

CHEM 4320/5320: BIOCHEMISTRY 1



Unknown

Chem 4320/5320: Biochemistry 1

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

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

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



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

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

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

Have questions or comments? For information about adoptions or adaptations 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 08/22/2026

TABLE OF CONTENTS

Licensing

1: Amino Acids

1.1: Nomenclature of Amino acids

1.2: Structure of Amino Acids

1.3: Properties of Amino Acids

1.3.1: Charged Nature of Amino Acids

1.3.2: Stereochemistry of Amino Acids

1.4: Reactions of Amino Acids

1.4.1: Acid-base Chemistry of Amino Acids

1.4.2: Acid-Base Reactions of Amino Acids

1.4.3: Introduction to Amino Acid Reactivity

1.5: Amino Acid Analysis and Chemical Sequencing

1.5.1: Polypeptides and Proteins

1.5.2: Sequence Determination Using Mass Spectrometry

1.5.3: The Edman Degradation

2: Protein Structure

2.1: The Structure of Proteins

2.2: Protein Folding

2.3: Denaturation of proteins

2.4: Amino Acids and Proteins (Exercises)

3: Methods of Protein Purification and Characterization

3.1: Protein Purification

3.2: Cell Disruption

3.3: Fractionation

3.4: Chromatography

3.4.1: Affinity Chromatography

3.4.2: Gel Exclusion Chromatography

3.4.3: Ion Exchange Chromatography

3.5: Electrophoresis

3.5.1: Blotting and Hybridization

3.5.2: DNA Analysis - Gel Electrophoresis

3.5.3: Exercises

4: Overview of Hemoglobin and Myoglobin

4.1: Myoglobin, Hemoglobin, and their Ligands

4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin

4.3: Exercises

5: Michaelis-Menten Enzyme Kinetics

5.1: Catalytic Efficiency of Enzymes

5.2: Michaelis-Menten Kinetics

5.3: Enzyme Parameters

5.4: Enzyme Inhibition

5.4.1: Enzyme Inhibition

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

5.6: Multi-Substrate Sequential Mechanisms

5.7: Double displacement reaction

6: Catalytic Strategies of Enzymes

6.1: Serine Proteases

6.2: Chymotrypsin

6.3: Trypsin

7: Enzyme Active Site and Substrate Specificity

8: Enzyme Regulation

8.1: Regulation of Enzyme Activity

8.2: Control of Metabolism Through Enzyme Regulation

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

8.4: Exercises

9: Carbohydrates

9.1: Carbohydrates Fundamentals

9.2: Monosaccharides

9.3: Disaccharides

9.4: Oligosaccharides

9.5: Polysaccharides

9.6: Glycosides

9.7: Exercises

10: The Generation of Energy from Carbohydrate Metabolism

11: Oxidation

11.1: The Electron Transport Chain

11.2: ATP and Oxidative Phosphorylation

11.3: Exercises

12: The Flow of Genetic Information

12.1: The Structure of DNA

12.2: DNA Replication

12.2.1: Replication

12.3: RNA

[12.3.1: Protein-RNA Recognition](#)

[12.3.2: RNA - Transcription](#)

[12.3.3: RNA Processing](#)

[12.3.4: Types of RNA](#)

[12.4: Regulation of Transcription](#)

[12.5: The Genetic Code](#)

[12.6: Translation](#)

[Index](#)

[Glossary](#)

[Detailed Licensing](#)

Licensing

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

1: AMINO ACIDS

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

1.1: Nomenclature of Amino acids

1.2: Structure of Amino Acids

1.3: Properties of Amino Acids

1.3.1: Charged Nature of Amino Acids

1.3.2: Stereochemistry of Amino Acids

1.4: Reactions of Amino Acids

1.4.1: Acid-base Chemistry of Amino Acids

1.4.2: Acid-Base Reactions of Amino Acids

1.4.3: Introduction to Amino Acid Reactivity

1.5: Amino Acid Analysis and Chemical Sequencing

1.5.1: Polypeptides and Proteins

1.5.2: Sequence Determination Using Mass Spectrometry

1.5.3: The Edman Degradation

1.1: Nomenclature of Amino acids

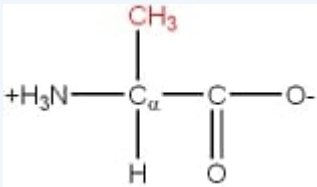
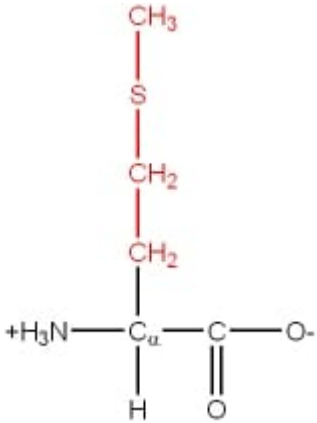
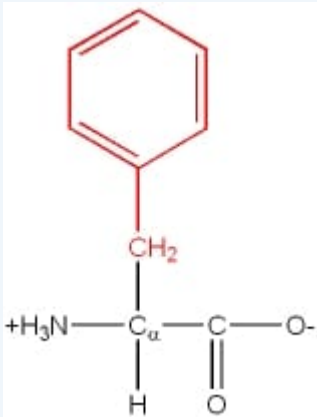
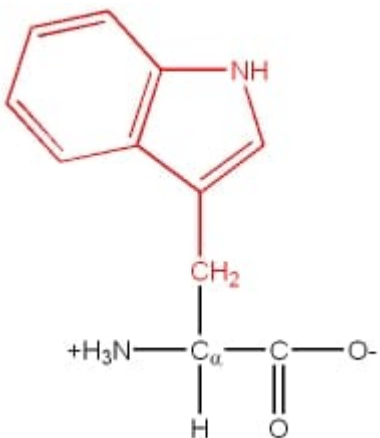
Common amino acids

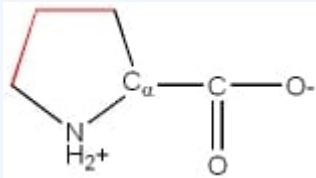
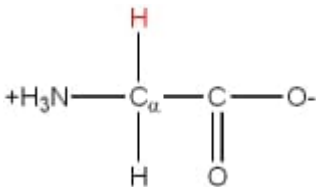
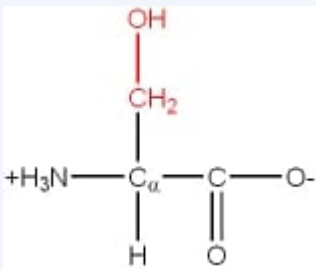
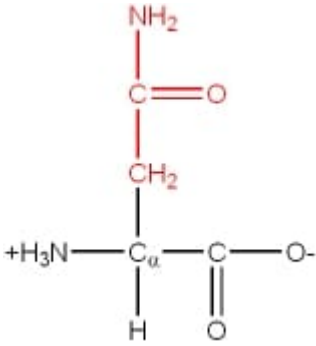
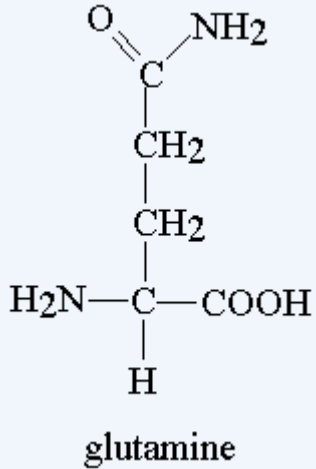
There are 20 common amino acids. They are composed of C, H, O, N and S atoms. They are structurally and chemically different, and also differ in size and volume. Some are branched structures, some are linear, some have ring structures. One of the 20 common amino acids is actually an imino acid. A typical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

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

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

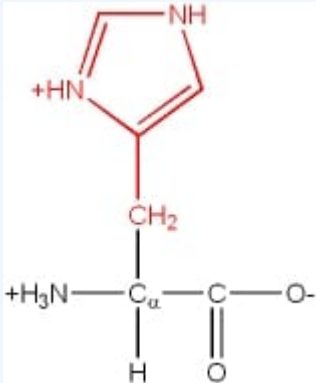
Type	Name	R-group Structure
Nonpolar	Leucine Leu, L	$ \begin{array}{c} \text{CH}_3 \\ \\ \text{CH}-\text{CH}_3 \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \quad \\ \text{H} \quad \quad \text{O} \end{array} $
	Isoleucine Ile, I	$ \begin{array}{c} \text{CH}_3 \\ \\ \text{CH}_2 \\ \\ \text{CH}-\text{CH}_3 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \quad \\ \text{H} \quad \quad \text{O} \end{array} $
	Valine Val, V	$ \begin{array}{c} \text{CH}_3 \\ \\ \text{CH}-\text{CH}_3 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \quad \\ \text{H} \quad \quad \text{O} \end{array} $

Type	Name	R-group Structure
	Alanine Ala, A	
	Methionine Met, M	
	Phenylalanine Phe, F	
	Tryptophan Trp, W	

Type	Name	R-group Structure
	Proline Pro, P	
	Glycine Gly, G (note: sometimes included in polar group)	
Polar, uncharged	Serine Ser, S	
	Asparagine Asn, N	
	Glutamine Gln, Q	 glutamine

Type	Name	R-group Structure
	Threonine Thr, T	$ \begin{array}{c} \text{CH}_3 \\ \\ \text{CH}-\text{OH} \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \parallel \\ \text{H} \quad \text{O} \end{array} $
	Cysteine Cys, C	$ \begin{array}{c} \text{SH} \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \parallel \\ \text{H} \quad \text{O} \end{array} $
	Tyrosine Tyr, Y	$ \begin{array}{c} \text{OH} \\ \\ \text{C}_6\text{H}_4 \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \parallel \\ \text{H} \quad \text{O} \end{array} $
Acidic	Aspartic acid Asp, D	$ \begin{array}{c} \text{O}^- \\ \\ \text{C}=\text{O} \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \parallel \\ \text{H} \quad \text{O} \end{array} $

Type	Name	R-group Structure
	Glutamic acid Glu, E	$ \begin{array}{c} \text{O}^- \\ \\ \text{C}=\text{O} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \\ \text{H} \quad \text{O} \end{array} $
Basic	Lysine Lys, K	$ \begin{array}{c} \text{NH}_3^+ \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \\ \text{H} \quad \text{O} \end{array} $
	Arginine Arg, R	$ \begin{array}{c} \text{NH}_3^+ \\ \\ \text{C}=\text{NH} \\ \\ \text{NH} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{+H}_3\text{N}-\text{C}_\alpha-\text{C}-\text{O}^- \\ \quad \\ \text{H} \quad \text{O} \end{array} $

Type	Name	R-group Structure
	Histidine His, H	

Uncommon amino acids

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

- [Hydroxylysine](#) and [hydroxyproline](#). These are found in the protein collagen. Collagen is a fibrous protein made up of three polypeptides that form a stable assembly, but only if the proline and lysine residues are hydroxylated. (requires vitamin C for reduction of these amino acids to hydroxy form)
- [Thyroxine](#), an iodinated derivative of tyrosine, found in thyroglobulin (produced by thyroid gland; requires iodine in diet)
- g-carboxyglutamic acid (i.e. glutamic acid with two carboxyl groups) found in certain blood clotting enzymes (requires vitamin K for production)
- N-methyl arginine and n-acetyl lysine. Found in some DNA binding proteins known as histones

Amino acid derivatives not found in proteins

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

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

Contributors

- [Mike Blaber \(Florida State University\)](#)

Thumbnail: 3D model of L-tryptophan. Image used with permission (Public Domain; [Benjah-bmm27](#)).

1.2: Structure of Amino Acids

Contributors and Attributions

- [Mike Blaber \(Florida State University\)](#)

This page titled [1.2: Structure of Amino Acids](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [Structure of Amino Acids](#) is licensed [CC BY-NC-SA 4.0](#).

1.3.1: Charged Nature of Amino Acids

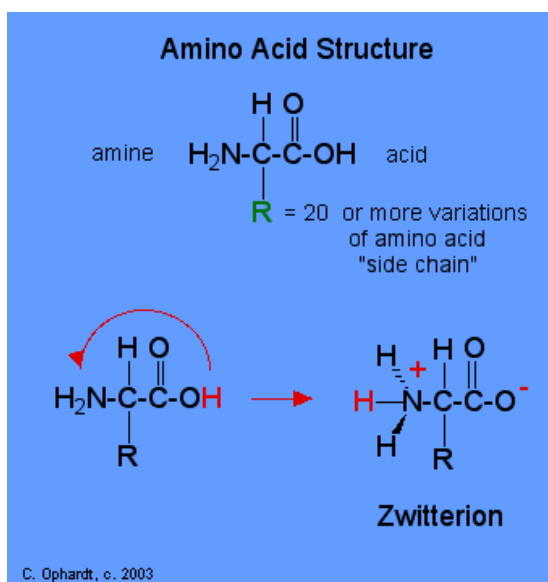
Although we are studying only about 20 amino acids, there are about six more found in the body. Many others are also known from a variety of sources. Amino acids are the building blocks used to make proteins and peptides. The different amino acids have interesting properties because they have a variety of structural parts which result in different polarities and solubilities.

Each amino acid has at least one amine and one acid functional group as the name implies. See graphic on the left. The different properties result from variations in the structures of different R groups. The R group is often referred to as the amino acid "side chain". Amino acids have special common names, however, a three letter abbreviation for the name is used most of the time. Consult the amino acid table on the next page for structure, names, and abbreviations.

Zwitterion

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

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



Contributors

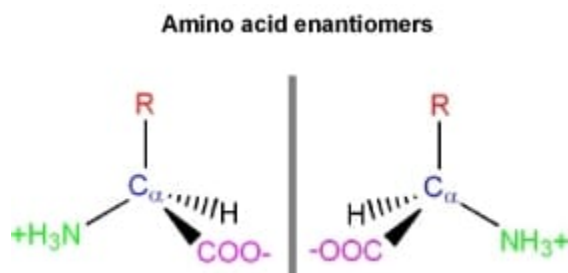
Charles Ophardt (Professor Emeritus, Elmhurst College); [Virtual Chembook](#)

This page titled [1.3.1: Charged Nature of Amino Acids](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [Charged Nature of Amino Acids](#) is licensed [CC BY-NC-SA 4.0](#).

1.3.2: Stereochemistry of Amino Acids

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



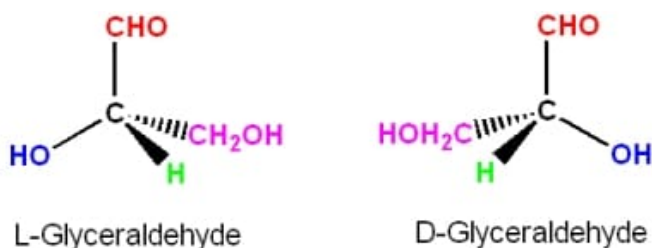
Exercise 1.3.2.1

How are these two uniquely different structures in the figure ago distinguished?

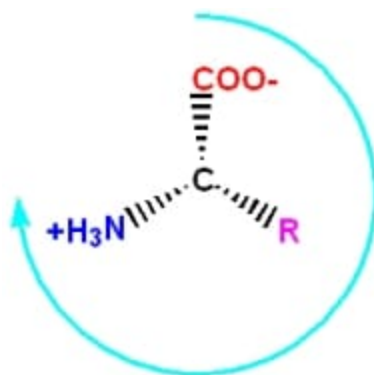
Answer available at the end of this section.

The D, L system

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



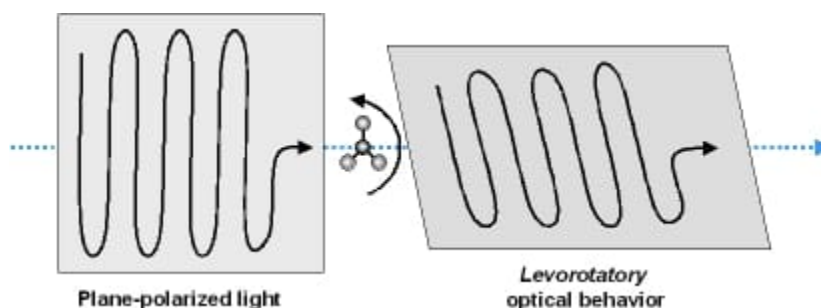
Even though the two enantiomers would seem to be essentially equivalent to each other, all common amino acids are found in the "L" enantiomer in living systems. When looking down the H-C, a bond towards the C_α there is a mnemonic to identify the L-enantiomer of amino acids (note: in this view the three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



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

Optical Activity

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



A source of potential confusion...

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

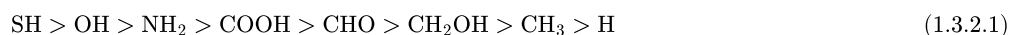
- L(+)-alanine (this is the L-enantiomer and it is dextrorotatory)
- L(-)-serine (this is the L-enantiomer and it is levorotatory)

Multiple chiral centers

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

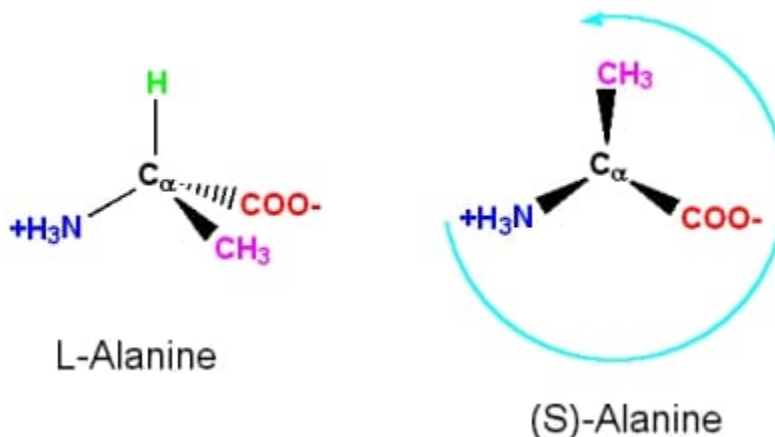
The R,S system of naming chiral centers

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



- A chiral center has four different functional groups. Identify the functional group with the lowest priority

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



Spectroscopic properties of amino acids

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

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

Separation and analysis of amino acid mixtures

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

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

Amino Acid	Mass (Da)	Isoelectric Point	Amino Acid	Mass (Da)	Isoelectric Point
Aspartic Acid	114.11	2.98	Isoleucine	113.16	6.038
Glutamic Acid	129.12	3.08	Glycine	57.05	6.064
Cysteine	103.15	5.02	Alanine	71.09	6.107

Amino Acid	Mass (Da)	Isoelectric Point	Amino Acid	Mass (Da)	Isoelectric Point
Tyrosine	163.18	5.63	Proline	97.12	6.3
Serine	87.08	5.68	Histidine	137.14	7.64
Methionine	131.19	5.74	Lysine	128.17	9.47
Tryptophan	186.12	5.88	Arginine	156.19	10.76
Phenylalanine	147.18	5.91	Threonine	101.11	-
Valine	99.14	6.002	Asparagine	115.09	-
Leucine	113.16	6.036	Glutamine	128.14	-

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

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

References

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

Contributors and Attributions

- [Mike Blaber \(Florida State University\)](#)

Answers

Exercise 1.3.2.1

Based off the stereochemistry at the C_α .

