterial cells to which the plasmid has been added are plated on agar containing the antibiotic, the cells which took up the plasmid will be able to grow, while the others will not.

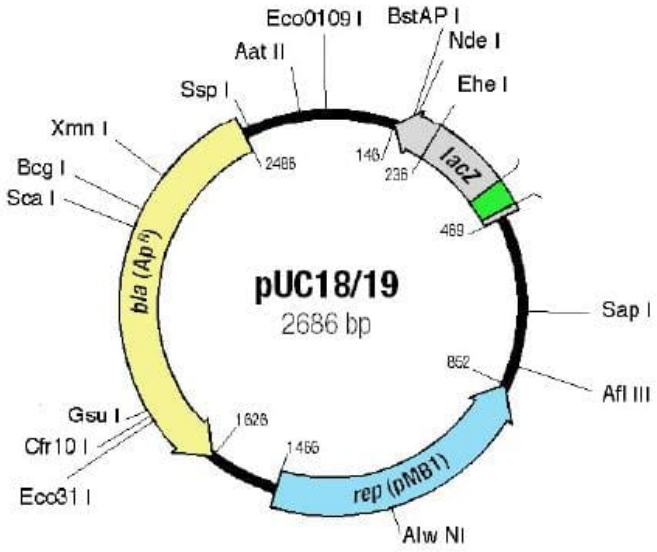


Figure 8.6.4: Restriction site map for the pUC 18/19 plasmids, a classic plasmid vector. Genes identified by arrows. Numbers correspond to the pUC 18/19 numbering convention

Expression cloning

As mentioned above, a gene of interest may be inserted into a vector and the recombinant plasmid be placed into a cell where the gene can be expressed. For instance, one might desire to clone the gene coding for human growth hormone or insulin or other medically important proteins and have a bacterium or yeast make large quantities of it very cheaply. Remember that these are human proteins, and thus it is not feasible to extract the proteins in any quantity from human subjects.

To clone a gene so that it can be expressed, one needs to set up the proper conditions in order for the human protein to be made in the bacterial cells. This typically involves the use of specially designed plasmids. These plasmids have been engineered to 1) replicate in high numbers; 2) carry markers that allow researchers to identify cells carrying them (antibiotic resistance, for example) and 3) contain sequences (such as a promoter and Shine Dalgarno sequence) necessary for expression of the desired protein, with convenient sites for insertion of the gene of interest in the appropriate place relative to the control sequences. A plasmid which has all of these features is referred to as an expression vector. In addition to plasmids that can be used for expression in bacterial cells, expression vectors are also available that allow protein expression in a variety of eukaryotic cells.

Many sophisticated variations on such vectors have been created that have made it easy to produce and purify large amounts of any protein of interest for which the gene has been cloned. A handy feature in some expression vectors is a sequence encoding an affinity tag either up- or downstream of the gene being expressed. This sequence allows a short affinity tag (such as a run of histidine residues) to be fused onto the encoded protein. The tag can be used to readily purify the protein, as described in the section on affinity chromatography.

This page titled 8.6: Isolating Genes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.7: Polymerase Chain Reaction (PCR)

Molecular cloning was the first method available to isolate a gene of interest and make many copies of it to obtain sufficient amounts of the DNA to study. Today, there is a faster and easier way to obtain large amounts of a DNA sequence of interest -the polymerase chain reaction (PCR). PCR allows one to use the power of DNA replication to amplify DNA enormously in a short period of time. As you know, cells replicate their DNA before they divide, and in doing so, double the amount of the cell's DNA. PCR essentially mimics cellular DNA replication in the test tube, repeatedly copying the target DNA over and over, to produce large quantities of the desired DNA.

Selective Replication

In contrast to cellular DNA replication, which amplifies all of a cell's DNA during a replication cycle, PCR does targeted amplification to replicate only a segment of DNA bounded by the two primers that determine where DNA polymerase begins replication. Figure 8.34 illustrates the process. Each cycle of PCR involves three steps, denaturing, annealing and extension, each of which occurs at a different temperature.

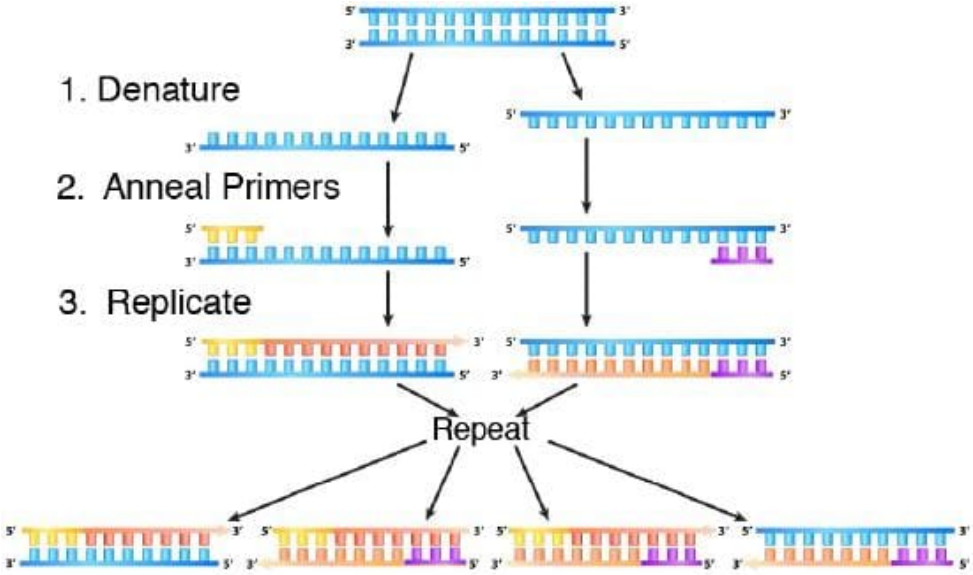


Figure 8.34 - Steps in the polymerase chain reaction. Image by Aleia Kim

The Starting Materials

Since PCR is, basically, replication of DNA in a test-tube, all the usual ingredients needed for DNA replication are required:

- A template (the DNA containing the target sequence that is being copied)
- Primers (to initiate the synthesis of the new DNA strands)
- Thermostable DNA polymerase (to carry out the synthesis). The polymerase needs to be heat stable, because heat is used to separate the template DNA strands in each cycle.
- dNTPs (DNA nucleotides to build the new DNA strands).
- The template is the DNA that contains the target you want to amplify (the "target" is the specific region of the DNA you want to amplify).