This page titled [1.3.2: Stereochemistry of Amino Acids](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [Stereochemistry of Amino Acids](#) is licensed [CC BY-NC-SA 4.0](#).

1.4.1: Acid-base Chemistry of Amino Acids

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

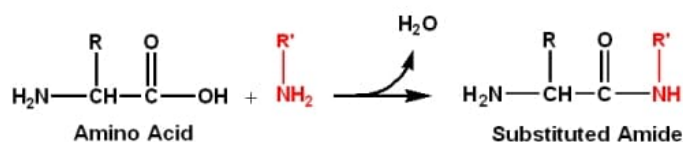
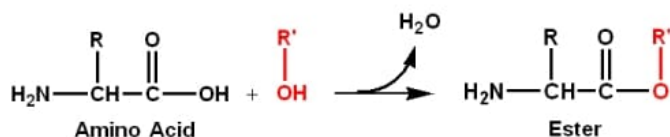
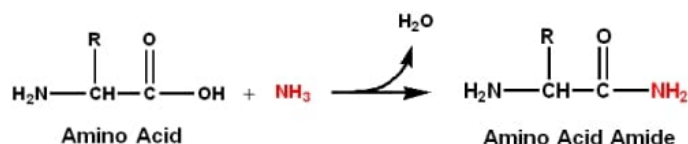
Amino acid	Functional Group	Side chain pKa
Cysteine	-SH	8.3
Serine	-OH	13
Threonine	-OH	13
Tyrosine	-OH	10.1
Aspartic acid	-COOH	3.9
Glutamic acid	-COOH	4.3
Histidine	Imidazole ring	6.0
Arginine	Guanidino	12.5
Lysine	-NH ₂	10.5

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

Reactions of amino acids

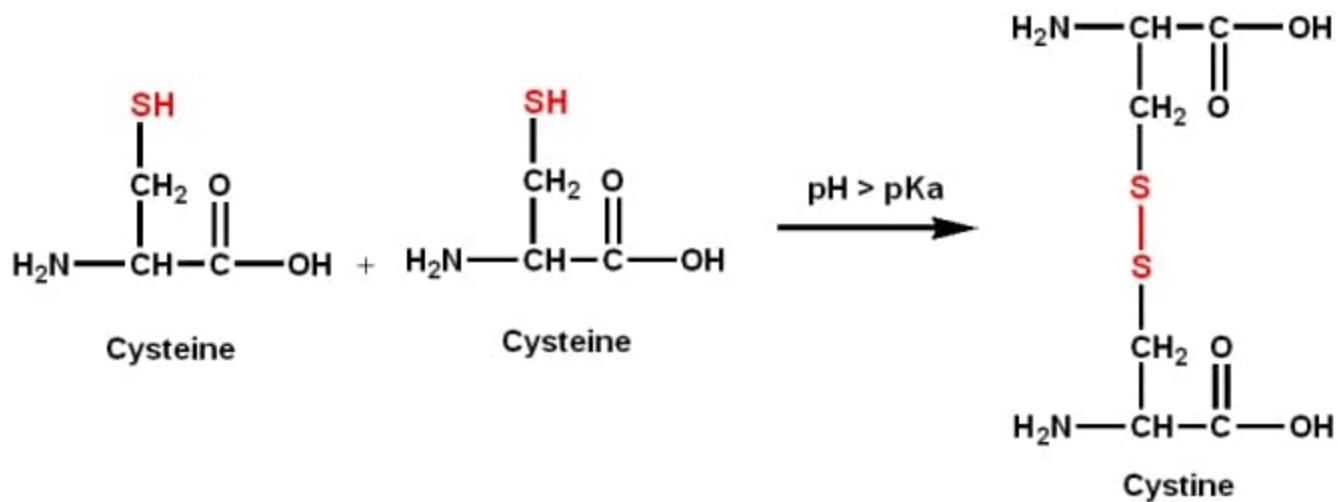
- Free amino acids (excluding proline) share similar chemical reactivities due to the common amino and carboxyl groups.
- Different amino acid side chains have different chemical reactivities. Therefore,
- reactivities of different proteins reflects the composition of the unique sequence of amino acids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

Some common carboxyl-group reactivities:



NOTE: This can lead to polymerization of amino acids

A common side chain reaction involving cysteine:



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

Contributors and Attributions

- Mike Blaber (Florida State University)

Template:HideTOC

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

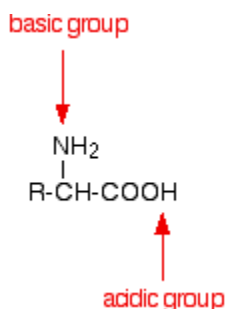
- Acid-base Chemistry of Amino Acids is licensed CC BY-NC-SA 4.0.

1.4.2: Acid-Base Reactions of Amino Acids

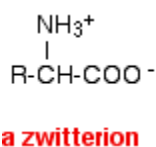
This page looks at what happens to amino acids as you change the pH by adding either acids or alkalis to their solutions. For simplicity, the page only looks at amino acids which contain a single -NH_2 group and a single -COOH group.

Amino acids as Zwitterions

An amino acid has both a basic amine group and an acidic carboxylic acid group.



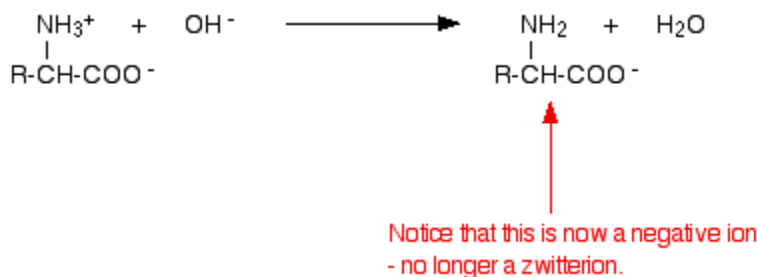
There is an internal transfer of a hydrogen ion from the -COOH group to the -NH_2 group to leave an ion with both a negative charge and a positive charge. This is called a zwitterion.



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

Adding an alkali to an amino acid solution

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



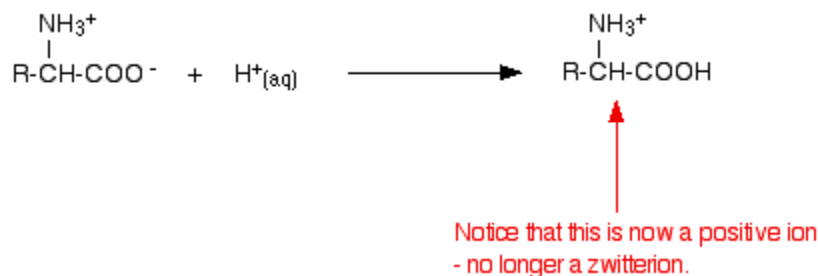
You could show that the amino acid now existed as a negative ion using [electrophoresis](#). In its simplest form, electrophoresis can just consist of a piece of moistened filter paper on a microscope slide with a crocodile clip at each end attached to a battery. A drop of amino acid solution is placed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

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

to travel towards the anode (the positive electrode).

Adding an acid to an amino acid solution

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



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

Shifting the pH from one extreme to the other

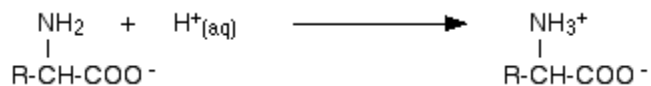
Suppose you start with the ion we've just produced under acidic conditions and slowly add alkali to it. That ion contains two acidic hydrogens - the one in the -COOH group and the one in the -NH_3^+ group. The more acidic of these is the one in the -COOH group, and so that is removed first - and you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



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



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



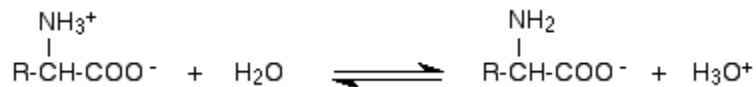
... and, of course, you can keep going by then adding a hydrogen ion to the -COO^- group.



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

When an amino acid dissolves in water, the situation is a little bit more complicated than we tend to pretend at this level. The

zwitterion interacts with water molecules - acting as both an acid and a base. As an acid:



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

As a base:



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

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

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

Contributors and Attributions

- Jim Clark (Chemguide.co.uk)

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

- [Acid-Base Reactions of Amino Acids](#) by [Jim Clark](#) is licensed [CC BY-NC 4.0](#).

1.3: PROPERTIES OF AMINO ACIDS

1.3.1: Charged Nature of Amino Acids

1.3.2: Stereochemistry of Amino Acids

This page titled [1.3: Properties of Amino Acids](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

1.4: REACTIONS OF AMINO ACIDS

1.4.1: Acid-base Chemistry of Amino Acids

1.4.2: Acid-Base Reactions of Amino Acids

1.4.3: Introduction to Amino Acid Reactivity

This page titled [1.4: Reactions of Amino Acids](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

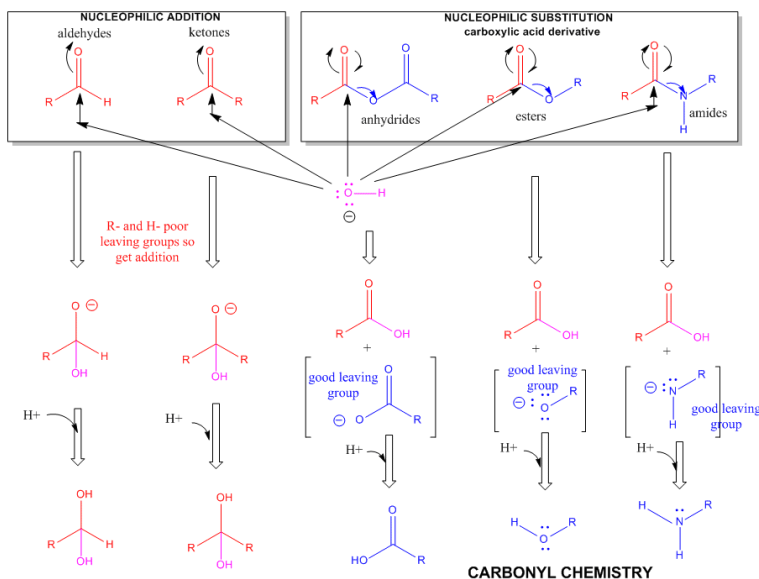
1.4.3: Introduction to Amino Acid Reactivity

You should be able to identify which side chains contain H bond donors and acceptors. Likewise, some are acids and bases. You should be familiar with the approximate pKa's of the side chains, and the N and C terminal groups. Three of the amino acid side chains (Trp, Tyr, and Phe) contribute significantly to the UV absorption of a protein at 280 nm. This section will deal predominantly with the chemical reactivity of the side chains, which is important in understanding the properties of the proteins. Many of the side chains are nucleophiles. Nucleophilicity is a measure of how rapidly molecules with lone pairs of electrons can react in nucleophilic substitution reactions. It correlates with basicity, which measures the extent to which a molecule with lone pairs can react with an acid (Bronsted or Lewis). The properties of the atom which holds the lone pair are important in determining both nucleophilicity and basicity. In both cases, the atom must be willing to share its unbonded electron pair. If the atoms holding the nonbonded pair is more electronegative, it will be less likely to share its electrons, and that molecule will be a poorer nucleophile (nu:) and weaker base. Using these ideas, it should be clear that RNH₂ is a better nucleophile than ROH, OH⁻ is a better than H₂O and RSH is a better than H₂O. In the latter case, S is bigger and its electron cloud is more polarizable - hence it is more reactive. The important side chain nucleophiles (in order from most to least nucleophilic) are Cys (RSH, pKa 8.5-9.5), His (pKa 6-7), Lys (pKa 10.5) and Ser (ROH, pKa 13).

An understanding of the chemical reactivity of the various R group side chains of the amino acids in a protein is important since chemical reagents that react specifically with a given amino acid side chain can be used to:

- identify the presence of the amino acids in unknown proteins or
- determine if a given amino acid is critical for the structure or function of the protein. For example, if a reagent that covalently interacts with only Lys is found to inhibit the function of the protein, a lysine might be considered to be important in the catalytic activity of the protein.

Figure: A REVIEW SUMMARY OF THE CHEMISTRY OF ALDEHYDES, KETONES, AND CARBOXYLIC ACID DERIVATIVES



The side chain of serine is generally no more reactive than ethanol. It is a potent nucleophile in a certain class of proteins (proteases, for example) when it is deprotonated. The amino group of lysine is a potent nucleophile only when deprotonated.

Contributors and Attributions

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

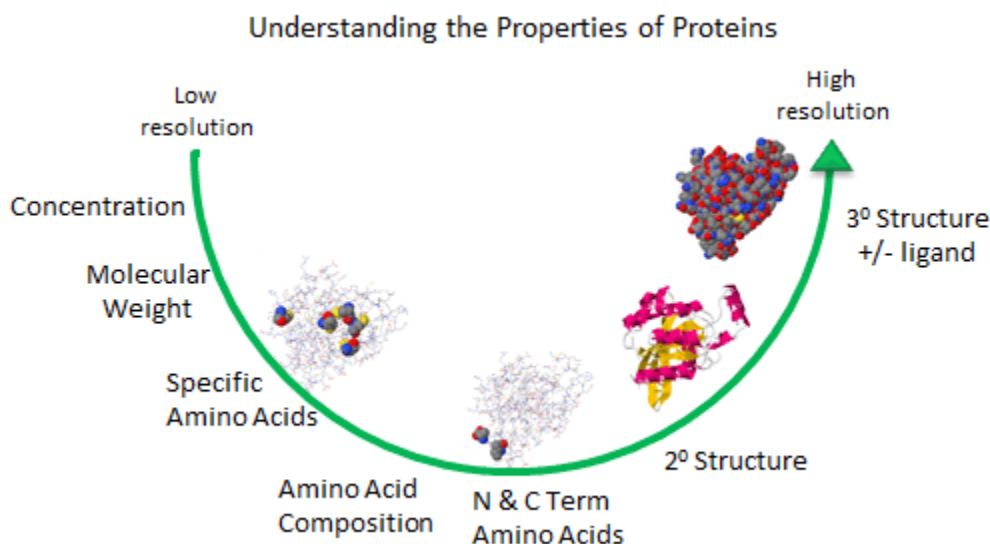
This page titled [1.4.3: Introduction to Amino Acid Reactivity](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Henry Jakubowski](#).

- [Bloggers](#) has no license indicated.

1.5: Amino Acid Analysis and Chemical Sequencing

As described in the Introduction to Proteins, we can understand proteins structure at varying level of complexity.

Figure: Protein Analysis from low to high resolution.



In the last chapter section, we learned about the charge and chemical reactivity properties of isolated amino acids and amino acids in proteins. The analysis of a whole protein is complicated since each different amino acid might be represented many times in the sequence. Each protein has an N-terminal and C-terminal amino acid and secondary structure. Some proteins exist biologically as multisubunit proteins, which adds to the complexity of the analyses since now the proteins would have multiple N- and C-terminal ends. In addition, isolated proteins might have chemical modifications (post-translational) which add to the functionalities of the proteins but also add to the complexities of the analyses. To illustrate some of these issues, view the structure of the RhoA protein below.



Updated RhoA - a cytoplasmic protein - The complexity of protein analysis Jmol14 (Java) | JSMol (HTML5)

Amino Acid Composition

At a low level of resolution, we can determine the amino acid composition of the protein by hydrolyzing the protein in 6 N HCl, 100°C, under vacuum for various time intervals. After removing the HCl, the hydrolysate is applied to an ion-exchange or hydrophobic interaction column, and the amino acids eluted and quantitated with respect to known standards. A non naturally-occurring amino acid like norleucine is added in known amounts as an internal standard to monitor quantitative recovery during the reactions. The separated amino acids are often derivitized with ninhydrin or phenylisothiocyanate to facilitate their detection. The reaction is usually allowed to proceed for 24, 36, and 48 hours, since amino acids with OH (like ser) are destroyed. A time course allows the concentration of Ser at time $t=0$ to be extrapolated. Trp is also destroyed during the process. In addition, the amide links in the side chains of Gln and Asn are hydrolyzed to form Glu and Asp, respectively.

- AA Analysis: Iowa State University Protein Facility

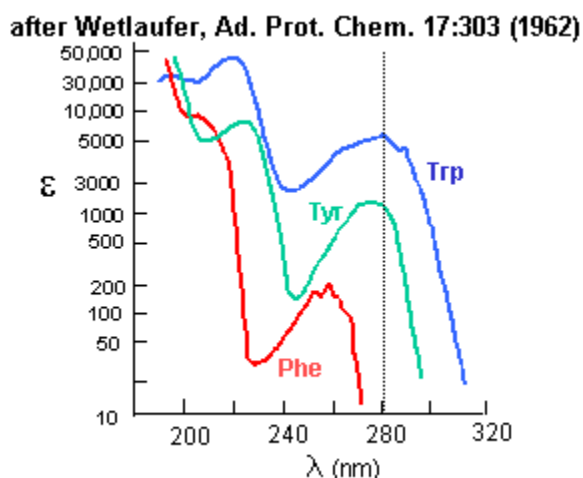
N- and C-Terminal Amino Acid Analysis

The amino acid composition does not give the sequence of the protein. The N-terminus of the protein can be determined by reacting the protein with fluorodinitrobenzene (FDNB) or dansyl chloride, which reacts with any free amine in the protein, including the epsilon amino group of lysine. The amino group of the protein is linked to the aromatic ring of the DNB through an amine and to the dansyl group by a sulfonamide, and are hence stable to hydrolysis. The protein is hydrolyzed in 6 N HCl, and the amino acids separated by TLC or HPLC. Two spots should result if the protein was a single chain, with some Lys residues. The labeled amino acid other than Lys is the N-terminal amino acid. The C-terminal amino acid can be determined by addition of carboxypeptidases, enzymes which cleave amino acids from the C-terminal. A time course must be done to see which amino acid is released first. N-terminal analysis can also be done as part of sequencing the entire protein as discussed below (Edman degradation reaction).

Analysis for Specific Amino Acids

Aromatic amino acids can be detected by their characteristic absorbance profiles. Amino acids with specific functional groups can be determined by chemical reactions with specific modifying groups, as shown in section 2A.

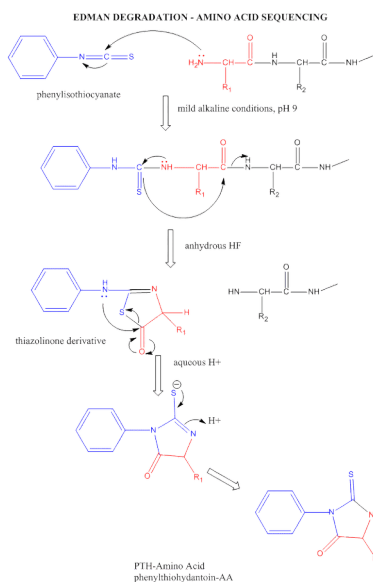
Figure: amino acid absorbance profiles



Amino Acid Sequence - Edman Degradation

Two methods exist to determine the entire sequence of a protein. In one, the protein is sequenced; in the other, the DNA encoding the protein is sequenced, from which the amino acid sequence can be derived. The actually protein can be sequenced by automated, sequential Edman Degradation.