The primers are short synthetic single-stranded DNA molecules whose sequence matches a region flanking the target sequence. It is possible to chemically synthesize DNA molecules of any given base sequence, to use as primers. To make primers of the correct sequence that will bind to the template DNA, it is necessary to know a little bit of the template sequence on either side of the region of DNA to be amplified. DNA polymerases and dNTPs are commercially available from biotechnology supply companies.

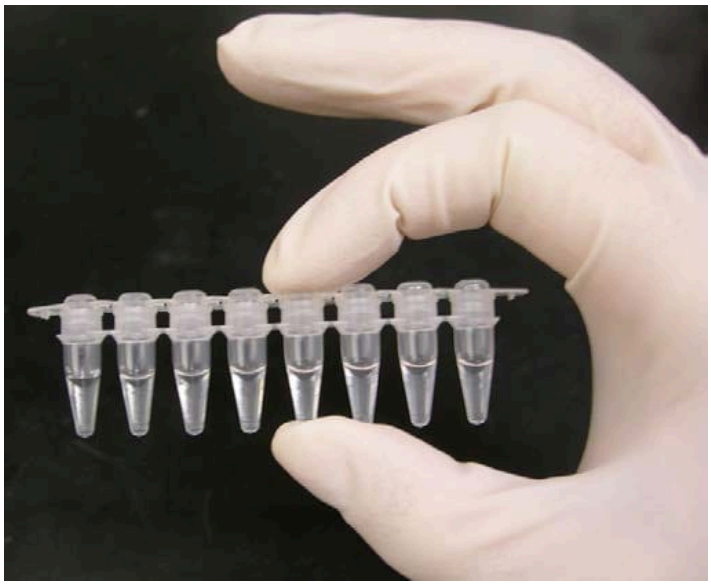


Figure 8.35 - PCR tubes with DNA samples ready for reaction. Wikipedia

First, all of the reagents are mixed together. Primers are present in millions of fold excess over the template. This is important because each newly made DNA strand starts from a primer. The first step of the process involves separating the strands of the target DNA by heating to near boiling.

Next, the solution is cooled to a temperature that favors complementary DNA sequences finding each other and making base pairs, a process called annealing. Since the primers are present in great excess, the complementary sequences they target are readily found and base-paired to the primers. These primers direct the synthesis of DNA. Only where a primer anneals to a DNA strand will replication occur, since DNA polymerases require a primer to begin synthesis of a new strand.



Figure 8.36 - A PCR thermocycler system. Wikipedia

Extension

In the third step in the process, the DNA polymerase replicates DNA by extension from the 3' end of the primer, making a new DNA strand. At the end of the first cycle, there are twice as many DNA molecules, just as in cellular replication. But in PCR, the process is repeated, usually for between 25 and 30 cycles. At the end of the process, there is a theoretical yield of 2^{30} (over 1 billion times) more DNA than there was to start. (This enormous amplification power is the reason that PCR is so useful for forensic investigations, where very tiny amounts of DNA may be available at a crime scene.)

The temperature cycles are controlled in a thermocycler, which repeatedly raises and lowers temperatures according to the set program. Since each cycle can be completed in a couple of minutes, the entire amplification can be completed very rapidly. The resulting DNA is analyzed on a gel to ensure that it is of the expected size, and depending on what it is to be used for, may also be sequenced, to be certain that it is the desired fragment.

Mutagenesis

PCR is frequently used to obtain gene sequences to be cloned into vectors for protein expression, for example. Besides simplicity and speed, PCR also has other advantages. Because primers can be synthesized that differ from the template sequence at any given position, it is possible to use PCR for site-directed mutagenesis. That is, PCR can be used to mutate a gene at a desired position in the sequence. This allows the proteins encoded by the normal and mutant genes to be expressed, purified and compared.

Analysis of gene expression

PCR can also be used to measure gene expression. Where in PCR the amount of amplified product is not determined till the end of all the cycles, a variation called quantitative real-time PCR is used, in which the accumulation of product is measured at each cycle. This is possible because real-time PCR machines have a detector module that can measure the levels of a fluorescent marker in the reaction, with the amount of fluorescence proportional to the amount of amplified product. By following the accumulation of product over the cycles it is possible to calculate the amount of starting template. To measure gene expression, the template used is mRNA reverse-transcribed into cDNA (see below). This coupling of reverse transcription with quantitative real-time PCR is called qRT-PCR.

This page titled 8.7: Polymerase Chain Reaction (PCR) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.8: Reverse Transcription

In the central dogma, DNA codes for mRNA, which codes for protein. One known exception to the central dogma is exhibited by retroviruses. These RNA-encoded viruses have a phase in their life cycle in which their genomic RNA is converted back to DNA by a virally-encoded enzyme known as reverse transcriptase. The ability to convert RNA to DNA is a method that is desirable in the laboratory for numerous reasons. For example, converting RNAs of interest to cDNA is used in RT-PCR as well as in other applications like microarray analysis.

Process

First, one creates a DNA oligonucleotide to serve as a primer for reverse transcriptase to use on a target RNA. The primer must, of course, be complementary to a segment (near the 3' end) of the RNA to be amplified. The RNA, reverse transcriptase, the primer, and four dNTPs are mixed. With one round of replication, the RNA is converted to a single strand of DNA. Denaturation frees the single stranded cDNA, which can be used as is, or converted to double-stranded cDNA, depending on the application.

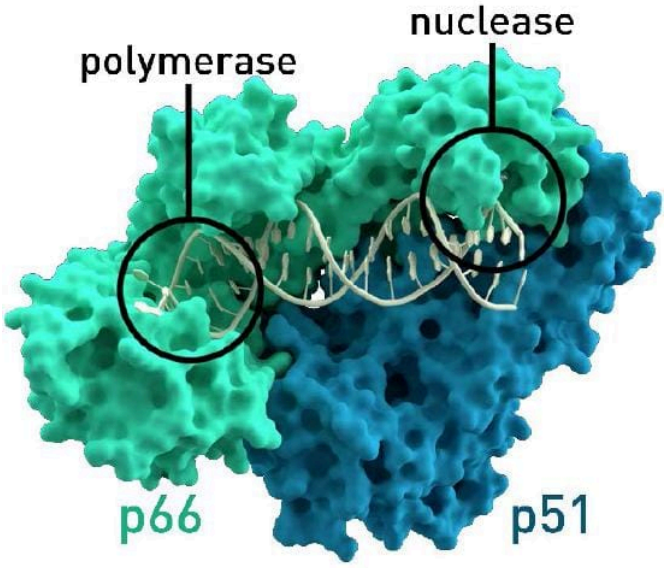


Figure 8.37 - Reverse transcriptase of HIV. The nuclease function is needed for the viral life cycle, but not for lab use. Wikipedia

Detecting molecular interactions

The study of biochemistry is basically the study of the interactions of cellular molecules. Methods for detecting interactions among biomolecules are, for this reason, very useful to biochemists. We will now discuss a couple of very different methods for detecting these inter-molecular interactions.

Yeast two-hybrid system (Y2H)



Figure 8.38 - Structure of a eukaryotic transcriptional activator, showing the DNA-binding domain (DBD) and transcriptional activation domain (TAD). Wikipedia

Yeast two-hybrid screening is a sophisticated technique for identifying which protein(s), out of a collection of all of a cell's proteins, interacts with a specific protein of interest. The method relies on the interaction between two proteins to reconstitute a functional transcriptional activator within yeast cells. You may remember that many transcriptional activators are modular proteins that have a domain that binds to DNA and another domain that activates transcription (Figure 8.38).

If the transcription factor is split, so that the binding domain is attached to one protein, and the activation domain to another

protein, a functional transcriptional activator can only be re-created if the two “carrier”proteins come into close proximity - that is, they interact. The presence of this functional activator can be detected by the expression of a reporter gene.

A simple way to understand this idea is by thinking of a transcriptional activator as a device, like a flashlight, that has two parts, the battery and the lamp, that must be together in order to function. Neither a person who has just a battery nor one who has only the lamp will be able to see in a dark room. But if the two interact by coming close enough to insert the battery in the flashlight, their interaction can be detected by the fact that the flashlight will now be functional as evidenced by the light produced.

It takes two to tango

Figure 8.39 (A) shows the normal yeast transcriptional activator, GAL4, with both the DNA-binding (DBD) and Activation domains (AD). It is able to stimulate transcription of the downstream reporter gene, lac z. Panels B and C show constructs that produce the GAL4 DBD and AD, respectively, fused to other proteins, one of which is termed the “bait” and the other as “prey”. Neither of these fusion proteins can stimulate transcription of the lac z gene. When constructs encoding both the bait and prey are in the same yeast cell, if the bait protein interacts with the prey, the DBD and AD of the GAL4 will be brought together to reconstitute a functional GAL4. The presence of functional GAL4 is readily detectable because it will stimulate expression of the lac z reporter gene. If the bait and prey proteins do not interact, then there will be no lac z expression. When interaction is detected through expression of the reporter gene, the specific prey protein can then be identified.

The yeast two-hybrid system allows for simultaneous screening of many prey proteins, by constructing large collections of fusion constructs, with each potential protein partner of the bait protein fused to the GAL4 activation domain.

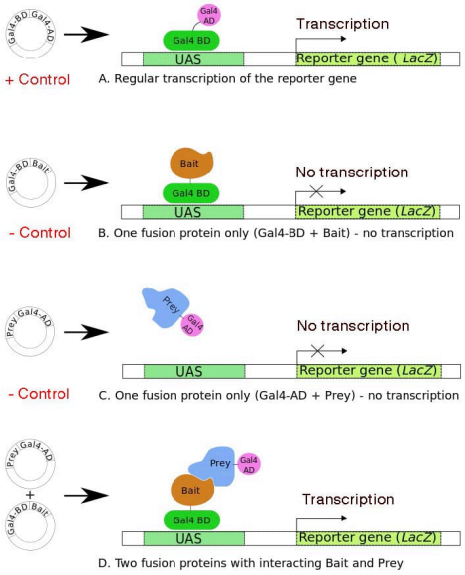


Figure 8.39 - Four scenarios for the yeast two-hybrid system. UAS = Upstream Activator Sequences - acts like a promoter. Scenario A shows that the two transcription factors start out as one protein. Wikipedia

This page titled 8.8: Reverse Transcription is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.9: FRET

Another method for detecting molecular interactions is Fluorescence resonance energy transfer (FRET) - also called Förster resonance energy transfer, resonance energy transfer (RET) or electronic energy transfer (EET). The technique is based on the observation that a molecule excited by the absorption of light can transfer energy to a nearby molecule if the emission spectrum of the first molecule overlaps with the excitation spectrum of the second (Figure 8.9.1) This transfer of energy can only take place if the two molecules are sufficiently close together (no more than a few nanometers apart).