Figure: Edman Degradation



In this technique, a protein adsorbed to a solid phase reacts with phenylisothiocyanate. An intramolecular cyclization and cleavage of the N-terminal amino acid results, which can be washed from the adsorbed protein and detected by HPLC analysis. The yields in this technique are close to 100%. However, with time, more chains accumulate in which an N-terminal amino acid has not been removed. If it is removed on the next step, two amino acids will elute, creating increasing "noise" in the elution step - i.e. more than 1 amino acid derivative will be detected. Hence the maximal length of the peptide which can be sequenced is about 50 amino acids. Most proteins are larger than that. Hence, before the protein can be sequenced, it must be cleaved with specific enzymes called endoproteases which cleave proteins after specific side chains. For example, trypsin cleaves proteins within a chain after Lys and Arg, while chymotrypsin cleaves after aromatic amino acids, like Trp, Tyr, and Phe. Chemical cleavage by small molecules can be used as well. Cyanogen bromide, CNBr, cleaves proteins after methionine side chains. The individual proteins must be cleaved using two different methods, and each peptide fragment isolated and sequenced. Then the order of the cleaved peptides with known sequence can be pieced together by comparing the peptide sequences obtained using different cleavage methods. Many proteins also have disulfide bonds connecting Cys side chains distal to each other in the polypeptide chain. Proteolytic or chemical cleavage of the protein would lead to the formation of a fragment containing two peptides linked by disulfides. Edman degradation would release two amino acids from such fragments. To avoid this problem, the protein is oxidized with performic acid, which irreversibly oxidizes free Cys, or Cys-Cys disulfides to cysteic acid residues. A summary of the steps involved in protein sequencing are shown below:

PROTEIN SEQUENCING STRATEGY - 8 STEPS

1. If the protein contains more than one polypeptide chain, the chains are separated and purified. If disulfide bonds connect two different chains, the S-S bond must be cleaved (as described in step 2) and each peptide independently purified.
2. Intrachain S-S bonds between Cys side chains are cleaved with performic acid. (See above for interchain S-S bonds).
3. The amino acid composition of each chain is determined
4. The N-terminal and C-terminal residues are identified.
5. Each polypeptide chain is cleaved into smaller fragments, and the amino acid composition and sequence of each fragment is determined.
6. Step 5 is repeated, using a different cleavage procedure to generate a different and overlapping set of peptide fragments.
7. The overall amino acid sequence of the protein is reconstructed from the sequences in overlapping fragments.
8. The position of the S-S is located. (See online problem set - Proteins)

Contributors and Attributions

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

This page titled [1.5: Amino Acid Analysis and Chemical Sequencing](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Henry Jakubowski](#).

1.5.1: Polypeptides and Proteins

Learning Objectives

1. Define or describe the following:
 - a. amino acid
 - b. "R" group
 - c. peptide bond
 - d. peptide
 - e. polypeptide
 - f. primary protein structure
 - g. secondary protein structure
 - h. tertiary protein structure
 - i. quaternary protein structure
 - j. gene
2. Describe how the primary structure of a protein or polypeptide ultimately determines its final three-dimensional shape.
3. Describe how the order of nucleotide bases in DNA ultimately determines the final three-dimensional shape of a protein or polypeptide.

Amino acids are the building blocks for proteins. All amino acids contain an amino or NH_2 group and a carboxyl (acid) or COOH group. There are 20 different amino acids commonly found in proteins and often 300 or more amino acids per protein molecule. Each amino acid differs in terms of its "R" group. The "R" group of an amino acid is the remainder of the molecule, that is, the portion other than the amino group, the acid group, and the central carbon. Each different amino acid has a unique "R" group and the unique chemical properties of an amino acid depend on that of its "R" group (Figure 1.5.1.1).

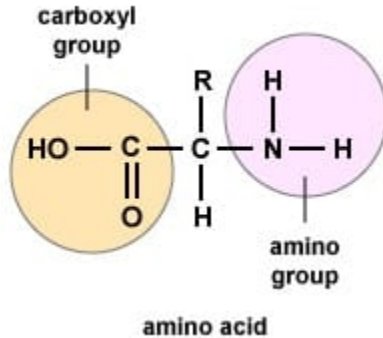


Figure 1.5.1.1: Amino Acids. Structure of an amino acid.

To form polypeptides and proteins, amino acids are joined together by peptide bonds, in which the amino or NH_2 of one amino acid bonds to the carboxyl (acid) or COOH group of another amino acid as shown in (Figure 1.5.1.2 and Figure 1.5.1.3).

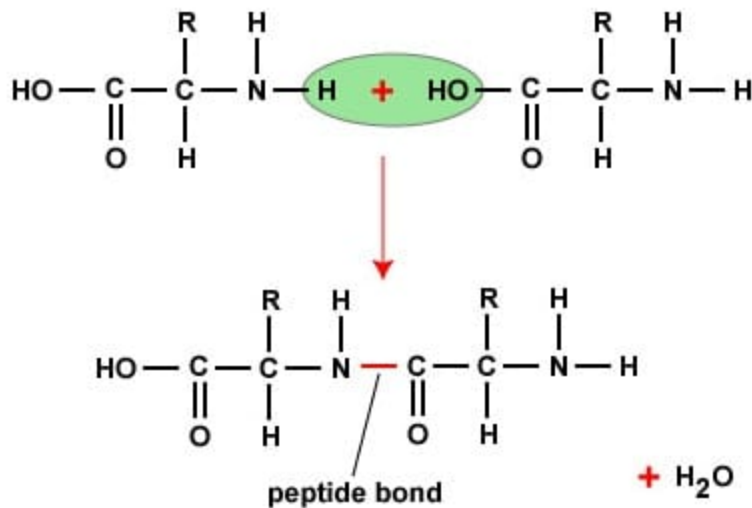


Figure 1.5.1.2: Peptide Bonds. A peptide bond forms when the amino group of one amino acid bonds to the carboxyl group of another amino acid.

A peptide is two or more amino acids joined together by peptide bonds, and a polypeptide is a chain of many amino acids. A protein contains one or more polypeptides. Therefore, proteins are long chains of amino acids held together by peptide bonds.

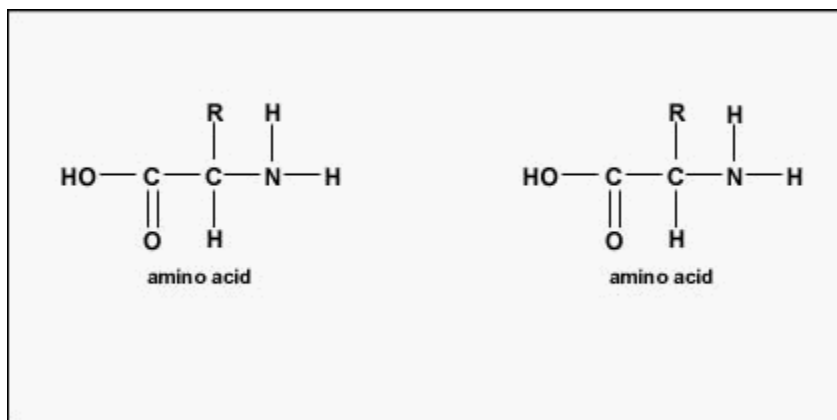


Figure 1.5.1.3: Formation of a Peptide Bond. A peptide bond forms when the amino group of one amino acid bonds to the carboxyl group of another amino acid.

The actual order of the amino acids in the protein is called its primary structure (Figure 1.5.1.4) and is determined by DNA. As will be seen later in this unit, DNA is divided into functional units called genes. A gene is a sequence of deoxyribonucleotide bases along one strand of DNA that codes for a functional product - a specific molecule of messenger RNA, transfer RNA, or ribosomal RNA. The product is usually messenger RNA (mRNA) and mRNA ultimately results in the synthesis of a polypeptide or a protein. Therefore, it is commonly said that the order of deoxyribonucleotide bases in a gene determines the amino acid sequence of a particular protein. Since certain amino acids can interact with other amino acids in the same protein, this primary structure ultimately determines the final shape and therefore the chemical and physical properties of the protein.

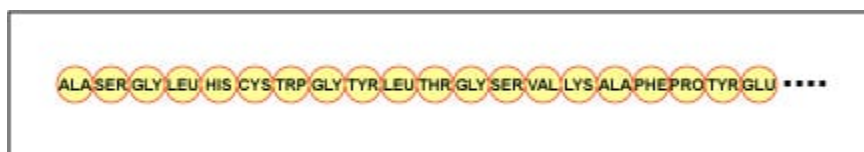
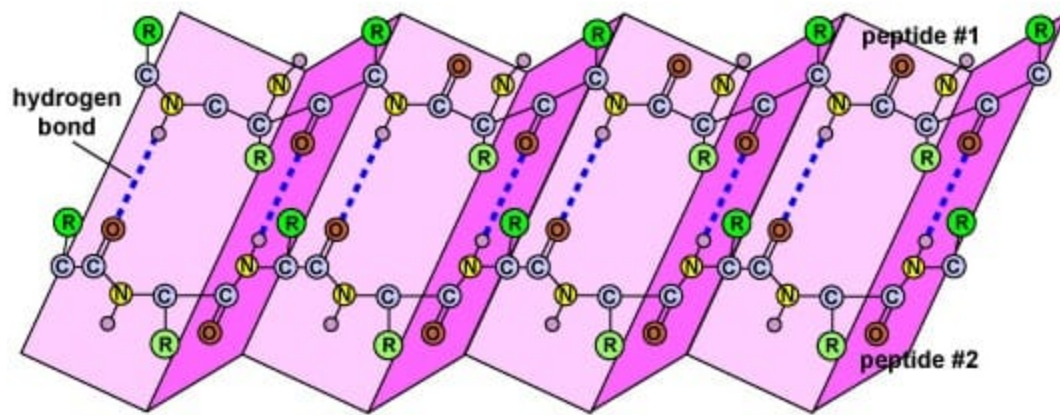


Figure 1.5.1.4: Primary Structure of a Protein or Polypeptide. The primary structure of a protein or polypeptide is the actual sequence of its amino acids. Primary structure is determined by the order of the deoxyribonucleotide bases in genes.

The secondary structure of the protein is due to hydrogen bonds that form between the oxygen atom of one amino acid and the nitrogen atom of another. This gives the protein or polypeptide the two-dimensional form of an alpha-helix or a beta-pleated sheet (Figure 1.5.1.5).



beta pleated sheet

Figure 1.5.1.5: Secondary

Structure of a Protein or Polypeptide. (left) The secondary structure of a protein or polypeptide is due to hydrogen bonds forming between an oxygen atom of one amino acid and a nitrogen atom of another. There are two possible types of secondary structure: an alpha helix and a beta sheet. In the case of an alpha helix, the hydrogen bonding causes the polypeptide to twist into a helix. With a beta sheet the hydrogen bonding enables the polypeptide to fold back and forth upon itself like a pleated sheet. (right) The secondary structure of a protein or polypeptide is due to hydrogen bonds forming between an oxygen atom of one amino acid and a nitrogen atom of another. There are two possible types of secondary structure: an alpha helix and a beta sheet. In the case of an alpha helix, the hydrogen bonding causes the polypeptide to twist into a helix. With a beta sheet the hydrogen bonding enables the polypeptide to fold back and forth upon itself like a pleated sheet.

In globular proteins such as enzymes, the long chain of amino acids becomes folded into a three-dimensional functional shape or tertiary structure. This is because certain amino acids with sulfhydryl or SH groups form disulfide (S-S) bonds with other amino acids in the same chain. Other interactions between R groups of amino acids such as hydrogen bonds, ionic bonds, covalent bonds, and hydrophobic interactions also contribute to the tertiary structure (Figure 1.5.1.6). In some proteins, such as antibody molecules and hemoglobin, several polypeptides may bond together to form a quaternary structure (Figure 1.5.1.7).

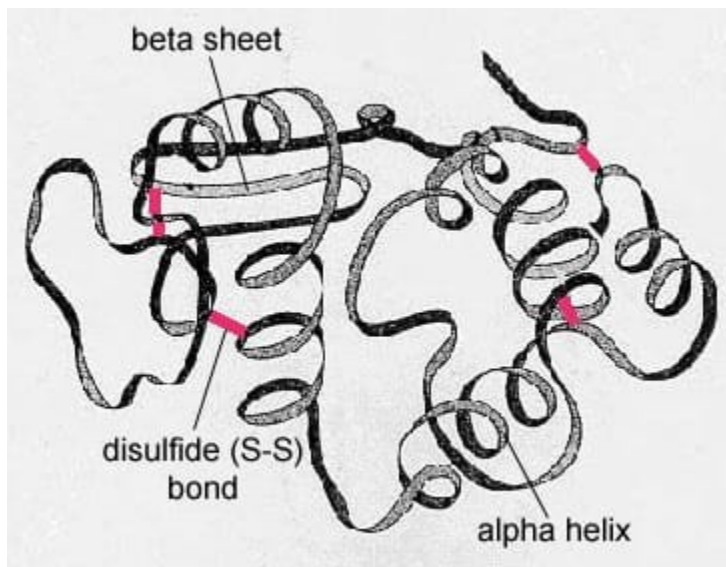


Figure 1.5.1.6: Tertiary Structure of a Protein or Polypeptide.

In globular proteins such as enzymes, the long chain of amino acids becomes folded into a three-dimensional functional shape or tertiary structure. This is because certain amino acids with sulfhydryl or SH groups form disulfide (S-S) bonds with other amino acids in the same chain. Other interactions between R groups of amino acids such as hydrogen bonds, ionic bonds, covalent bonds, and hydrophobic interactions also contribute to the tertiary structure.

As will be seen later in this unit, during protein synthesis, the order of nucleotide bases along a gene gets transcribed into a complementary strand of mRNA which is then translated by tRNA into the correct order of amino acids for that polypeptide or protein. Therefore, the order of deoxyribonucleotide bases along the DNA determines the order of amino acids in the proteins.

Because certain amino acids can interact with other amino acids, the order of amino acids for each protein determines its final three-dimensional shape, which in turn determines the function of that protein (e.g., what substrate an enzyme will react with, what epitopes the Fab of an antibody will combine with, what receptors a cytokine will bind to).

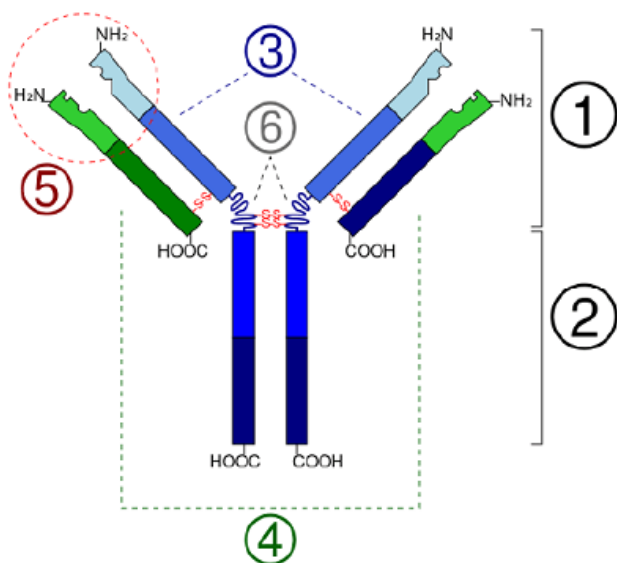


Figure 1.5.1.7: Quaternary Structure of a Protein. The quaternary structure of a protein is due to several polypeptides joining together, as in the case of antibody molecules. Schematic diagram of the basic unit of immunoglobulin (antibody) Fab Fc heavy chain (consist of VH, CH1, hinge, CH2 and CH3 regions: from N-term) light chain (consist of VL and CL regions: from N-term) antigen binding site hinge regions (*) -S-S- mean disulfide bonds. (CC-SA-BY 3.0; Y_tambe).

Summary

1. Amino acids are the building blocks for proteins. There are 20 different amino acids commonly found in proteins and often 300 or more amino acids per protein molecule.
2. All amino acids contain an amino or NH_2 group and a carboxyl (acid) or COOH group.
3. To form polypeptides and proteins, amino acids are joined together by peptide bonds, in which the amino or NH_2 of one amino acid bonds to the carboxyl (acid) or COOH group of another amino acid.
4. A peptide is two or more amino acids joined together by peptide bonds; a polypeptide is a chain of many amino acids; and a protein contains one or more polypeptides. Therefore, proteins are long chains of amino acids held together by peptide bonds.
5. The actual order of the amino acids in the protein is called its primary structure and is determined by DNA.
6. The order of deoxyribonucleotide bases in a gene determines the amino acid sequence of a particular protein. Since certain amino acids can interact with other amino acids in the same protein, this primary structure ultimately determines the final shape and therefore the chemical and physical properties of the protein.
7. The secondary structure of the protein is due to hydrogen bonds that form between the oxygen atom of one amino acid and the nitrogen atom of another and gives the protein or polypeptide the two-dimensional form of an alpha-helix or a beta-pleated sheet.
8. In globular proteins such as enzymes, the long chain of amino acids becomes folded into a three-dimensional functional shape or tertiary structure. This is because certain amino acids with sulfhydryl or SH groups form disulfide (S-S) bonds with other amino acids in the same chain. Other interactions between R groups of amino acids such as hydrogen bonds, ionic bonds, covalent bonds, and hydrophobic interactions also contribute to the tertiary structure.
9. In some proteins, such as antibody molecules, several polypeptides may bond together to form a quaternary structure.

1.5.2: Sequence Determination Using Mass Spectrometry

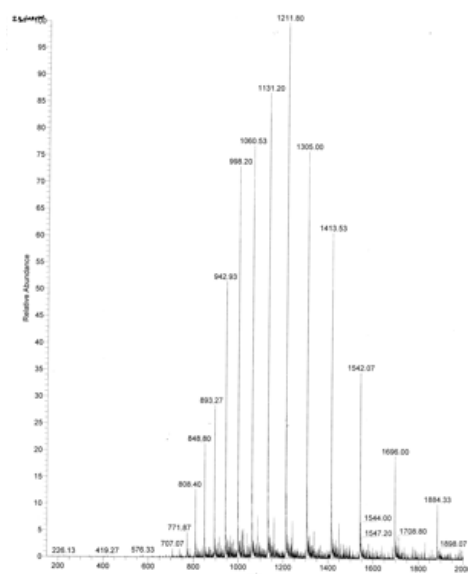
Mass spectrometry is supplanting more traditional methods (see above) as the choice to determine the molecular mass and structure of a protein. Its power comes from its exquisite sensitivity and modern computational methods to determine structure through comparisons of ion fragment data with computer databases of known protein structures. In mass spectrometry, a molecule is first ionized in an ion source. The charged particles are then accelerated by an electric field into a mass analyzer where they are subjected to an external magnetic field. The external magnetic field interacts with the magnetic field arising from the movement of the charged particles, causing them to deflect. The deflection is proportional to the mass to charge ratio, m/z . Ions then enter the detector which is usually a photomultiplier. Sample introduction into the ion source occurs through simple diffusion of gases and volatile liquids from a reservoir, by injection of a liquid sample containing the analyte by spraying a fine mist, or for very large proteins by desorbing a protein from a matrix using a laser. Analysis of complex mixtures is done by coupling HPLC with mass spectrometry in a LCMS.