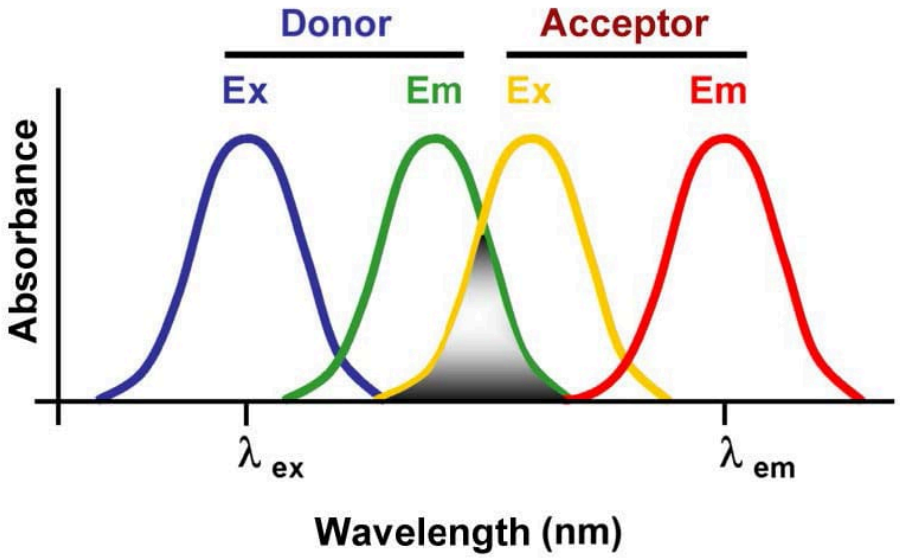


Figure 8.9.1: Excitation and emission spectra for donor and acceptor fluorophores in FRET. Wikipedia

In the technique, a donor fluorophore or an acceptor fluorophore is covalently attached to two molecules of interest. The acceptor fluorophore is designed to accept energy from the donor molecule (orange dotted line in Figure 8.9.2) and fluoresce at a unique wavelength (red arrow) when it receives that energy from the donor.

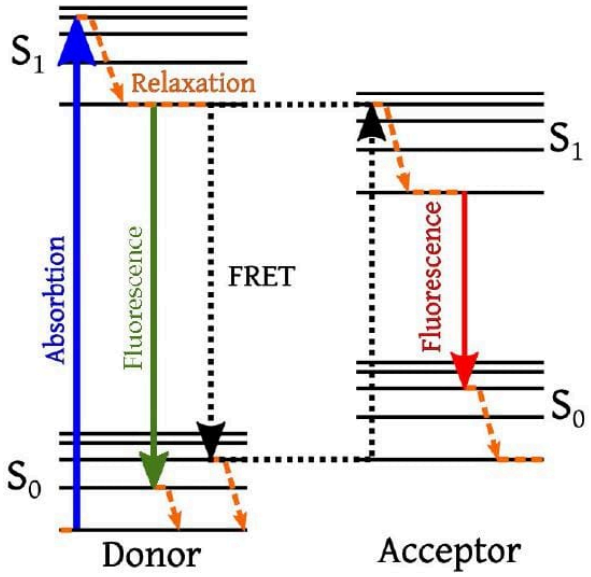


Figure 8.9.2***:*** *Fluorescence resonance energy transfer between donor and acceptor chromophores. Image by Pehr Jacobson*

Further, the wavelength of light that the donor absorbs is uniquely tailored for the donor fluorophore and has no effect on the acceptor fluorophore. The only way the acceptor can fluoresce is if it is close enough to receive energy transferred from the donor

(red arrow). This fluorescence will have a unique wavelength, as well. If the donor and acceptor are not close enough together, the donor fluoresces and emits light corresponding to the green or black arrow. These are different wavelengths than that of the red arrow.

The experiment begins in the cell with one protein with a donor fluorophore and the other protein with an acceptor fluorophore. Light of a wavelength that excites the donor fluorophore is shined on the cell. If a protein with a donor interacts with the protein carrying an acceptor, then energy transfer occurs from the donor fluorophore to the acceptor and the unique fluorescence (red line) of the acceptor is detected. If the two proteins do not interact, then little or no fluorescence from the acceptor is detected.

This page titled 8.9: FRET is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.10: Genome Editing (CRISPR)

The development of tools that would allow scientists to make specific, targeted changes in the genome has been the Holy Grail of molecular biology. An ingenious new tool that is both simple and effective in making precise changes is poised to revolutionize the field, much as PCR did in the 1980s. Known as the CRISPR/Cas9 system, and often abbreviated simply as CRISPR, it is based on a sort of bacterial immune system that allows bacteria to recognize and inactivate viral invaders.

CRSPR

CRISPR stands for Clustered Regularly Interspaced Short Palindromic Repeats, short repeated sequences found in prokaryotic DNA, separated by spacer sequences derived from past encounters with, for example, a bacteriophage. Like the glass slipper left behind by Cinderella that was later used to identify her, the pieces of the invader's sequences are a way for the bacteria to identify the virus if it attacks again. Inserted into the bacterial genome, these sequences can later be transcribed into a guide RNA that matches, and base-pairs with, sections of the viral genome if it was encountered again. A nuclease associated with the guide RNA then cleaves the sequence base-paired with the guide RNA. (The nucleases are named Cas for CRISPR-associated.)

The essential elements of this system are a guide RNA that homes in on the target sequence and a nuclease that can make a cut in the sequence that is bound by the guide RNA. By engineering guide RNAs complementary to a target gene, it is possible to target the nuclease to cleave within that gene. In the CRISPR/Cas9 system, the Cas9 endonuclease cuts both strands of the gene sequence targeted by the guide RNA (Figure 8.10.1). This generates a double-strand break that the cell attempts to repair.

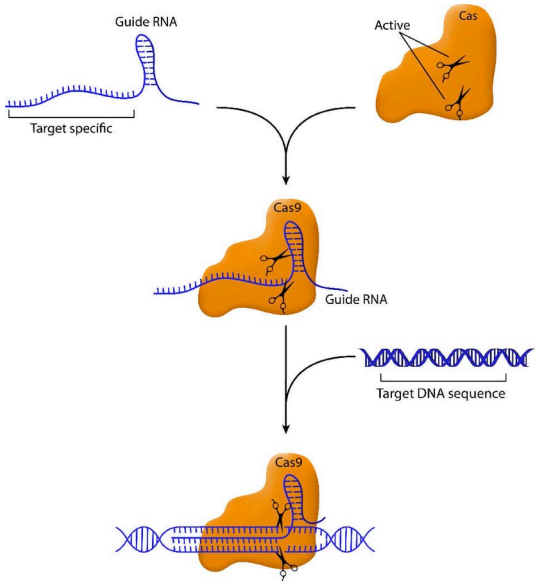


Figure 8.10.1: A guide RNA directs the Cas9 nuclease to its target gene. Image by Pehr Jacobsen

As you may remember, double-strand breaks in DNA can be repaired by simple, nonhomologous end joining (NHEJ) or by homologous recombination. When a break is fixed by NHEJ, there is good chance that there will be deletions or insertions that will inactivate the gene they are in. Thus, targeted cleavage of a site by CRISPR/Cas9 can easily and specifically inactivate a gene, making it easy to characterize the gene's function.