Ion source: There are many methods to ionize molecules, including atmospheric pressure chemical ionization (APCI), chemical ionization (CI), or electron impact (EI). The most common methods for protein/peptide analyses are electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI).

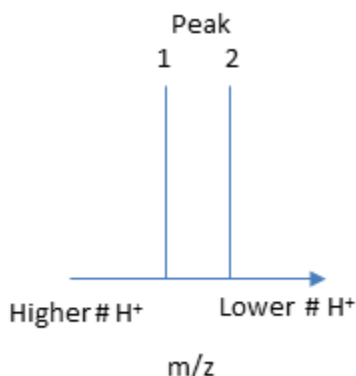
Electrospray ionization (ESI) - The analyte, dissolved in a volatile solvent like methanol or acetonitrile, is injected through a fine stainless steel capillary at a slow flow rate into the ion source. A high voltage (3-4 kV) is maintained on the capillary giving it a positive charge with respect to the other oppositely charged electrode. The flowing liquid becomes charged with same polarity as the polarity of the positively charged capillary. The high field leads to the emergence of the sample as a charged aerosol spray of charged microdrops which reduces electrostatic repulsions in the liquid. This method essentially uses electrical energy to produce the aerosol instead of mechanical energy to produce a liquid aerosol, as in the case of a perfume atomizer. Surrounding the capillary is a flowing gas (nitrogen) which helps to move the aerosol towards the mass analyzer. The microdrops become smaller in size as the volatile solvent evaporates, increasing the positive charge density on the drops. Eventually electrostatic repulsions cause the drops to explode in a series of steps, ultimately producing analyte devoid of solvent. This gentle method of ionization produces analytes that are not cleaved but ready for introduction into the mass analyzer. Protein emerge from this process with a roughly Gaussian distribution of positive charges on basic side chains. In organic chemistry you studied mass spectrums of small molecules induced by electron bombardment. This produces ions of +1 charge as an electron is stripped away from the neutral molecule. The highest m/z peak in the spectrum is the parent ion or M^+ ion. The highest m/z ratio detectable in the mass spectrum is in the thousands. However, large peptides and proteins with large molecular masses can be detected and resolved since the charge on the ions are great than +1. In 2002, John Fenn was awarded a Noble Prize in Chemistry for the development and use of ESI to study biological molecules.

An example of an ESI spectrum of apo-myoglobin is shown below. Note the roughly Gaussian distribution of the peaks, each of which represents the intact protein with charges differing by +1. Protein have positive charges by virtue of both protonation of amino acid side chains as well as charges induced during the electrostray process itself. Based on the amino acid sequence of myoglobin and the assumption that the pK_a of the side chains are the same in the protein as for isolated amino acids, the calculated average net charges of apoMb would be approximately +30 at pH 3.5, +20 at pH 4.5, +9 at pH 6, and 0 at pH 7.8 (the calculated pI). The mass spectrum below was taken by direct injection into the MS of apoMb in 0.1% formic acid (pH 2.8). Charges on the peptide are a combined results of charges present on the peptide before the electrospray and changes in charges induced during the process. www.chm.bris.ac.uk/ms/theory/...onisation.html.

Figure: ESI Mass Spectrum of Apo-Myoglobin



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



If M is the molecular mass of the analyte protein, and n is the number of positive charges on the protein represented in a given m/z peak, then the following equations give the molecular mass M of the protein for each peak:

$$M_{peak2} = n(m/z)_{peak2} - n(1.008) \quad (1.5.2.1)$$

$$M_{peak1} = (n + 1)(m/z)_{peak1} - (n + 1)(1.008) \quad (1.5.2.2)$$

where 1.008 is the atomic weight of H. Since there is only one value of M , the two equations can be set equal to each other, giving:

$$n(m/z)_{peak2} - n(1.008) = (n + 1)(m/z)_{peak1} - (n + 1)(1.008) \quad (1.5.2.3)$$

Solving for n gives:

$$n = [(m/z)_{peak1} - 1.008] / [(m/z)_{peak2} - (m/z)_{peak1}]. \quad (1.5.2.4)$$

Knowing n , the molecular mass M the protein can be calculated for each m/z peak. The best value of M can then be determined by averaging the M values determined from each peak (16,956 from the above figure). For peaks from m/z of 893-1542, the calculated values of n ranged from +18 to +10.

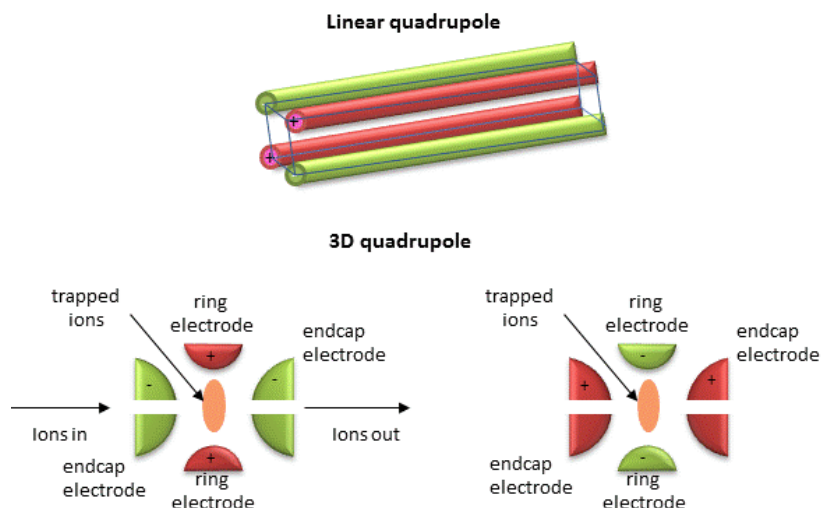
Matrix assisted laser desorption ionization (MALDI): In this technique, used for larger biomolecules like proteins and polysaccharides, the analyte is mixed with an absorbing matrix material. Laser excitation is used to excite the matrix, leading to energy transfer that results in ionization and "launching" of the matrix and analyte in ion form from the solid mixture. Parent ion

peaks of $(M+H)^+$ and $(M-H)^-$ are formed.

Mass Analyzer

Quadrupole ion trap (used in ESI) - A complex mixture of ions can be contained (or trapped) in this type of mass analyzer. Two common type are linear and 3D quadrupoles.

Figure: Linear and 3D Quadrupoles



As dipoles display positive and negative charge separation on a linear axis, quadrupoles have either opposite electrical charges or opposite magnetic fields at the opposing ends of a square or cube. In charge separation, the monopole (sum of the charges) and dipoles cancel to zero, but the quadrupole moment does not. The quadrupole traps ions using a combination of fixed and alternating electric fields. The trap contains He at 1 mTorr. For the 3D trap, The ring electrode has a oscillating RF voltage which keeps the ions trapped. The end caps also have an AC voltage. Ions oscillate in the trap with a "secular" frequency determined by the frequency of the RF voltage, and of course, the m/z ratio. By increasing the the amplitude of the RF field across the ring electron, ion motion in the trap becomes destabilized and leads to ion ejection into the detector. When the secular frequency of ion motion matches the applied AC voltage to the endcap electrodes, resonance occurs and the amplitude of motion of the ions increases, also allowing leakage out of the ion trap into the detector.

- [EXTERNAL](#) Ion Trap Tutorial
- [EXTERNAL](#) Ion Trap Animation (requires download) from Thermo Scientific

Time of Flight (TOF) tube (used in MALDI) - a long tubes is used and the time required for ion detection is determined. The small molecular mass ions take the shortest time to reach the detector.

Tandem Mass Spectrometry (MS/MS)

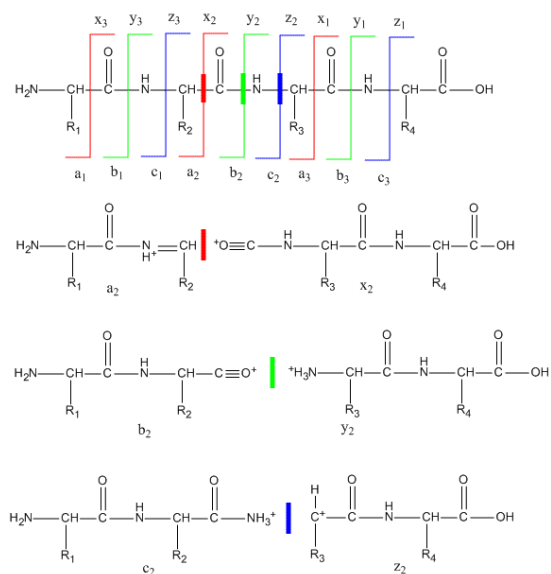
Quadrupole mass analyzers which can select ions of varying m/z ratios in the ion traps are commonly used in for tandem mass spectrometry (MS/MS). In this technique, the selected ions are further fragmented into smaller ions by a process called collision induced dissociation (CID). When performed on all of the initial ions present in the ion trap, the sequence of a peptide/protein can be determined. This techniques usually requires two mass analyzers with a collision cell in-between where selected ions are fragmented by collision with an inert gas. It can also be done in a single mass analyzer using a quadrupole ion trap.

In a typical MS/MS experiment to determine a protein sequence, a protein is cleaved into protein fragments with an enzyme such as trypsin, which cleaves on carboxyl side of positively charge Lys and Arg side chains. The average size of proteins in the human proteome is approximately 50,000. If the average molecular mass of an amino acid in a protein is around 110 (18 subtracted since water is released on amide bond formation), the average number of amino acids in the protein would be around 454. If 10%

of the amino acids are Arg and Lys, the on average there would be approximately 50 Lys and Arg, and hence 50 tryptic peptides of average molecular mass of 1000. The fragments are introduced in the MS where a peptide fragment fingerprint analysis can be performed. The MWs of the fragments can be identified and compared to known peptide digestion fragments from known proteins to identify the analyte protein.

To get sequence information, a tryptic peptide with a specific m/z ratio (optimally with a single +1 charge) is further selected in the ion trap and fragmented on collision with an inert agent (MS/MS). Since the m/e range of mass spectrometers is in the thousands, tryptic fragments with a single charge can easily be detected and targeted for MS/MS. The likely and observed cleavages for a tetrapeptide and the resulting ions with a +1 charge are illustrated below. Ions with the original N terminus are denoted as a, b, and c, while ions with the original C terminus are denoted as x, y, and z. c and y ions gain an extra proton from the peptide to form positively charged $-NH_3^+$ groups. The actual ions observed depend on many factors including the sequence of the peptide, its original charge, the energy of the collision inducing the fragmentation, etc. Low energy fragmentation of peptides in ion traps usually produce a, b, and y ions, along with peaks resulting from loss of NH_3 (a^* , b^* and y^*) or H_2O (ao , bo and yo). No peaks resulting from fragmentation of side chains are observed. Fragmentation at two sites in the peptide (usually at b and y sites in the backbone) form an internal fragment.

Figure: Peptide Fragmentation and Sequencing by MS/MS



The y1 peak represents the C-terminal Lys or Arg (in this example) of the tryptic peptide. Peak y2 has one additional amino acid compared to y1 and the molecular mass difference identifies the extra amino acid. Peak y3 is likewise one amino acid larger than y2. All three y fragment peaks have a common Lys/Arg C-terminal and the charged fragment contains the C-terminal end of the original peptide. All b fragment peaks for a given peptide contain a common N terminal amino acid with b1 the smallest. Note that the subscript represents the number of amino acids in the fragment. By identifying b and y peaks the actual sequence of small peptide can be determined. Usually spectra are matched to databases to identify the structure of each peptide and ultimately that of the protein. The actual m values for fragments can be calculated as follows, where (N) is the molecular mass of the neutral N terminal group, (C) is the molecular mass of the neutral c terminal group, and (M) is the molecule mass of the neutral amino acids.

- a: $(N)+(M)-CHO$
- b: $(N)+(M)$
- y: $(C)+(M)+H$ (note in the figure above that the amino terminus of the y peptides has an extra proton in the +1 charged peptides.)

m/z values can be calculated from the calculated m values and by adding the one H mass to the overall z if the overall charge is +1, etc.

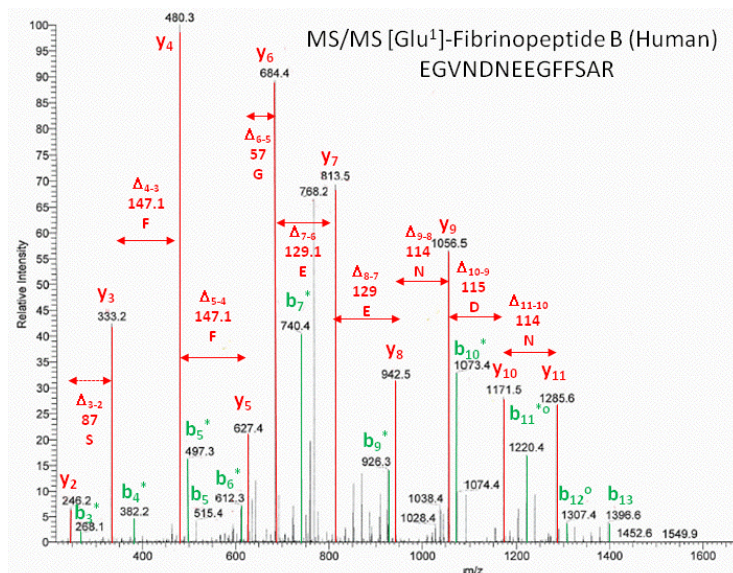
Table: Masses of amino acid residues in a protein. (For N terminal amino acid, add 1 H. for C terminus add OH)

Residue	Code	Monoisotopic Mass	Average Mass
Ala	A	71.03714	71.0779
Arg	R	156.101111	156.1857
Asn	N	114.042927	114.1026
Asp	D	115.026943	115.0874
Cys	C	103.009185	103.1429
Glu	E	129.042593	129.114
Gln	Q	128.058578	128.1292
Gly	G	57.021464	57.0513
His	H	137.058912	137.1393
Ile	I	113.084064	113.1576
Leu	L	113.084064	113.1576
Lys	K	128.092963	128.1723
Met	M	131.040485	131.1961
Phe	F	147.068414	147.1739
Pro	P	97.052764	97.1152
Ser	S	87.032028	87.0773

Residue	Code	Monoisotopic Mass	Average Mass
Thr	T	101.047679	101.1039
Trp	W	186.079313	186.2099
Tyr	Y	163.06332	163.1733
Val	V	99.068414	99.1311

As an example, using these MW values, the sequence of the human Glu1- fibrinopeptide B can be determined from MS/MS spectra shown in an annotated form below. Note that most of the b peaks are b* resulting from loss of NH₃ from the N terminus.

Figure: Annotated MS/MS spectra of human Glu1- fibrinopeptide B



- EXTERNAL [MS review from Organic Chemistry](#)
- EXTERNAL [Mass Spectroscopy Educational Resources](#)
- EXTERNAL [Mass Spec Resources](#)
- EXTERNAL [Sequence Determination: Animation](#)
- EXTERNAL [An Introduction to Mass Spectrometry](#)
- EXTERNAL [Ion Traps](#)
- EXTERNAL [What is electrospray?](#)
- EXTERNAL [Protein Prospector: Proteomics tools for mining sequence databases in conjunction with Mass Spectrometry experiments](#)
- EXTERNAL [Proteomics Research for Integrative Biology](#)
- EXTERNAL [Mascot from Matrix Science: a powerful search engine which uses mass spectrometry data to identify proteins from primary sequence databases](#)

-  Peptide Mass Fingerprinting: A tutorial

Contributors and Attributions

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

This page titled [1.5.2: Sequence Determination Using Mass Spectrometry](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Henry Jakubowski](#).

1.5.3: The Edman Degradation

Objectives

After completing this section, you should be able to

1. describe how an Edman degradation is used to determine the sequence of the amino acid residues in peptides containing up to 20 such residues.
2. describe, briefly, how the procedure is modified to deal with peptides and proteins containing more than 20 amino acid residues.
3. write a detailed mechanism for the Edman degradation.
4. determine the structure of a peptide, given a list of the fragments that are produced by a partial acid hydrolysis.
5. determine the structure of a peptide, given a list of the fragments that are produced when the peptide is cleaved by a specific enzyme and the details of the types of bonds cleaved by that enzyme.
6. predict the fragments that would be produced when a peptide of known structure is cleaved by a specific enzyme, given sufficient information about the types of bonds that are cleaved by the enzyme in question.

Key Terms

Make certain that you can define, and use in context, the key term below.

- Edman degradation

Study Notes

The reagent used in the Edman degradation is phenyl isothiocyanate. You may find it helpful to review the relationship between cyanates, isocyanates, thiocyanates and isothiocyanates.

$R-O-C\equiv N$ cyanate (such compounds do not exist)

$R-N=C=O$ isocyanate

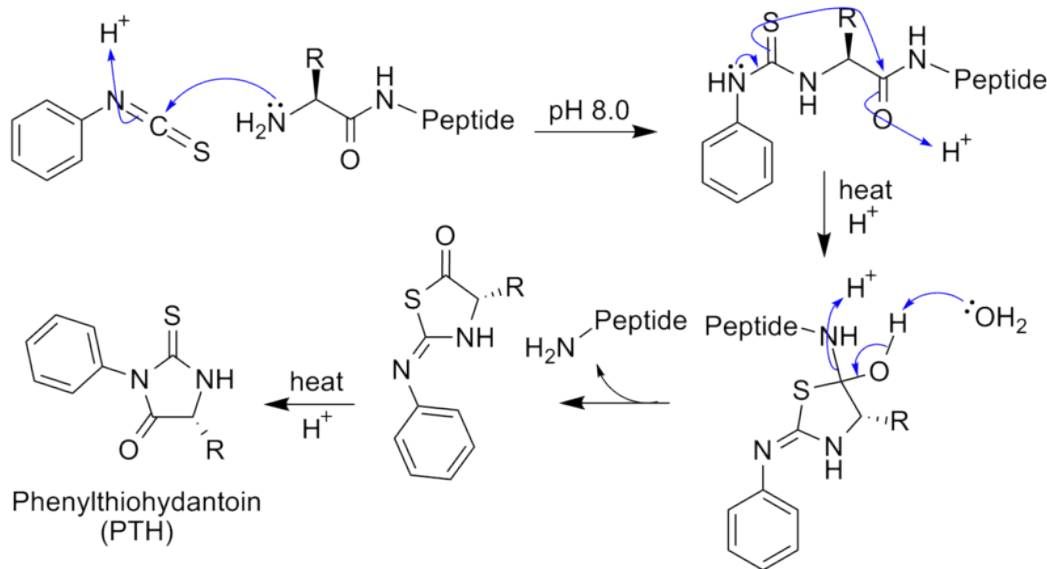
$R-S-C\equiv N$ thiocyanate

$R-N=C=S$ isothiocyanate

You need not memorize the specific peptide bonds that are broken by the enzymes trypsin and chymotrypsin.

Edman degradation is the process of purifying protein by sequentially removing one residue at a time from the amino end of a peptide. To solve the problem of damaging the protein by hydrolyzing conditions, Pehr Edman created a new way of labeling and cleaving the peptide. Edman thought of a way of removing only one residue at a time, which did not damage the overall sequencing. This was done by adding Phenyl isothiocyanate, which creates a phenylthiocarbamoyl derivative with the N-terminal.