But, what if you wished to simply mutate the gene at a specific site to study the effect of the mutation? This, too, can be achieved. If a homologous sequence bearing the specific mutation is provided, homologous recombination can repair the break, and at the same time insert the exact mutation desired. It is obvious that if you can insert a mutation as just described, it should be possible to correct a mutation in the genome by cleaving at the appropriate spot and providing the correct sequence as a template for repair by homologous recombination. The simplicity of the system holds great promise for curing genetic diseases.

Scientists have also come up with some creative variations on the CRISPR/Cas9 system. For instance, one variant inactivates the nuclease activity of Cas9. The guide RNA in this system pairs with the target sequence, but the Cas9 does not cleave it. Instead,

the Cas9 blocks the transcription of the downstream gene (Figure 8.10.2) This method allows specific genes to be turned off without actually altering the DNA sequence.

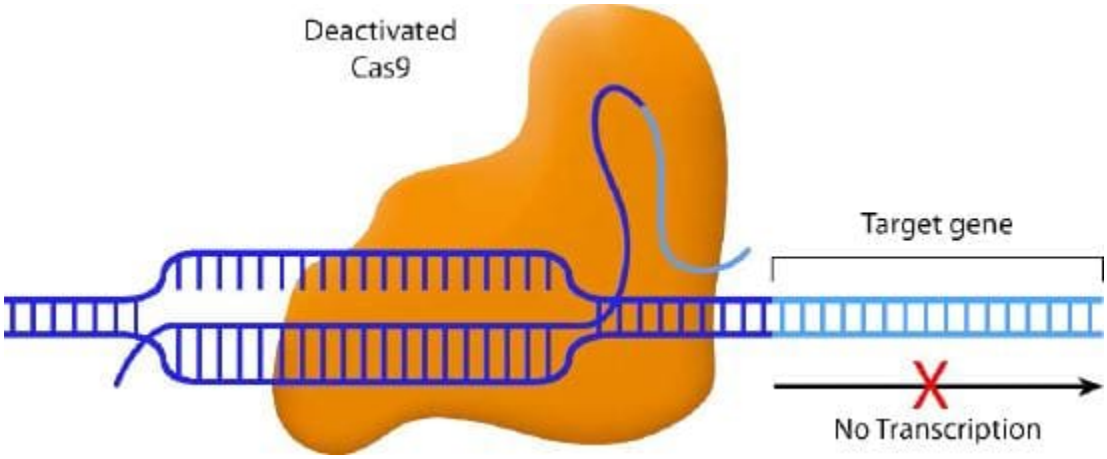


Figure 8.10.2: Inactive Cas9 can block transcription of a target gene. Image by Pehr Jacobsen

Another variation also uses a disabled Cas9, but this time, the Cas9 is fused to a transcriptional activation domain. In this situation, the guide RNA positions the Cas9-activator domain in a place where it can enhance transcription from a specific promoter (Figure 8.10.3). Other variations on this theme attach histone-modifying enzymes or DNA methylases to the inactive Cas9. Again, the guide RNA positions the Cas9 in the desired spot, and the enzyme attached to Cas9 can methylate the DNA or modify the histones in that region.

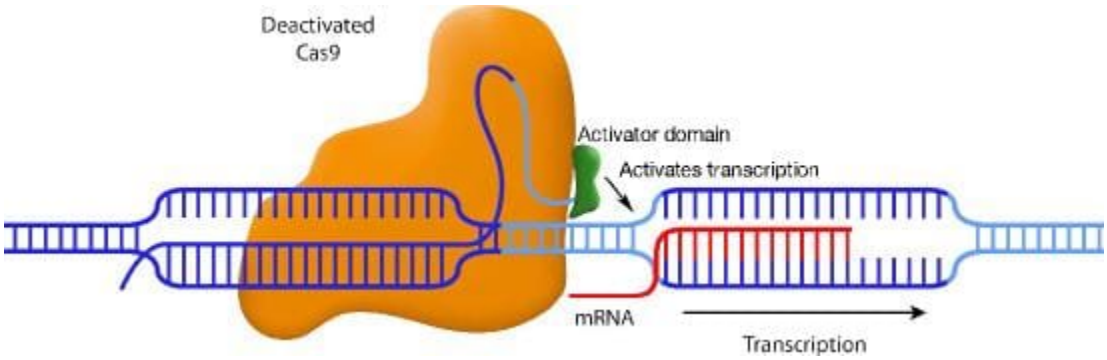


Figure 8.10.3: A Cas9-activator domain fusion can activate transcription of a target gene. Image by Pehr Jacobsen

CRISPR has already been used to edit genomes in a wide variety of species (and in human cell cultures). It may not be long before the technique is approved for clinical use. In the meanwhile, CRISPR is transforming molecular biology.

This page titled 8.10: Genome Editing (CRISPR) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.11: Protein Cleavage

Because of their large size, intact proteins can be difficult to study using analytical techniques, such as mass spectrometry. Consequently, it is often desirable to break a large polypeptide down into smaller pieces. Proteases are enzymes that typically break peptide bonds by binding to specific amino acid sequences in a protein and catalyzing their hydrolysis.

Chemical reagents, such as cyanogen bromide, which cleaves peptide bonds on the C-terminal side of a methionine residue can also be used to cut larger proteins into smaller peptides. Common proteins performing this activity are found in the digestive system and are shown below.