The N-terminal is then cleaved under less harsh acidic conditions, creating a cyclic compound of phenylthiohydantoin PTH-amino acid. This does not damage the protein and leaves two constituents of the peptide. This method can be repeated for the rest of the residues, separating one residue at a time.



Edman degradation is very useful because it does not damage the protein. This allows sequencing of the protein to be done in less time. Edman sequencing is done best if the composition of the amino acid is known. As we saw in Section 26.5, to determine the composition of the amino acid, the peptide must be hydrolyzed. This can be done by denaturing the protein and heating it and adding HCl for a long time. This causes the individual amino acids to be separated, and they can be separated by ion exchange chromatography. They are then dyed with ninhydrin and the amount of amino acid can be determined by the amount of optical absorbance. This way, the composition but not the sequence can be determined.

Sequencing Larger Proteins

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

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

For example, trypsin can be used on the initial peptide to cleave it at the carboxyl side of arginine and lysine residues. Using trypsin to cleave the protein and sequencing them individually with Edman degradation will yield many different individual results. Although the sequence of each individual cleaved amino acid segment is known, the order is scrambled. Chymotrypsin, which cleaves on the carboxyl side of aromatic and other bulky nonpolar residues, can be used. The sequence of these segments overlap with those of the trypsin. They can be overlapped to find the original sequence of the initial protein. However, this method is limited in analyzing larger sized proteins (more than 100 amino acids) because of secondary hydrogen bond interference. Other weak intermolecular bonding such as hydrophobic interactions cannot be properly predicted. Only the linear sequence of a protein can be properly predicted assuming the sequence is small enough.

Contributors and Attributions

- Wikibooks (Structural Biopchemistry). The content on this page is licensed under a CC-SA-BY 3.0 license.

This page titled [1.5.3: The Edman Degradation](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [26.7: The Edman Degradation](#) by Dietmar Kennepohl, Steven Farmer is licensed [CC BY-SA 3.0](#).

2: Protein Structure

Learning Objectives

- describe the differences between primary, secondary, supersecondary, tertiary, quaternary and domain protein structure
- explain the basis of CD measurements for secondary structure
- describe the similarities between torsion angles and an energy vs torsion angle plot for the rotation of the C2-C3 torsion angle with phi/psi angles of peptide bonds and the 2D plots of allowed conformations around a given amino acid in a protein (Ramachandran plot).
- (from reading give explanation for observed propensities of amino acids for different secondary structure)

Template:HideTOC

Thumbnail: Structure of human hemoglobin. The proteins α and β subunits are in red and blue, and the iron-containing heme groups in green. From PDB: 1GZX. (GNU; Proteopedia Hemoglobin).

Contributors and Attributions

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

This page titled [2: Protein Structure](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Henry Jakubowski](#).

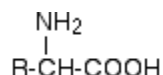
2.1: The Structure of Proteins

This page explains how amino acids combine to make proteins and what is meant by the primary, secondary and tertiary structures of proteins. Quaternary structure isn't covered. It only applies to proteins consisting of more than one polypeptide chain.

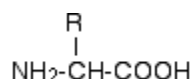
The primary structure of proteins

Drawing the amino acids

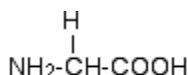
In chemistry, if you were to draw the structure of a general 2-amino acid, you would probably draw it like this:



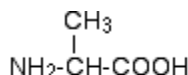
However, for drawing the structures of proteins, we usually twist it so that the "R" group sticks out at the side. It is much easier to see what is happening if you do that.



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



glycine

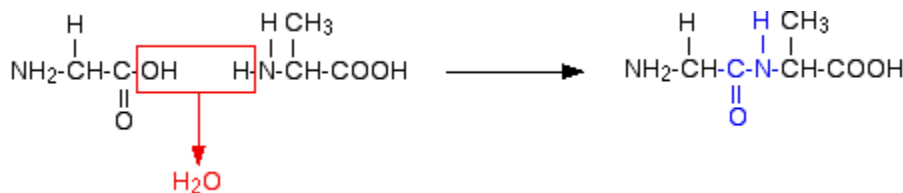


alanine

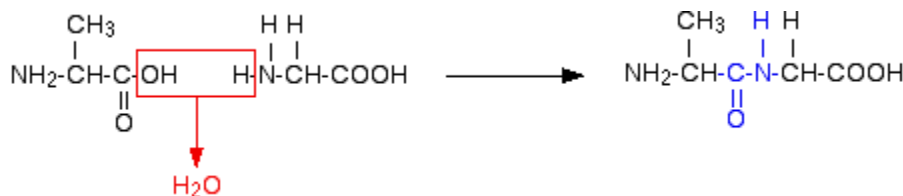
Peptides and polypeptides

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

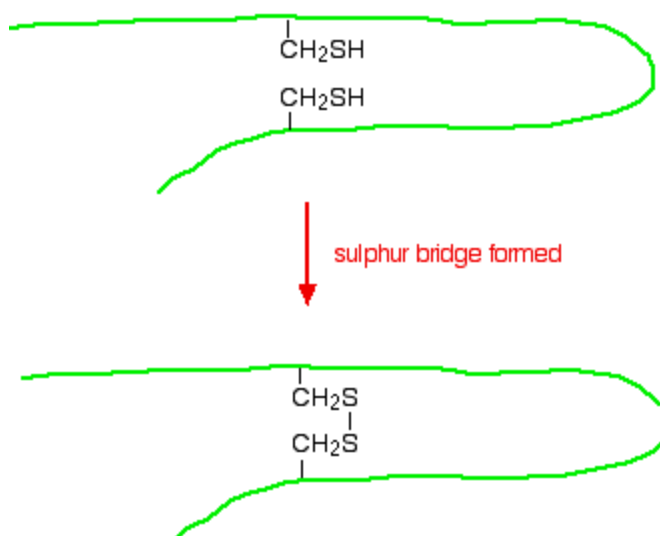
Either:



Or:



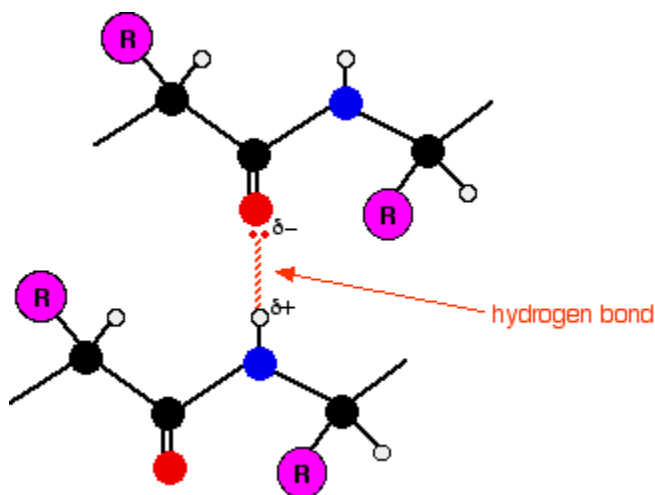
In each case, the linkage shown in blue in the structure of the dipeptide is known as a peptide link. In chemistry, this would also be known as an amide link, but since we are now in the realms of biochemistry and biology, we'll use their terms.



Because of the way sulphur bridges affect the way the protein folds, other people count this as a part of the tertiary structure (see below). This is obviously a potential source of confusion!

The secondary structure of proteins

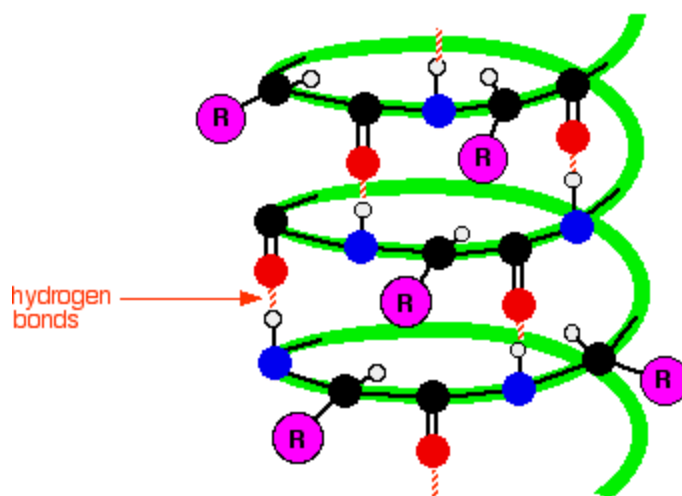
Within the long protein chains there are regions in which the chains are organised into regular structures known as alpha-helices (alpha-helices) and beta-pleated sheets. These are the secondary structures in proteins. These secondary structures are held together by hydrogen bonds. These form as shown in the diagram between one of the lone pairs on an oxygen atom and the hydrogen attached to a nitrogen atom:



Although the hydrogen bonds are always between C=O and H-N groups, the exact pattern of them is different in an alpha-helix and a beta-pleated sheet. When you get to them below, take some time to make sure you see how the two different arrangements works.

The alpha-helix

In an alpha-helix, the protein chain is coiled like a loosely-coiled spring. The "alpha" means that if you look down the length of the spring, the coiling is happening in a clockwise direction as it goes away from you. The next diagram shows how the alpha-helix is held together by hydrogen bonds. This is a very simplified diagram, missing out lots of atoms. We'll talk it through in some detail after you have had a look at it.



What's wrong with the diagram? Two things:

First of all, only the atoms on the parts of the coils facing you are shown. If you try to show all the atoms, the whole thing gets so complicated that it is virtually impossible to understand what is going on.

Secondly, I have made no attempt whatsoever to get the bond angles right. I have deliberately drawn all of the bonds in the backbone of the chain as if they lie along the spiral. In truth they stick out all over the place. Again, if you draw it properly it is virtually impossible to see the spiral.

So, what do you need to notice?

Notice that all the "R" groups are sticking out sideways from the main helix.

Notice the regular arrangement of the hydrogen bonds. All the N-H groups are pointing upwards, and all the C=O groups pointing downwards. Each of them is involved in a hydrogen bond.

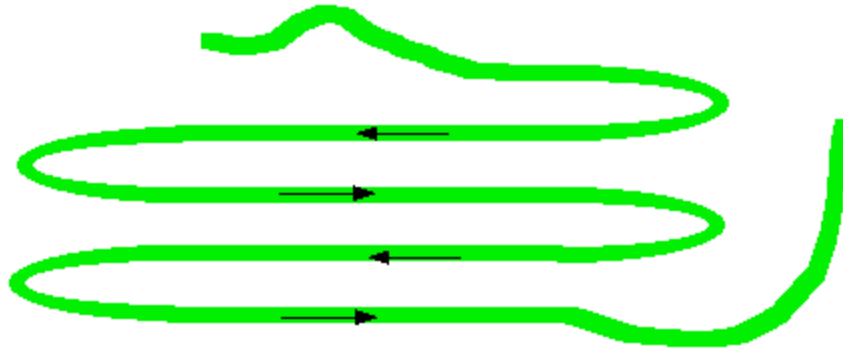
And finally, although you can't see it from this incomplete diagram, each complete turn of the spiral has 3.6 (approximately) amino acid residues in it.

If you had a whole number of amino acid residues per turn, each group would have an identical group underneath it on the turn below. Hydrogen bonding can't happen under those circumstances.

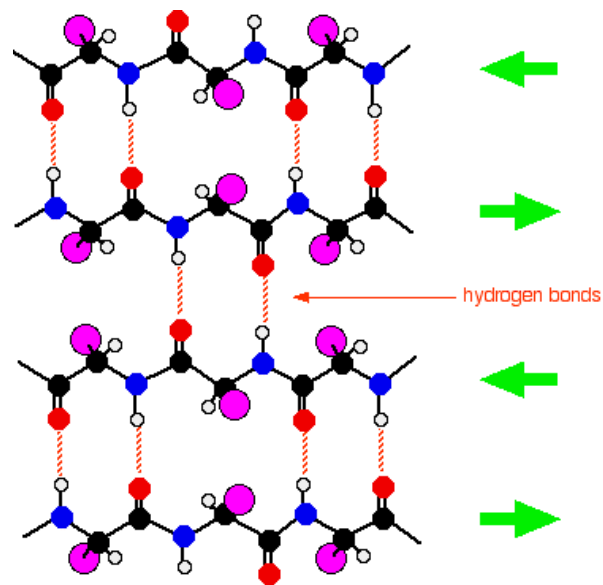
Each turn has 3 complete amino acid residues and two atoms from the next one. That means that each turn is offset from the ones above and below, such that the N-H and C=O groups are brought into line with each other.

Beta-pleated sheets

In a beta-pleated sheet, the chains are folded so that they lie alongside each other. The next diagram shows what is known as an "anti-parallel" sheet. All that means is that next-door chains are heading in opposite directions. Given the way this particular folding happens, that would seem to be inevitable.



It isn't, in fact, inevitable! It is possible to have some much more complicated folding so that next-door chains are actually heading in the same direction. The folded chains are again held together by hydrogen bonds involving exactly the same groups as in the alpha-helix.



The tertiary structure of proteins

The tertiary structure of a protein is a description of the way the whole chain (including the secondary structures) folds itself into its final 3-dimensional shape. This is often simplified into models like the following one for the enzyme dihydrofolate reductase. Enzymes are, of course, based on proteins.



The model shows the alpha-helices in the secondary structure as coils of "ribbon". The beta-pleated sheets are shown as flat bits of ribbon ending in an arrow head. The bits of the protein chain which are just random coils and loops are shown as bits of "string". The color coding in the model helps you to track your way around the structure - going through the spectrum from dark blue to end up at red. You will also notice that this particular model has two other molecules locked into it (shown as ordinary molecular models). These are the two molecules whose reaction this enzyme catalyses.

What holds a protein into its tertiary structure?

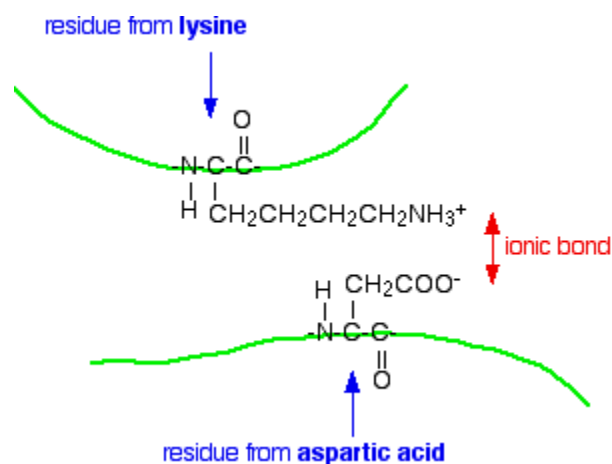
The tertiary structure of a protein is held together by interactions between the the side chains - the "R" groups. There are several ways this can happen.

Ionic interactions

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

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

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

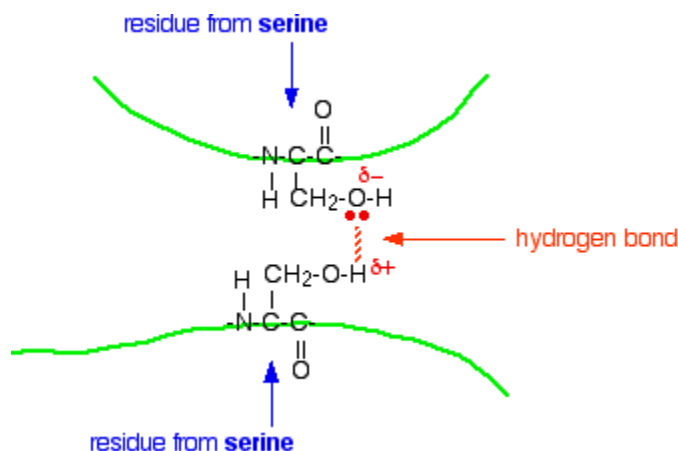


Hydrogen bonds

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

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

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

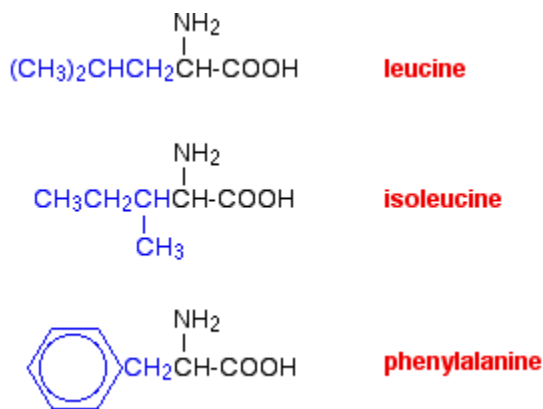


You could easily imagine similar hydrogen bonding involving -OH groups, or -COOH groups, or -CONH₂ groups, or -NH₂ groups in various combinations - although you would have to be careful to remember that a -COOH group and an -NH₂ group would form a zwitterion and produce stronger ionic bonding instead of hydrogen bonds.

van der Waals dispersion forces

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

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



Sulfur bridges

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

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

- [The Structure of Proteins](#) by [Jim Clark](#) is licensed [CC BY-NC 4.0](#).

2.2: Protein Folding

Introduction and Protein Structure

Proteins have several layers of structure each of which is important in the process of protein folding. The first most basic level of this structure is the sequence of amino acids themselves.¹ The sequencing is important because it will determine the types of interactions seen in the protein as it is folding. A novel sequence-based method based on the assumption that protein-protein interactions are more related to amino acids at the surface than those at the core.² This study shows that not only is the amino acids that are in a protein important but also the order in which they are sequenced. The interactions of the amino acids will determine what the secondary and tertiary structure of the protein will be.

The next layer in protein structure is the secondary structure. The secondary structure includes architectural structures that extend in one dimension.¹ Secondary structure includes α -Helices and β -sheets (Figure 2.2.1). The α -helices, the most common secondary structure in proteins, the peptide $-\text{CO}-\text{NH}$ -groups in the backbone form chains held together by NH OC hydrogen bonds.³ The α -helices form the backbone of proteins and help to aid in the folding process. The β -sheets form in two distinct ways. They are able to form in both parallel β -pleated sheets and anti parallel β -pleated sheets.¹ When the α -helix or β -sheet is formed, the excluded volumes generated by the backbone and side chains overlap, leading to an increase in the total volume available to the translational displacement of water molecules.⁴ This is important because it leads to a more thermodynamically stable conformation and leads to less strain on the protein as a whole and thus are aided by the conformation.

The tertiary structure is the next layer in protein structure. This takes the α -Helices and β -sheets and allows them to fold into a three dimensional structure.¹ Most proteins take on a globular structure once folded. The description of globular protein structures as an ensemble of contiguous closed loops or tightened end fragments reveals fold elements crucial for the formation of stable structures and for navigating the very process of protein folding.⁵ The globular proteins generally have a hydrophobic core surrounded by a hydrophilic outer layer. These interactions are important because they lead to the global structure and help create channels and binding sites for enzymes.

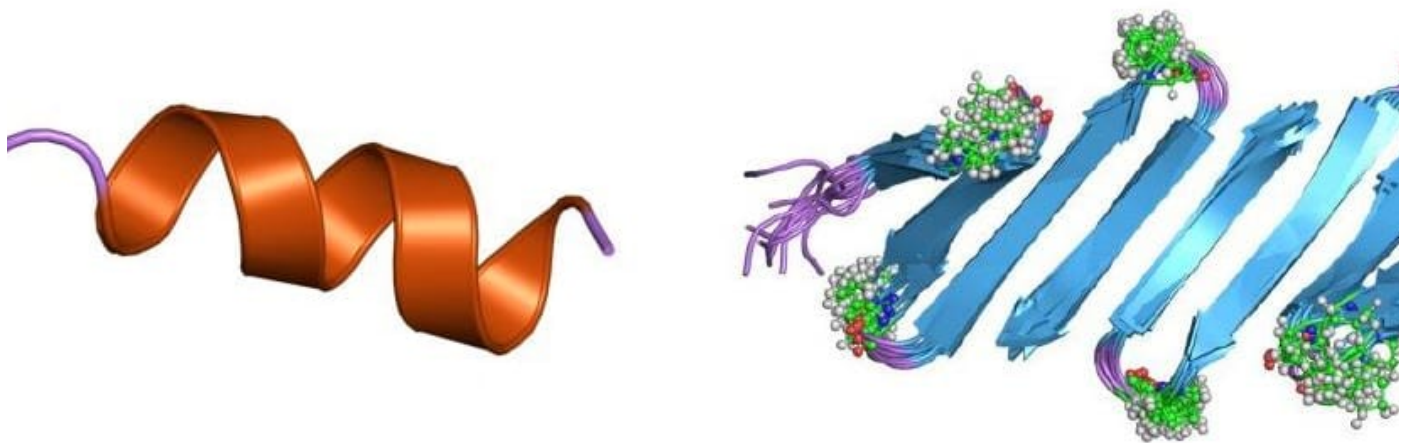


Figure 2.2.1: (left) typical example to an α -helix, from Wikimedia Commons. (right) Typical example of a β -sheet (Public Domain; Jawahar Swaminathan and MSD staff at the [European Bioinformatics Institute](#) via [Wikipedia](#))

The last layer of protein structure is the quaternary structure. The folding transition and the functional transitions between useful states are encoded in the linear sequence of amino acids, and a long- term goal of structural biology is to be able to predict both the structure and function of molecules from the information in the sequence.⁶ The Subunit organization is the last level of

structure in protein molecules.¹ The organization of the subunits is important because that determines the types of interactions that can form and dictates its use in the body.

Protein Folding

Proteins are folded and held together by several forms of molecular interactions. The molecular interactions include the thermodynamic stability of the complex, the hydrophobic interactions and the disulfide bonds formed in the proteins. The figure below (Figure 2.2.2) is an example of protein folding.

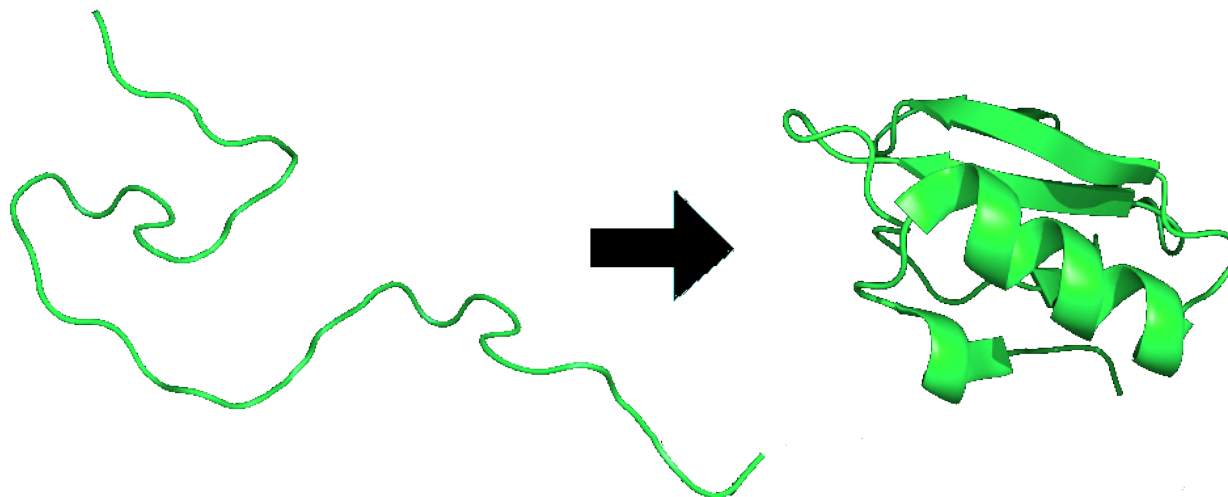


Figure 2.2.2: Protein Folding. (Public Domain; [DrKjaergaard](#) via [Wikipedia](#))

The biggest factor in a proteins ability to fold is the thermodynamics of the structure. The interaction scheme includes the short-range propensity to form extended conformations, residue-dependent long-range contact potentials, and orientation-dependent hydrogen bonds.⁷ The thermodynamics are a main stabilizing force within a protein because if it is not in the lowest energy conformation it will continue to move and adjust until it finds its most stable state. The use of energy diagrams and maps are key in finding out when the protein is in the most stable form possible.

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

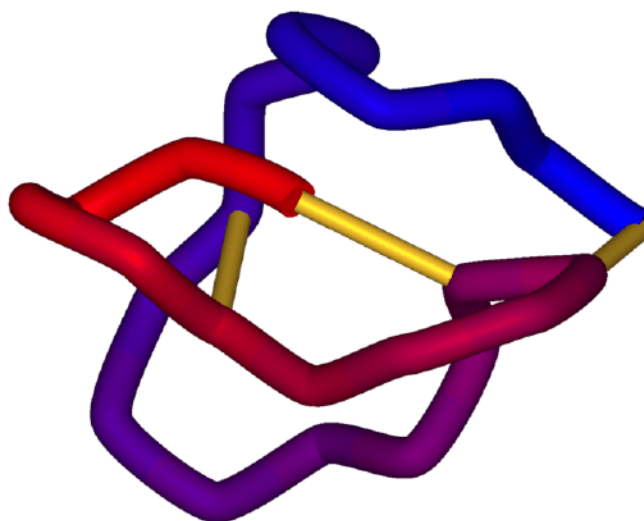


Figure 2.2.3: Disulfide Bonds. (Public Domain; [vasconcellos](#) via [Wikipedia](#))

Another type of interaction seen when the protein is folding is the disulfide linkages that form in the protein (Figure 2.2.3). The disulfide bond, a sulfur- sulfur chemical bond that results from an oxidative process that links nonadjacent (in most cases) cysteine's of a protein.⁹ These are a major way that proteins get into their folded form. The types of disulfide bonds are cysteine-cysteine linkage is a stable part of their final folded structure and those in which pairs of cysteines alternate between the reduced and oxidized states.⁹ The more common is the linkages that cause the protein to fold together and link back on itself compared to the cysteines that are changing oxidation states because the bonds between cysteines once created are fairly stable.

Misfunctions

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

Alzheimer's Disease

Alzheimer's Disease (AD) is a neurological degenerative disease that affects around 5 million Americans, including nearly half of those who are age 85 or older.¹⁰ The predominant risk factors of AD are age, family history, and heredity. Alzheimer's disease typically results in memory loss, confusion of time and place, misplacing places, and changes in mood and behavior.¹¹ AD results in dense plaques in the brain that are comprised of fibrillar β -amyloid proteins with a well-orders β -sheet secondary structure.¹² These plaques visually look like voids in the brain matter (see figure 5) and are directly connected to the deterioration of thought processes. It has been determined that AD is a protein misfolding disease, where the misfolded protein is directly related to the formation of these plaques in the brain.¹³

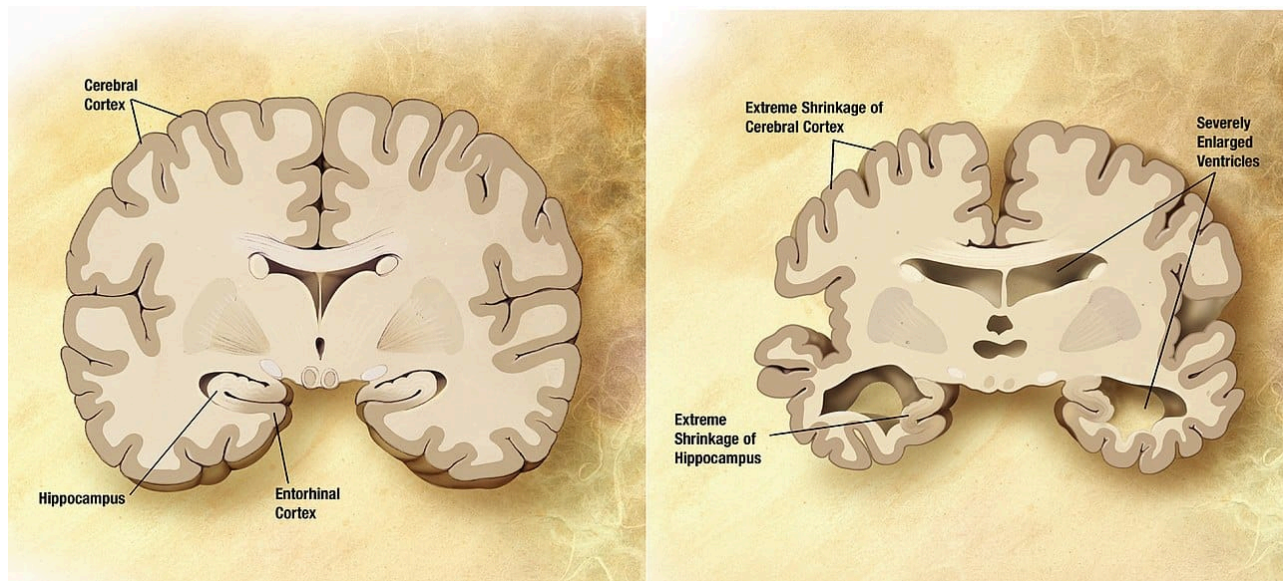


Figure 2.2.4: Comparison of healthy brain (left) with the brain with Alzheimer's (right). (Public Domain; NIH via Wikipedia)

It is yet to be fully understood what exactly causes this protein misfolding to begin, but several theories point to oxidative stress in the brain to be the initiating factor. This oxidation results in damage to the phospholipids in the brain, which has been found to result in a faster accumulation of amyloid β -proteins.¹⁴

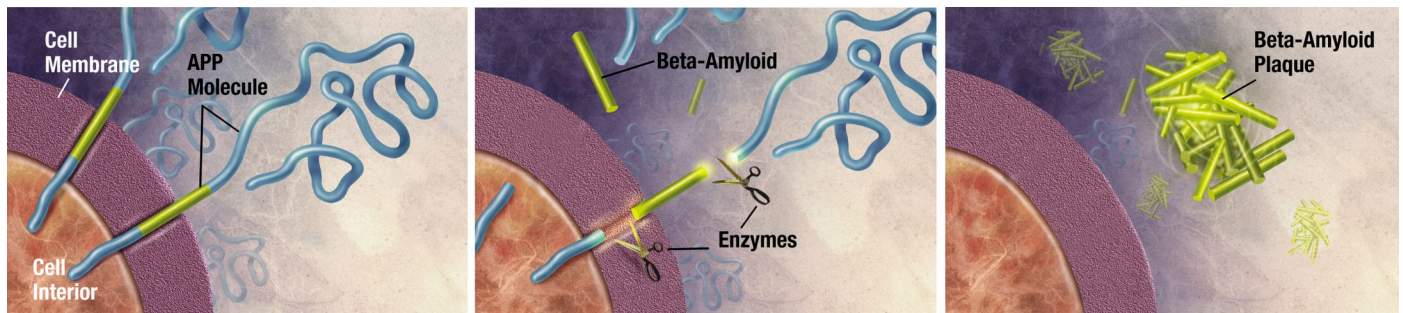


Figure 2.2.5: Beta-Amyloid Plaque Formation. (Public Domain; NIH via Wikipedia)

Cystic Fibrosis

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

Resources

1. Garrett, R.H., and Grisham, C.M. *Biochemistry fourth edition*; Brooks/Cole. Australia, 2010. p. 93-95, 135, 143, 160.
2. Chang, D.T., Syu Y., Lin, P., Predicting the protein-protein interactions using primary structures with predicted protein surface. *BCM Bioinformatics*. Jan 2010
3. Bodis, P., Schwartz, E., Koepf, M. Cornelissen, J. Rowan A.E., Roeland J. Nolte, M. and Woutersen S. Vibrational self-trapping in beta-sheet structures observed with femtosecond nonlinear infrared spectroscopy. *The Journal Of Chemical Physics* **131**, 2009
4. Yasuda, S., Yoshidome, T. Oshima, H. Kodama R., Harano, Y. and Kinoshita, M. Effects of side-chain packing on the formation

- of secondary structures in protein folding. *The Journal of Chemical Physics* **132**, 2010
5. Papandreoul, N., Berezovsky I.N., Lopes, A., Eliopoulos, E., and Chomilier, J. Universal positions in globular proteins from observation to simulation. *Eur. J. Biochem.* 271, 4762–4768 (2004)
 6. Hausratha, A.C. A Kinetic theory of tertiary contact formation coupled to the helix-coil transition in polypeptides. *The Journal of Chemical Physics.* 2006
 7. Cieplaka M. and Niewieczerza S. Hydrodynamic interactions in protein folding. *Journal of Chemical Physics.* **130** 2009
 8. Gront, D. Kolinskia, A. and Skolnick, J. A new combination of replica exchange Monte Carlo and histogram analysis for protein folding and thermodynamics. *Journal of Chemical Physics.* Vol. 115 2001
 9. Kadokura, H., Katzen F. and Beckwith, J. Protein disulfide bond formation in prokaryotes. *Annual Review Biochem* 2003.
 10. Alzheimer's Disease. Centers for Disease Control and Prevention. 2010, <http://www.cdc.gov/aging/aginginfo/alzheimers.htm>.
 11. Alzheimer's Association. 2011, www.alz.org.
 12. Glenner, G. G.; Wong, C. W. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.* **1984**, 120, 885–890.
 13. Hashimoto, M.; Rockenstein, E.; Crews, L.; Mashliah, E. Role of Protein Aggregation in Mitochondrial Dysfunction and Neurodegeneration in Alzheimer's and Parkinson's Diseases. *Neuromolecular Med.* **2003**, 4, 21–36
 14. The Cystic Fibrosis Foundation. 2011, <http://www.cff.org/AboutCF>.
 15. Cheung, J. C. and Deber, C. M. Misfolding of the Cystic Fibrosis Transmembrane Conductance Regulator and Disease. *Biochemistry* **2008**, 47, 1465–1473.
 16. Koppaka, V.; Axelsen, P. Accelerated Accumulation of Amyloid β Proteins on Oxidatively Damaged Lipid Membranes. *Biochemistry* **2000**, 39, 10011–10016.
 17. Riordan, J.M. et al. Identification of the cysticfibrosigene: cloning and characterization of complementary DNA. *Science* **1989**, 245, 1066–1073.

- [Protein Folding](#) is licensed [CC BY-NC-SA 4.0](#).

2.3: Denaturation of proteins

Learning Objectives

- Describe the four levels of protein structure.
- Identify the types of attractive interactions that hold proteins in their most stable three-dimensional structure.
- Explain what happens when proteins are denatured.
- Identify how a protein can be denatured.

Each of the thousands of naturally occurring proteins has its own characteristic amino acid composition and sequence that result in a unique three-dimensional shape. Since the 1950s, scientists have determined the amino acid sequences and three-dimensional conformation of numerous proteins and thus obtained important clues on how each protein performs its specific function in the body.

Proteins are compounds of high molar mass consisting largely or entirely of chains of amino acids. Because of their great complexity, protein molecules cannot be classified on the basis of specific structural similarities, as carbohydrates and lipids are categorized. The two major structural classifications of proteins are based on far more general qualities: whether the protein is (1) fiberlike and insoluble or (2) globular and soluble. Some proteins, such as those that compose hair, skin, muscles, and connective tissue, are fiberlike. These fibrous proteins are insoluble in water and usually serve structural, connective, and protective functions. Examples of fibrous proteins are keratins, collagens, myosins, and elastins. Hair and the outer layer of skin are composed of keratin. Connective tissues contain collagen. Myosins are muscle proteins and are capable of contraction and extension. Elastins are found in ligaments and the elastic tissue of artery walls.

Globular proteins, the other major class, are soluble in aqueous media. In these proteins, the chains are folded so that the molecule as a whole is roughly spherical. Familiar examples include egg albumin from egg whites and serum albumin in blood. Serum albumin plays a major role in transporting fatty acids and maintaining a proper balance of osmotic pressures in the body. Hemoglobin and myoglobin, which are important for binding oxygen, are also globular proteins.

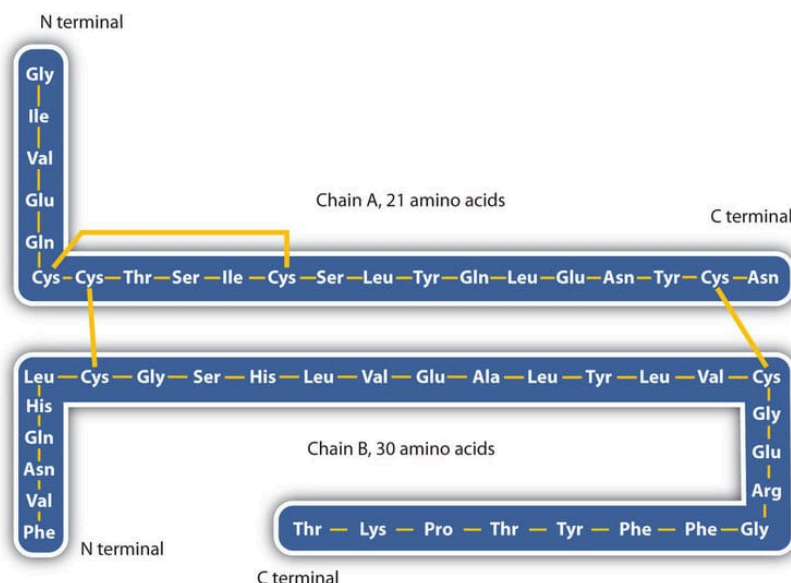


Figure 2.3.1: Primary Structure of Human Insulin. Human insulin, whose amino acid sequence is shown here, is a hormone that is required for the proper metabolism of glucose.

Levels of Protein Structure

The structure of proteins is generally described as having four organizational levels. The first of these is the primary structure, which is the number and sequence of amino acids in a protein's polypeptide chain or chains, beginning with the free amino group

and maintained by the peptide bonds connecting each amino acid to the next. The primary structure of insulin, composed of 51 amino acids, is shown in Figure 2.3.1.

A protein molecule is not a random tangle of polypeptide chains. Instead, the chains are arranged in unique but specific conformations. The term secondary structure refers to the fixed arrangement of the polypeptide backbone. On the basis of X ray studies, Linus Pauling and Robert Corey postulated that certain proteins or portions of proteins twist into a spiral or a helix. This helix is stabilized by *intrachain* hydrogen bonding between the carbonyl oxygen atom of one amino acid and the amide hydrogen atom four amino acids up the chain (located on the next turn of the helix) and is known as a right-handed α -helix. X ray data indicate that this helix makes one turn for every 3.6 amino acids, and the side chains of these amino acids project outward from the coiled backbone (Figure 2.3.2). The α -keratins, found in hair and wool, are exclusively α -helical in conformation. Some proteins, such as gamma globulin, chymotrypsin, and cytochrome c, have little or no helical structure. Others, such as hemoglobin and myoglobin, are helical in certain regions but not in others.