- Subtilisin - C-terminal side of large uncharged side chains
- Chymotrypsin - C terminal side of aromatics (Phe, Tyr, Trp)
- Trypsin - C-terminal side of lysine and arginines (not next to proline)
- Carboxypeptidase - N-terminal side of C-terminal amino acid
- Elastase - Hydrolyzes C-side of small AAs (Gly, Ala)
- Cyanogen Bromide (chemical) - Hydrolyzes C-side of Met

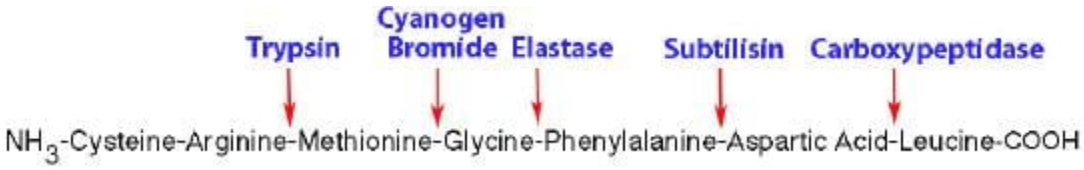


Figure 8.45 - Protease cleavage sites on a polypeptide

Determining mass and protein sequence

Mass spectrometry, as its name suggests, is a method that can be used to determine the masses of molecules. Once limited to analyzing small molecules, it has since been adapted and improved to allow the analysis of biologically important molecules like proteins and nucleic acids. Mass spectrometers use an electrical field to accelerate an ionized molecule toward a detector. The time taken by an ionized molecule to move from its point of ionization to the detector will depend on both its mass and its charge and is termed its time of flight (TOF).

MALDI-TOF



Figure 8.46 - A desktop MALDI-TOF system

MALDI-TOF (Matrix-assisted Laser Desorption Ionization - Time of Flight) is an analytical technique allowing one to determine the molecular masses of biologically relevant molecules with great precision. It is commonly used in proteomics and determination of masses of large biomolecules, including nucleic acids. The development of MALDI, which permits the production of ionic forms of relatively large molecules, was crucial to the successful use of mass spectrometry of biomolecules. Figure 8.46 shows a compact MALDI-TOF system.

The MALDI-TOF process involves three basic steps. First, the material to be analyzed is embedded in solid support material (matrix) that can be volatilized in a vacuum chamber by a laser beam. In the second part of the process, a laser focused on the matrix volatilizes the sample, causing the molecules within it to vaporize and, in the process, to form ions by either gaining or losing protons. Third, the ions thus created in the sample are accelerated by an electric field towards a detector. Their rate of movement towards the detector is a function of the ratio of their mass to charge (m/z). An ion with a mass of 100 and a charge of +1 will move twice as fast as an ion with a mass of 200 and a charge of +1 and at the same rate as an ion with a mass of 200 and a charge of +2. Thus, by precisely determining the time it takes for an ion to go from ionization (time zero of the laser treatment) to being detected, the mass to charge ratio for all of the molecules in a sample can be readily determined.

Ionization may result in destabilization of larger molecules, which fragment into smaller ones in the MALDI-TOF detection chamber. The size of each of the sub-fragments of a larger molecule allows one to determine its identity if this is not previously known. This fragmentation can be intentionally enhanced by having the accelerated ions collide with an inert gas, like argon.

Fragmentation of a molecule may also be carried out prior to analysis, as for example, by cleaving a protein into smaller peptides by the use of enzymes or chemical agents. The amino acid sequence of a protein may be determined by using MALDI-TOF by analyzing the precise molecular masses of the many short peptide fragments obtained from a protein. When one amino acid, for example, fragments from a larger peptide, this can be detected as the difference in mass between the fragment with and without the amino acid, since each amino acid will have a characteristic molecular mass. By peptide mass fingerprinting and analysis of smaller fragments of individual peptides, the entire sequence of a polypeptide can, thus, be determined.

This page titled 8.11: Protein Cleavage is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

8.12: Membrane Dynamics (FRAP)

Understanding the dynamics of movement in the membranes of cells is the province of the Fluorescence Recovery After Photobleaching (FRAP) technique (Figure 8.47). This optical technique is used to measure the two dimensional lateral diffusion of molecules in thin films, like membranes, using fluorescently labeled probes. It also has applications in protein binding.

In the method, a lipid bilayer is uniformly labeled with a fluorescent tag (Figure 8.47, Step A) and then a subset of the tag is bleached using a laser (Step B). The spread of the bleached molecules is followed using a microscope (Step C). Information obtained in this manner provides data about the rate of lateral diffusion occurring in a lipid bilayer (Step D).

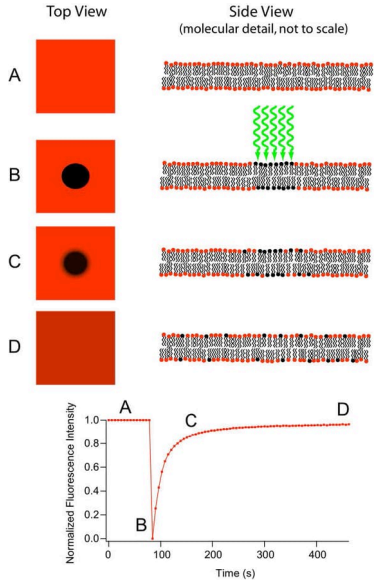


Figure 8.47 - FRAP analysis. Wikipedia

This page titled 8.12: Membrane Dynamics (FRAP) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

9.1: Section 1-

This page titled 9.1: Section 1- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

9.2: Section 2-

This page titled 9.2: Section 2- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

9.3: Section 3-

This page titled 9.3: Section 3- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

9.4: Section 4-

This page titled 9.4: Section 4- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

9.5: Section 5-

This page titled 9.5: Section 5- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

9.6: Section 6-

This page titled 9.6: Section 6- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

10.1: Section 1-

This page titled 10.1: Section 1- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

10.2: Section 2-

This page titled 10.2: Section 2- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

10.3: Section 3-

This page titled 10.3: Section 3- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

10.4: Section 4-

This page titled 10.4: Section 4- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