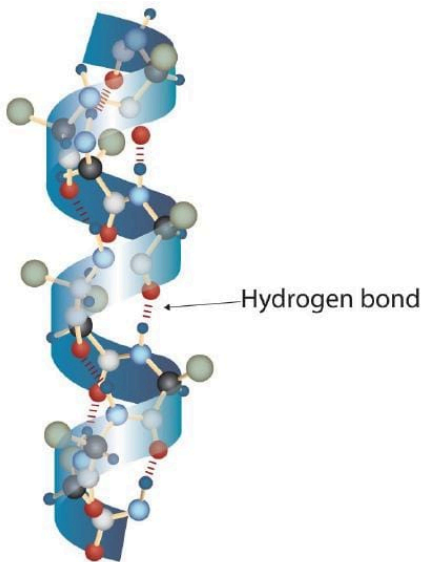


Figure 2.3.2 A Ball-and-Stick Model of an α -Helix. This ball-and-stick model shows the intrachain hydrogen bonding between carbonyl oxygen atoms and amide hydrogen atoms. Each turn of the helix spans 3.6 amino acids. Note that the side chains (represented as green spheres) point out from the helix.

Another common type of secondary structure, called the β -pleated sheet conformation, is a sheetlike arrangement in which two or more extended polypeptide chains (or separate regions on the same chain) are aligned side by side. The aligned segments can run either parallel or antiparallel—that is, the N-terminals can face in the same direction on adjacent chains or in different directions—and are connected by *interchain* hydrogen bonding (Figure 2.3.3). The β -pleated sheet is particularly important in structural proteins, such as silk fibroin. It is also seen in portions of many enzymes, such as carboxypeptidase A and lysozyme.

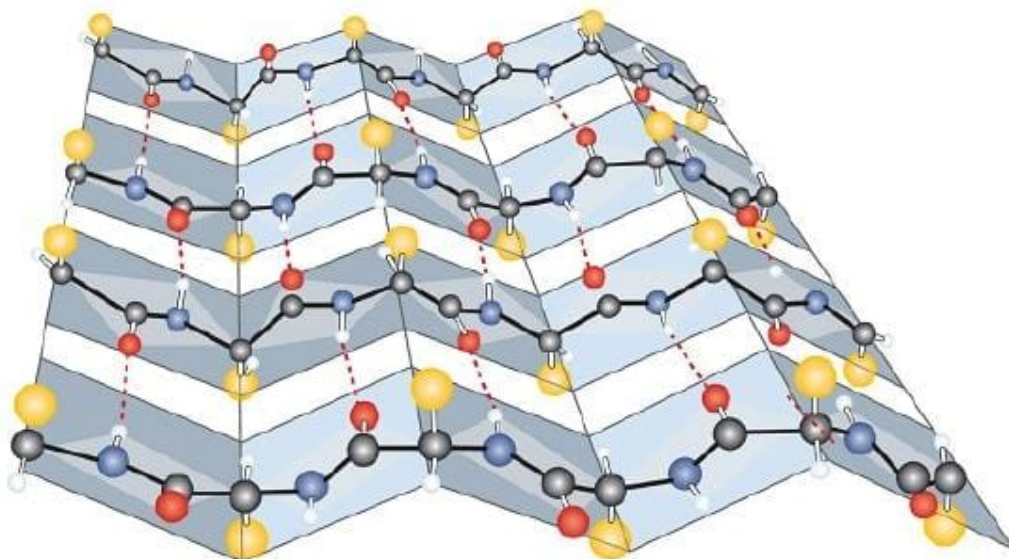


Figure 2.3.3: A Ball-and-Stick Model of the β -Pleated Sheet Structure in Proteins. The side chains extend above or below the sheet and alternate along the chain. The protein chains are held together by interchain hydrogen bonding.

Tertiary structure refers to the unique three-dimensional shape of the protein as a whole, which results from the folding and bending of the protein backbone. The tertiary structure is intimately tied to the proper biochemical functioning of the protein. Figure 2.3.4 shows a depiction of the three-dimensional structure of insulin.

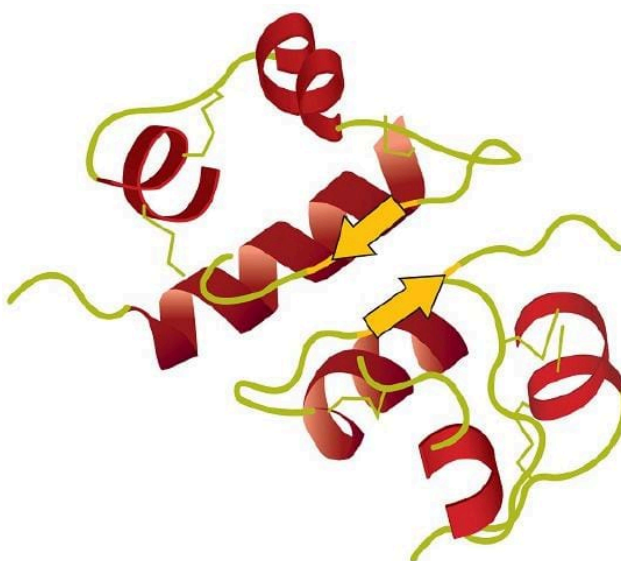
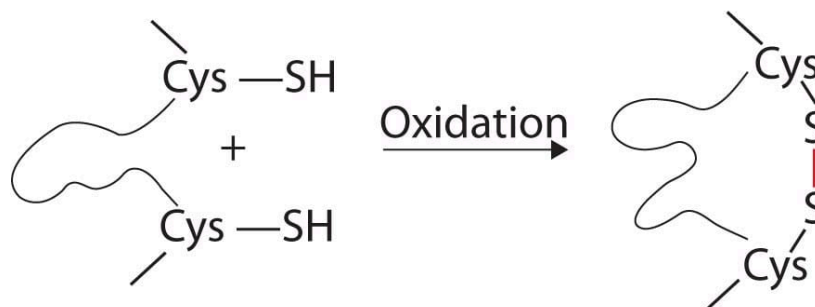


Figure 2.3.4: A Ribbon Model of the Three-Dimensional Structure of Insulin. The spiral regions represent sections of the polypeptide chain that have an α -helical structure, while the broad arrows represent β -pleated sheet structures.

Four major types of attractive interactions determine the shape and stability of the tertiary structure of proteins. You studied several of them [previously](#).

1. **Ionic bonding.** Ionic bonds result from electrostatic attractions between positively and negatively charged side chains of amino acids. For example, the mutual attraction between an aspartic acid carboxylate ion and a lysine ammonium ion helps to maintain a particular folded area of a protein (part (a) of Figure 2.3.5).
2. **Hydrogen bonding.** Hydrogen bonding forms between a highly electronegative oxygen atom or a nitrogen atom and a hydrogen atom attached to another oxygen atom or a nitrogen atom, such as those found in polar amino acid side chains. Hydrogen bonding (as well as ionic attractions) is extremely important in both the intra- and intermolecular interactions of proteins (part (b) of Figure 2.3.5).
3. **Disulfide linkages.** Two cysteine amino acid units may be brought close together as the protein molecule folds. Subsequent

oxidation and linkage of the sulfur atoms in the highly reactive sulfhydryl (SH) groups leads to the formation of cystine (part (c) of Figure 2.3.5). Intrachain disulfide linkages are found in many proteins, including insulin (yellow bars in Figure 2.3.1) and have a strong stabilizing effect on the tertiary structure.



4. Dispersion forces. Dispersion forces arise when a normally nonpolar atom becomes momentarily polar due to an uneven distribution of electrons, leading to an instantaneous dipole that induces a shift of electrons in a neighboring nonpolar atom. Dispersion forces are weak but can be important when other types of interactions are either missing or minimal (part (d) of Figure 2.3.5). This is the case with fibroin, the major protein in silk, in which a high proportion of amino acids in the protein have nonpolar side chains. The term *hydrophobic interaction* is often misused as a synonym for dispersion forces. Hydrophobic interactions arise because water molecules engage in hydrogen bonding with other water molecules (or groups in proteins capable of hydrogen bonding). Because nonpolar groups cannot engage in hydrogen bonding, the protein folds in such a way that these groups are buried in the interior part of the protein structure, minimizing their contact with water.

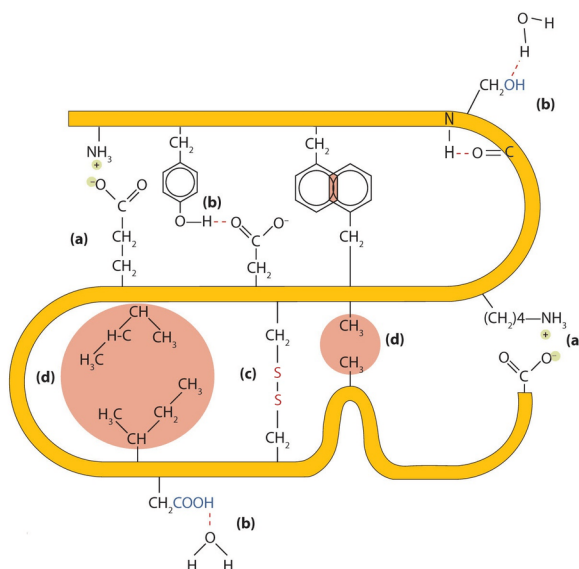


Figure 2.3.5: Tertiary Protein Structure Interactions. Four interactions stabilize the tertiary structure of a protein: (a) ionic bonding, (b) hydrogen bonding, (c) disulfide linkages, and (d) dispersion forces.

When a protein contains more than one polypeptide chain, each chain is called a *subunit*. The arrangement of multiple subunits represents a fourth level of structure, the quaternary structure of a protein. Hemoglobin, with four polypeptide chains or subunits, is the most frequently cited example of a protein having quaternary structure (Figure 2.3.6). The quaternary structure of a protein is produced and stabilized by the same kinds of interactions that produce and maintain the tertiary structure. A schematic representation of the four levels of protein structure is in Figure 2.3.7.

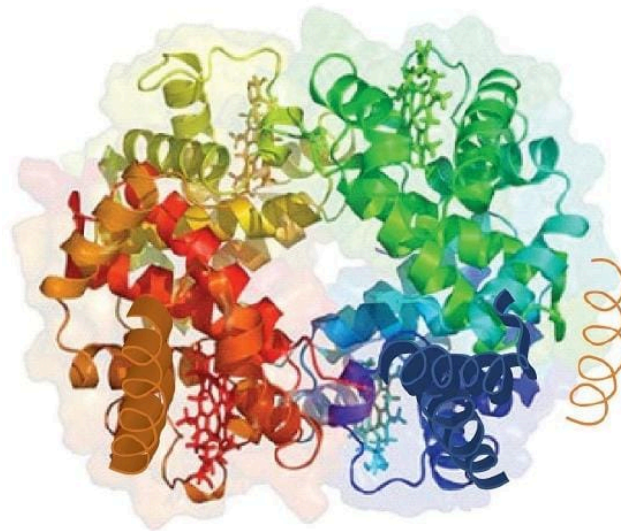


Figure 2.3.6 The Quaternary Structure of Hemoglobin. Hemoglobin is a protein that transports oxygen throughout the body.

Source: Image from the RCSB PDB ([www.pdb.org/opens in new window](http://www.pdb.org/opens%20in%20new%20window)) of PDB ID 1I3D (R.D. Kidd, H.M. Baker, A.J. Mathews, T. Brittain, E.N. Baker (2001) Oligomerization and ligand binding in a homotetrameric hemoglobin: two high-resolution crystal structures of hemoglobin Bart's (gamma(4)), a marker for alpha-thalassemia. Protein Sci. 1739-1749).

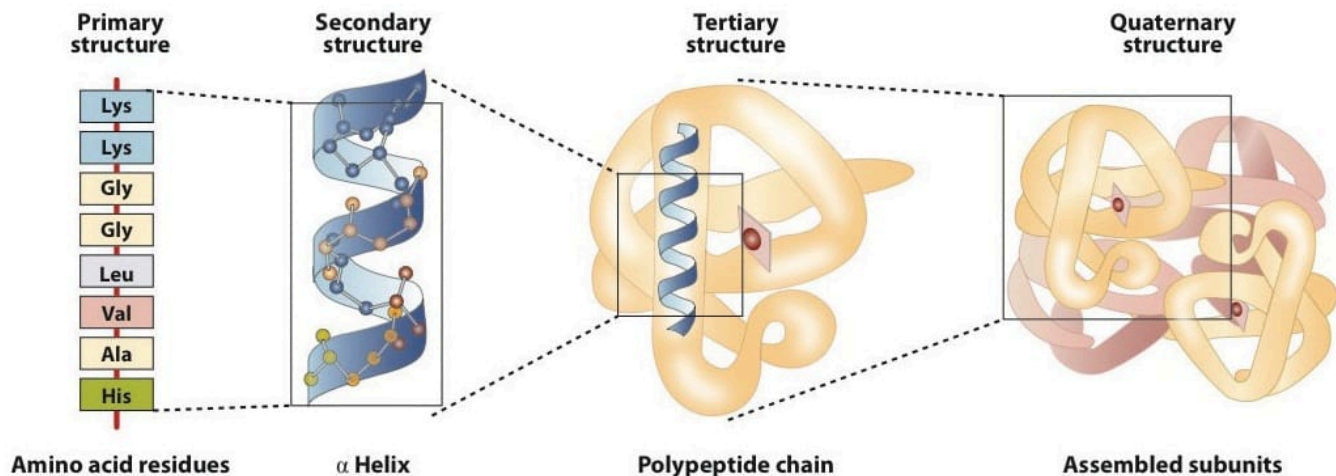


Figure 2.3.7: Levels of Structure in Proteins

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

Denaturation of Proteins

The highly organized structures of proteins are truly masterworks of chemical architecture. But highly organized structures tend to have a certain delicacy, and this is true of proteins. Denaturation is the term used for any change in the three-dimensional structure of a protein that renders it incapable of performing its assigned function. A denatured protein cannot do its job. (Sometimes denaturation is equated with the precipitation or coagulation of a protein; our definition is a bit broader.) A wide variety of reagents and conditions, such as heat, organic compounds, pH changes, and heavy metal ions can cause protein denaturation (Figure 2.3.1).

Figure 2.3.1: Protein Denaturation Methods

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

Anyone who has fried an egg has observed denaturation. The clear egg white turns opaque as the albumin denatures and coagulates. No one has yet reversed that process. However, given the proper circumstances and enough time, a protein that has unfolded under sufficiently gentle conditions can refold and may again exhibit biological activity (Figure 2.3.8). Such evidence suggests that, at least for these proteins, the primary structure determines the secondary and tertiary structure. A given sequence of amino acids seems to adopt its particular three-dimensional arrangement naturally if conditions are right.

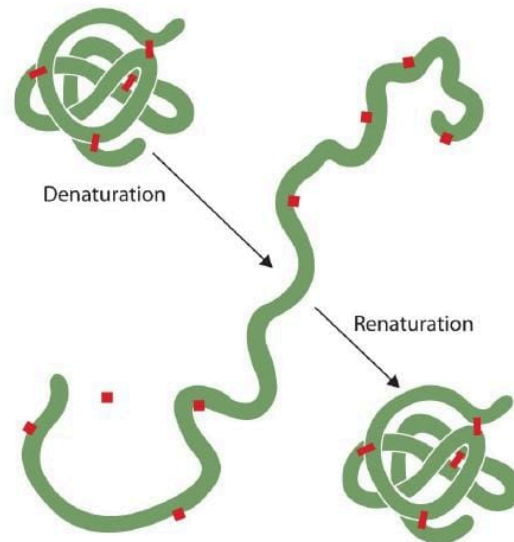


Figure 2.3.8: Denaturation and Renaturation of a Protein. The denaturation (unfolding) and renaturation (refolding) of a protein is depicted. The red boxes represent stabilizing interactions, such as disulfide linkages, hydrogen bonding, and/or ionic bonds.

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

Summary

Proteins can be divided into two categories: fibrous, which tend to be insoluble in water, and globular, which are more soluble in water. A protein may have up to four levels of structure. The primary structure consists of the specific amino acid sequence. The resulting peptide chain can form an α -helix or β -pleated sheet (or local structures not as easily categorized), which is known as secondary structure. These segments of secondary structure are incorporated into the tertiary structure of the folded polypeptide chain. The quaternary structure describes the arrangements of subunits in a protein that contains more than one subunit. Four major types of attractive interactions determine the shape and stability of the folded protein: ionic bonding, hydrogen bonding, disulfide linkages, and dispersion forces. A wide variety of reagents and conditions can cause a protein to unfold or denature.

This page titled [2.3: Denaturation of proteins](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [18.4: Proteins](#) by Anonymous is licensed [CC BY-NC-SA 3.0](#). Original source: <https://2012books.lardbucket.org/books/introduction-to-chemistry-general-organic-and-biological>.

2.4: Amino Acids and Proteins (Exercises)

These are homework exercises to accompany [Chapter 13](#) of the University of Kentucky's LibreText for [CHE 103 - Chemistry for Allied Health](#).

Questions

13.1: Amino Acids

[\(click here for solutions\)](#)

Q13.1.1

Read the material at <http://hyperphysics.phy-astr.gsu.edu/hbase/organic/essam.html> and answer the following questions:

- What are essential amino acids?
- What are nonessential amino acids?
- What happens if you are deficient in an amino acid?

Q13.1.2

Draw the functional groups present in all amino acids.

Q13.1.3

Complete the following for threonine, lysine, and tyrosine.