10.5: Section 5-

This page titled 10.5: Section 5- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

10.6: Section 6-

This page titled 10.6: Section 6- is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11: POINT BY POINT

- Topic hierarchy

11.1: In the Beginning
11.2: Structure and Function
11.3: Membranes
11.4: Catalysis
11.5: Energy
11.6: Metabolism
11.7: Information Processing
11.8: Techniques

This page titled 11: Point by Point is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.1: In the Beginning

This page titled 11.1: In the Beginning is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.2: Structure and Function

This page titled 11.2: Structure and Function is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.3: Membranes

This page titled 11.3: Membranes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.4: Catalysis

This page titled 11.4: Catalysis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.5: Energy

This page titled 11.5: Energy is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.6: Metabolism

This page titled 11.6: Metabolism is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.7: Information Processing

This page titled 11.7: Information Processing is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

11.8: Techniques

This page titled 11.8: Techniques is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, and Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.

INDEX

A

actin
2.5: Structure and Function- Protein Function II

C

Cardiolipins
2.8: Structure and Function - Lipids and Membranes
Chemiosmotic model
5.2: Electron Transport and Oxidative Phosphorylation
chromatography
8.2: Fractionation and Chromatography Techniques
Chymotrypsin
4.3: Mechanisms of Catalysis
citric acid cycle
6.2: Citric Acid Cycle & Related Pathways
clotting
4.4: Blood Clotting
collagen
2.4: Structure and Function- Proteins II
CRISPR
8.10: Genome Editing (CRISPR)

D

DNA (Z Form)
2.6: Structure and Function - Nucleic Acids
dyneins
2.5: Structure and Function- Protein Function II

E

electrophoresis
8.3: Electrophoresis
endocytosis
3.3: Other Considerations in Membranes
Expression Cloning
8.6: Isolating Genes
expression vectors
8.6: Isolating Genes

F

Fluorescence Recovery After Photobleaching (FRAP)
8.12: Membrane Dynamics (FRAP)
Fractionation
8.2: Fractionation and Chromatography Techniques
FRET
8.9: FRET

G

gluconeogenesis
6.1: Metabolism - Sugars
glycolytic pathway
6.1: Metabolism - Sugars

H

Hind III
8.6: Isolating Genes

I

isoelectric focusing
8.3: Electrophoresis
isoprenoids
6.4: Other Lipids

K

keratins
2.4: Structure and Function- Proteins II
Kinesins
2.5: Structure and Function- Protein Function II

L

lysis
8.1: Cell Lysis

M

Macropinocytosis
3.3: Other Considerations in Membranes
microarrays
8.4: Detection, identification and quantitation of specific nucleic acids and proteins
microfilaments
2.5: Structure and Function- Protein Function II
microtubules
2.5: Structure and Function- Protein Function II
Morpheein model
4.2: Control of Enzymatic Activity

N

northern blot
8.4: Detection, identification and quantitation of specific nucleic acids and proteins
nucleic acids
2.6: Structure and Function - Nucleic Acids

P

Palindromes
8.6: Isolating Genes
Perilipin
6.3: Fats and Fatty Acids
phagocytosis
3.3: Other Considerations in Membranes
Photophosphorylation
5.3: Energy - Photophosphorylation
Polyacrylamide gel electrophoresis
8.3: Electrophoresis
polymerase chain reaction (PCR)
8.7: Polymerase Chain Reaction (PCR)
Porphyria
6.4: Other Lipids
promoter
7.5: Transcription

R

replication
7.3: DNA Replication
restriction enzymes
8.6: Isolating Genes

S

secondary structure
2.4: Structure and Function- Proteins II
serine proteases
4.3: Mechanisms of Catalysis
signaling
7.9: Signaling
Size Exclusion Chromatography
8.2: Fractionation and Chromatography Techniques
SNARE proteins
3.3: Other Considerations in Membranes
Southern blot
8.4: Detection, identification and quantitation of specific nucleic acids and proteins
sphingolipids
6.4: Other Lipids
Squalene
6.4: Other Lipids
statins
6.4: Other Lipids
subtilisin
4.3: Mechanisms of Catalysis

T

tertiary structure
2.4: Structure and Function- Proteins II
transcription
7.5: Transcription
transcriptome
8.5: Transcriptomics
transglutaminase
4.4: Blood Clotting
translation
7.7: Translation
tubulin
2.5: Structure and Function- Protein Function II

U

urea cycle
6.5: Amino Acids and the Urea Cycle

V

von Willebrand factor
4.4: Blood Clotting

W

warfarin
4.4: Blood Clotting
Western blot
8.4: Detection, identification and quantitation of specific nucleic acids and proteins

This page titled Index is shared under a license and was authored, remixed, and/or curated by LibreTexts.

GLOSSARY

S

Sample Word 1 | Sample Definition 1

This page titled Glossary is shared under a license and was authored, remixed, and/or curated by LibreTexts.

DETAILED LICENSING

Overview

Title: Biochemistry Free For All (Ahern, Rajagopal, and Tan)

Webpages: 89

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 89.9% (80 pages)
- Undeclared: 10.1% (9 pages)

By Page

- Biochemistry Free For All (Ahern, Rajagopal, and Tan) - *CC BY-NC-SA 4.0*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: In The Beginning - *CC BY-NC-SA 4.0*
    - 1.1: Introduction - Basic Biology - *CC BY-NC-SA 4.0*
    - 1.2: Introduction - Basic Chemistry - *CC BY-NC-SA 4.0*
    - 1.3: Introduction - Water and Buffers - *CC BY-NC-SA 4.0*
  - 2: Structure and Function - *CC BY-NC-SA 4.0*
    - 2.1: Prelude to Structure and Function - *CC BY-NC-SA 4.0*
    - 2.4: Structure and Function- Proteins II - *CC BY-NC-SA 4.0*
    - 2.5: Structure and Function- Protein Function II - *CC BY-NC-SA 4.0*
    - 2.6: Structure and Function - Nucleic Acids - *CC BY-NC-SA 4.0*
    - 2.7: Structure and Function- Carbohydrates - *CC BY-NC-SA 4.0*
    - 2.8: Structure and Function - Lipids and Membranes - *CC BY-NC-SA 4.0*
    - 2.2: Structure & Function - Amino Acids - *CC BY-NC-SA 4.0*
    - 2.3: Structure & Function- Proteins I - *CC BY-NC-SA 4.0*
  - 3: Membranes - *CC BY-NC-SA 4.0*
    - 3.1: Basic Concepts in Membranes - *CC BY-NC-SA 4.0*
    - 3.2: Transport in Membranes - *CC BY-NC-SA 4.0*
    - 3.3: Other Considerations in Membranes - *CC BY-NC-SA 4.0*
  - 4: Catalysis - *CC BY-NC-SA 4.0*
    - 4.1: Basic Principles of Catalysis - *CC BY-NC-SA 4.0*
    - 4.2: Control of Enzymatic Activity - *CC BY-NC-SA 4.0*
    - 4.3: Mechanisms of Catalysis - *CC BY-NC-SA 4.0*
    - 4.4: Blood Clotting - *CC BY-NC-SA 4.0*
  - 5: Energy - *CC BY-NC-SA 4.0*
    - 5.1: Basics of Energy - *CC BY-NC-SA 4.0*
    - 5.2: Electron Transport and Oxidative Phosphorylation - *CC BY-NC-SA 4.0*
    - 5.3: Energy - Photophosphorylation - *CC BY-NC-SA 4.0*
  - 6: Metabolism - *CC BY-NC-SA 4.0*
    - 6.1: Metabolism - Sugars - *CC BY-NC-SA 4.0*
    - 6.2: Citric Acid Cycle & Related Pathways - *CC BY-NC-SA 4.0*
    - 6.3: Fats and Fatty Acids - *CC BY-NC-SA 4.0*
    - 6.4: Other Lipids - *CC BY-NC-SA 4.0*
    - 6.5: Amino Acids and the Urea Cycle - *CC BY-NC-SA 4.0*
    - 6.6: Nucleotides - *CC BY-NC-SA 4.0*
  - 7: Information Processing - *CC BY-NC-SA 4.0*
    - 7.1: Prelude to Information Processing - *CC BY-NC-SA 4.0*
    - 7.2: Genes and Genomes - *CC BY-NC-SA 4.0*
    - 7.3: DNA Replication - *CC BY-NC-SA 4.0*
    - 7.4: DNA Repair - *CC BY-NC-SA 4.0*
    - 7.5: Transcription - *CC BY-NC-SA 4.0*
    - 7.6: RNA Processing - *CC BY-NC-SA 4.0*
    - 7.7: Translation - *CC BY-NC-SA 4.0*
    - 7.8: Gene Expression - *CC BY-NC-SA 4.0*
    - 7.9: Signaling - *CC BY-NC-SA 4.0*
  - 8: Basic Techniques - *CC BY-NC-SA 4.0*
    - 8.1: Cell Lysis - *CC BY-NC-SA 4.0*
    - 8.2: Fractionation and Chromatography Techniques - *CC BY-NC-SA 4.0*

- 8.3: Electrophoresis - *CC BY-NC-SA 4.0*
- 8.4: Detection, identification and quantitation of specific nucleic acids and proteins - *CC BY-NC-SA 4.0*
- 8.5: Transcriptomics - *CC BY-NC-SA 4.0*
- 8.6: Isolating Genes - *CC BY-NC-SA 4.0*
- 8.7: Polymerase Chain Reaction (PCR) - *CC BY-NC-SA 4.0*
- 8.8: Reverse Transcription - *CC BY-NC-SA 4.0*
- 8.9: FRET - *CC BY-NC-SA 4.0*
- 8.10: Genome Editing (CRISPR) - *CC BY-NC-SA 4.0*
- 8.11: Protein Cleavage - *CC BY-NC-SA 4.0*
- 8.12: Membrane Dynamics (FRAP) - *CC BY-NC-SA 4.0*

◦ 9: Chapter 10 - *CC BY-NC-SA 4.0*
- 9.1: Section 1- - *CC BY-NC-SA 4.0*
- 9.2: Section 2- - *CC BY-NC-SA 4.0*
- 9.3: Section 3- - *CC BY-NC-SA 4.0*
- 9.4: Section 4- - *CC BY-NC-SA 4.0*
- 9.5: Section 5- - *CC BY-NC-SA 4.0*
- 9.6: Section 6- - *CC BY-NC-SA 4.0*

◦ 10: Chapter 11 - *CC BY-NC-SA 4.0*
- 10.1: Section 1- - *CC BY-NC-SA 4.0*
- 10.2: Section 2- - *CC BY-NC-SA 4.0*
- 10.3: Section 3- - *CC BY-NC-SA 4.0*
- 10.4: Section 4- - *CC BY-NC-SA 4.0*
- 10.5: Section 5- - *CC BY-NC-SA 4.0*
- 10.6: Section 6- - *CC BY-NC-SA 4.0*

◦ 11: Point by Point - *CC BY-NC-SA 4.0*
- 11.1: In the Beginning - *CC BY-NC-SA 4.0*
- 11.2: Structure and Function - *CC BY-NC-SA 4.0*
- 11.3: Membranes - *CC BY-NC-SA 4.0*
- 11.4: Catalysis - *CC BY-NC-SA 4.0*
- 11.5: Energy - *CC BY-NC-SA 4.0*
- 11.6: Metabolism - *CC BY-NC-SA 4.0*
- 11.7: Information Processing - *CC BY-NC-SA 4.0*
- 11.8: Techniques - *CC BY-NC-SA 4.0*

◦ Back Matter - *Undeclared*
- Index - *Undeclared*
- Glossary - *Undeclared*
- Detailed Licensing - *Undeclared*

This page titled Biochemistry Free For All (Ahern, Rajagopal, and Tan) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kevin Ahern, Indira Rajagopal, & Taralyn Tan via source content that was edited to the style and standards of the LibreTexts platform.