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

13.2: Peptides

[\(click here for solutions\)](#)

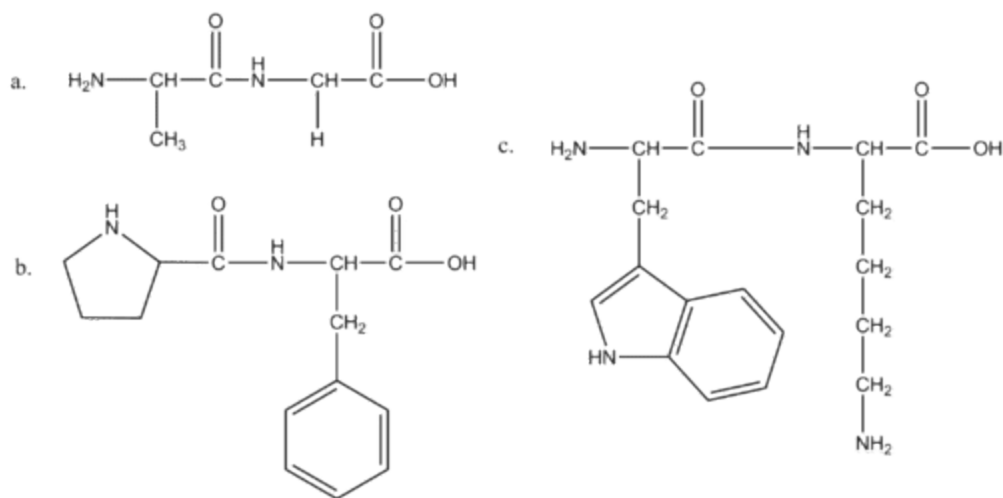
Q13.2.1

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

- tyrosine and lysine
- threonine and glutamine
- alanine and histidine

Q13.2.2

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



Q13.2.3

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

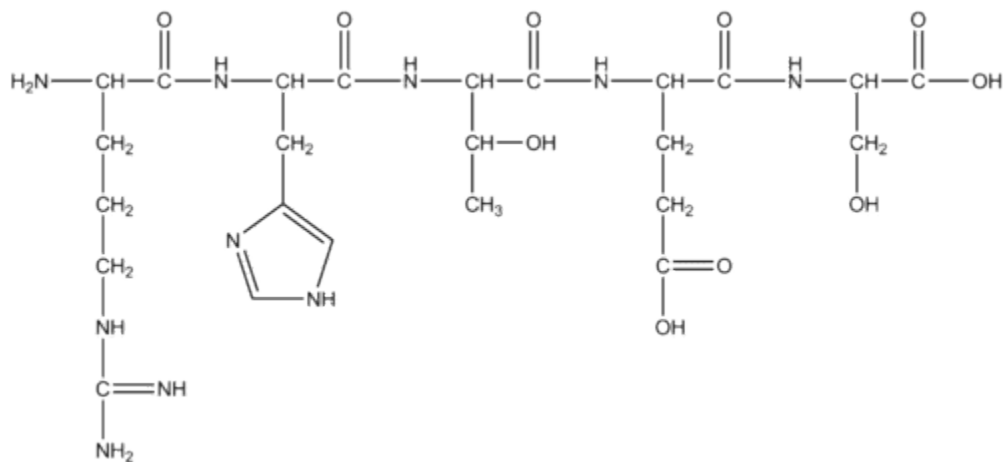
Q13.2.4

Draw the following polypeptides.

- Ser-Tyr-Gln
- Lys-Met-Gly

Q13.2.5

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



13.3: Protein Structure

[\(click here for solutions\)](#)

Q13.3.1

Describe the four levels of protein structure.

Q13.3.2

What levels of structure involve hydrogen bonding?

Q13.3.3

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

Q13.3.4

What types of interactions hold the secondary structure together?

Q13.3.5

What types of interactions hold the tertiary structure together?

Q13.3.6

What levels of structure are affected by denaturation?

Q13.3.7

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

Answers

13.1: Amino Acids

Q13.1.1

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

Q13.1.2

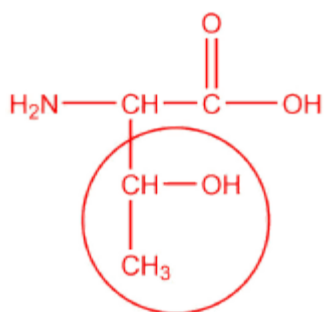
amine and carboxylic acid

Q13.1.3

Complete the following for threonine, lysine, and tyrosine.

threonine

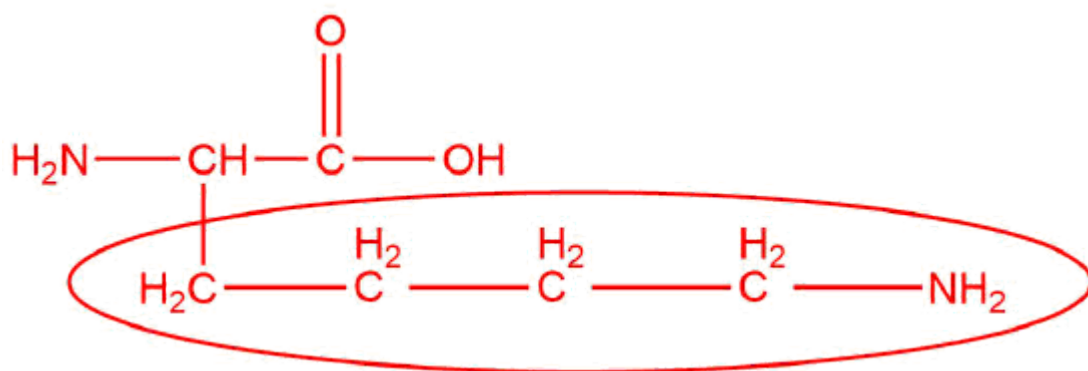
a.



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

lysine

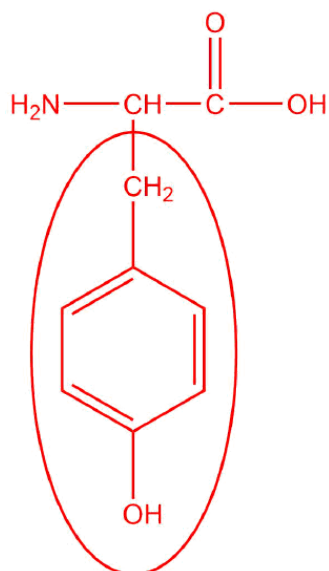
a.



- b.
- c. basic
- d. 9.47
- e. < 9.47
- f. > 9.47

tyrosine

a.



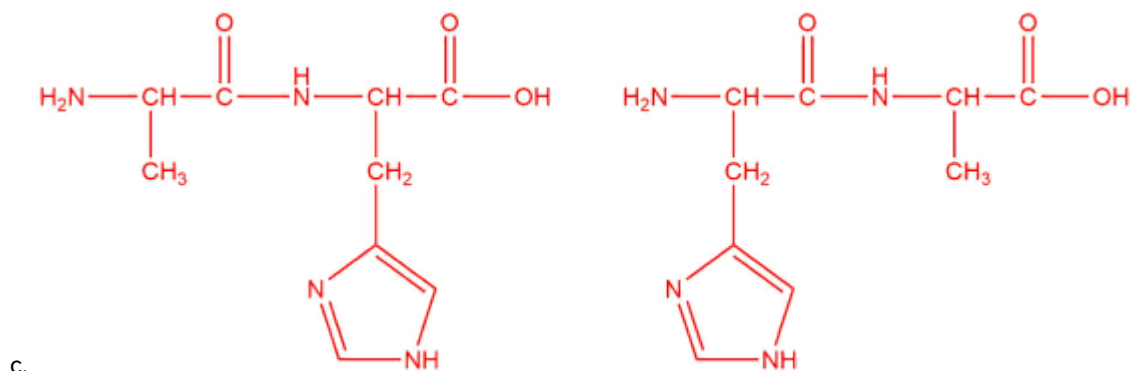
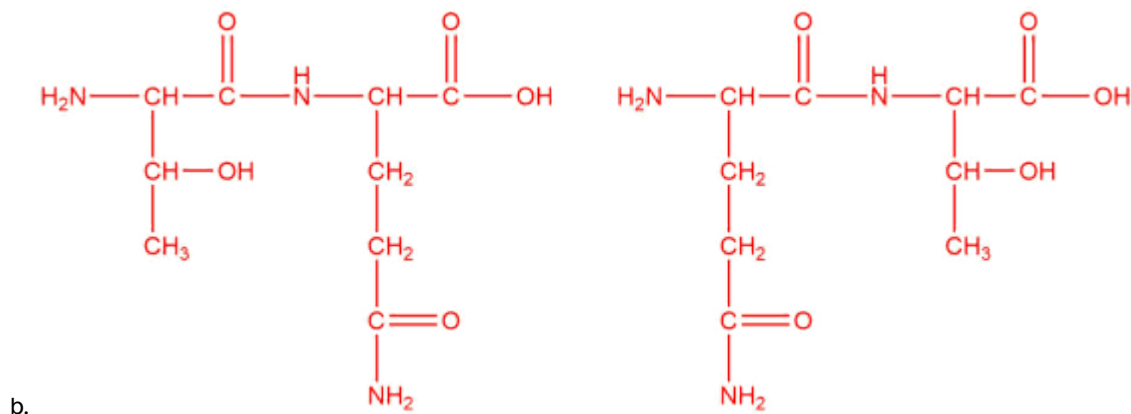
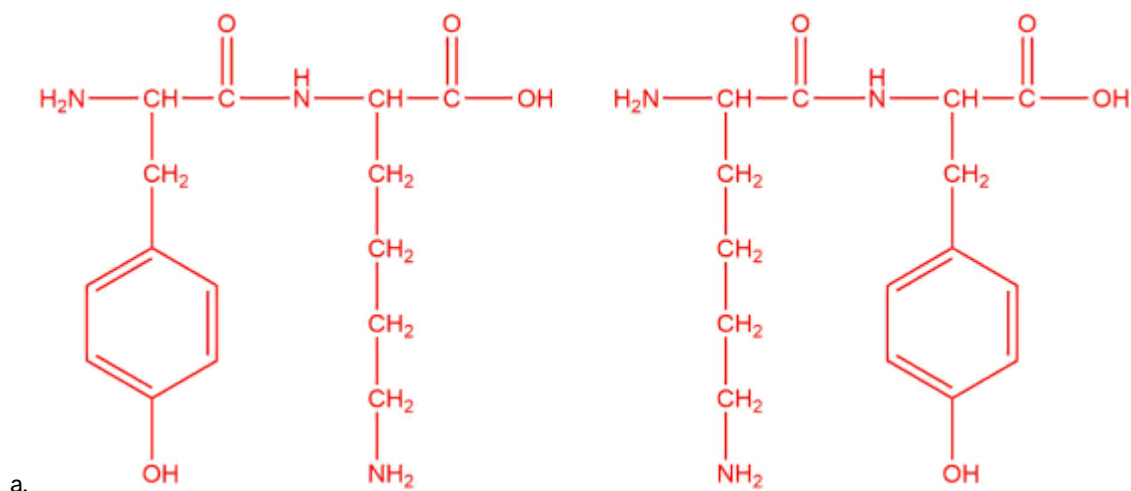
b.

- c. polar
- d. 5.63
- e. < 5.63
- f. > 5.63

13.2: Peptides

Q13.2.1

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



Q13.2.2

a.	$\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\ \\ \text{CH}_3 \end{array}$ alanine	$\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\ \\ \text{H} \end{array}$ glycine
b.	$\begin{array}{c} \text{O} \\ \parallel \\ \text{C}-\text{OH} \\ \\ \text{HN} \end{array}$ proline	$\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\ \\ \text{CH}_2 \\ \\ \text{C}_6\text{H}_5 \end{array}$ phenylalanine
c.	$\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\ \\ \text{CH}_2 \\ \\ \text{C}_8\text{H}_6\text{N} \end{array}$ tryptophan	$\begin{array}{c} \text{O} \\ \parallel \\ \text{H}_2\text{N}-\text{CH}-\text{C}-\text{OH} \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{CH}_2 \\ \\ \text{NH}_2 \end{array}$ lysine

Q13.2.3

Thr-Ala-Phe

Thr-Phe-Ala

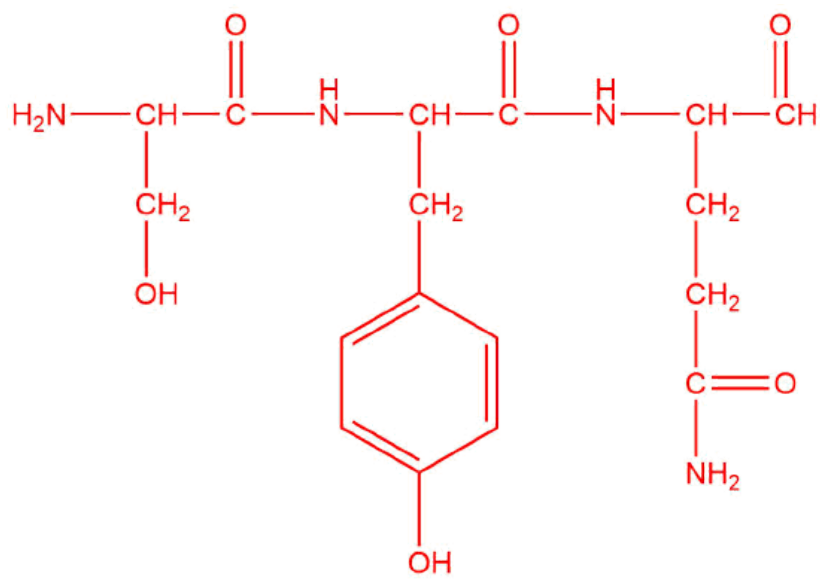
Ala-Thr-Phe

Ala-Phe-Thr

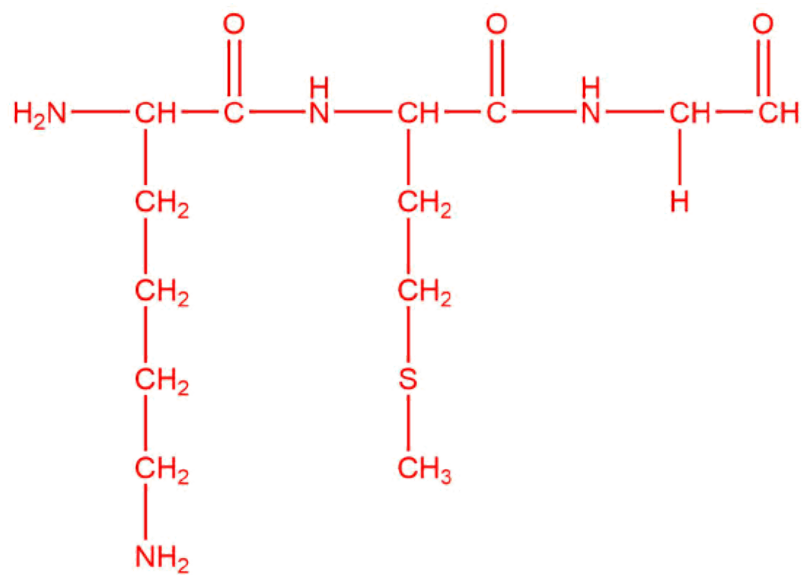
Phe-Ala-Thr

Phe-Thr-Ala

Q13.2.4

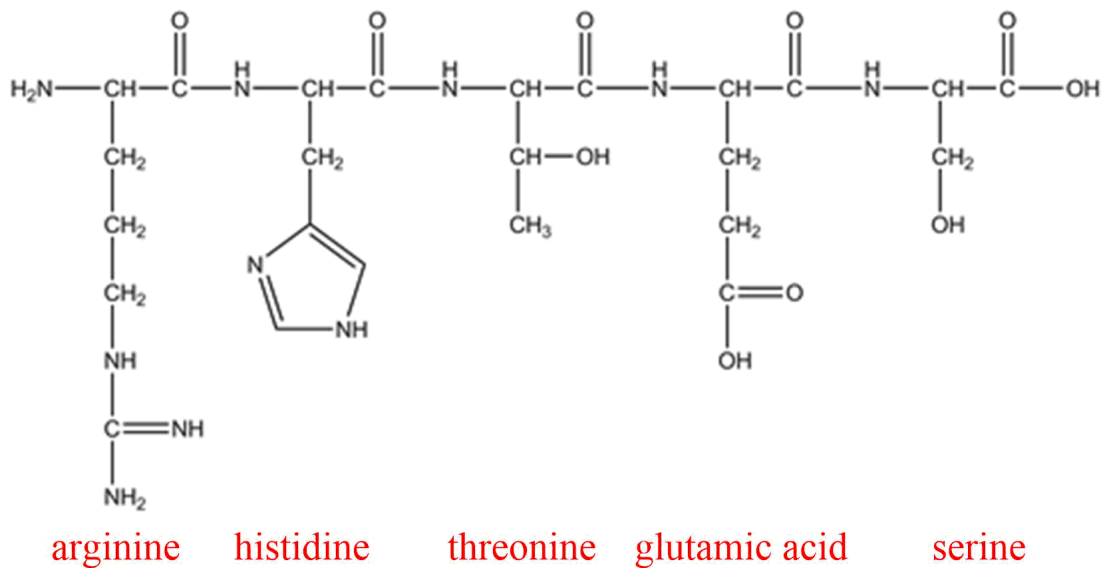


a.



b.

Q13.2.5



Arg-His-Thr-Glu-Ser

13.3: Protein Structure

Q13.3.1

Primary - sequence of amino acids

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

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

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

Q13.3.2

secondary, tertiary, and quaternary structures

Q13.3.3

tertiary structures

Q13.3.4

hydrogen bonds

Q13.3.5

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

Q13.3.6

secondary, tertiary, and quaternary

Q13.3.7

No, a quaternary structure must have multiple subunits.

This page titled [2.4: Amino Acids and Proteins \(Exercises\)](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Allison Soult](#).

3: METHODS OF PROTEIN PURIFICATION AND CHARACTERIZATION

3.1: Protein Purification

3.2: Cell Disruption

3.3: Fractionation

3.4: Chromatography

3.4.1: Affinity Chromatography

3.4.2: Gel Exclusion Chromatography

3.4.3: Ion Exchange Chromatography

3.5: Electrophoresis

3.5.1: Blotting and Hybridization

3.5.2: DNA Analysis - Gel Electrophoresis

3.5.3: Exercises

3.1: Protein Purification

Assays, Specific Activity, Initial Fractionation

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

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

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

Western Blotting

Antibodies can be used in a method called **Western blotting**, which is useful for determining levels of protein expression and for assaying proteins during purification. This method usually involves the following steps:

1. A protein sample is subjected to polyacrylamide gel electrophoresis.
2. After this the gel is placed over a sheet of nitrocellulose and the protein in the gel is electrophoretically transferred to the nitrocellulose.
3. The nitrocellulose is then soaked in gelatin to "block" its ability to non-specifically bind proteins.
4. The nitrocellulose is then incubated with the specific antibody for the protein of interest.
5. The nitrocellulose is then incubated with a second antibody which is specific for the first antibody. For example, if the first antibody was raised in rabbits, the second antibody might be termed "goat anti-rabbit immunoglobulin". What this means is that rabbit immunoglobulins were used to elicit an antibody response in goats. The goat antibodies (polyclonal) will include those which recognize the conserved region in the rabbit antibodies. Since the Fc region is conserved, it will bind to any and all rabbit antibodies, including those on the nitrocellulose paper.
6. The second antibody will typically have a covalently attached enzyme which, when provided with a chromogenic substrate, will cause a color reaction.
7. Thus the molecular weight and amount of the desired protein can be characterized from a complex mixture (e.g. crude cell extract) of other proteins.

In a variation of the above, the protein sample may be blotted directly on a nitrocellulose paper (called a **dot blot**) without first running a gel. This may be desirable if, for example, the antibody is monoclonal and recognizes an epitope which is dependent upon native structure (which would be destroyed upon running an SDS PAGE).

In addition to their varied uses, antibodies can also be used to purify proteins.

- If relatively large amounts of an antibody can be obtained, they can be covalently attached to a chromatography resin (e.g. sephadex beads).
- If a crude cell extract is run over such a column, only the protein of interest should bind, and everything else will flow through.
- The bound protein can then be eluted. This is typically achieved by moderately low pH conditions (using acetic acid). As long as the protein of interest is not irreversibly denatured by such conditions, the method will work quite well.
- One potential pitfall involves that of monoclonal antibodies being utilized to purify mutant proteins. The regions of the protein

comprising the epitope cannot be modified without destroying the ability of the antibody to bind. Thus, the use of monoclonal antibodies in a purification scheme may preclude its use in purifying certain mutants.

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

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

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

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

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

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

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

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

Initial steps in purification

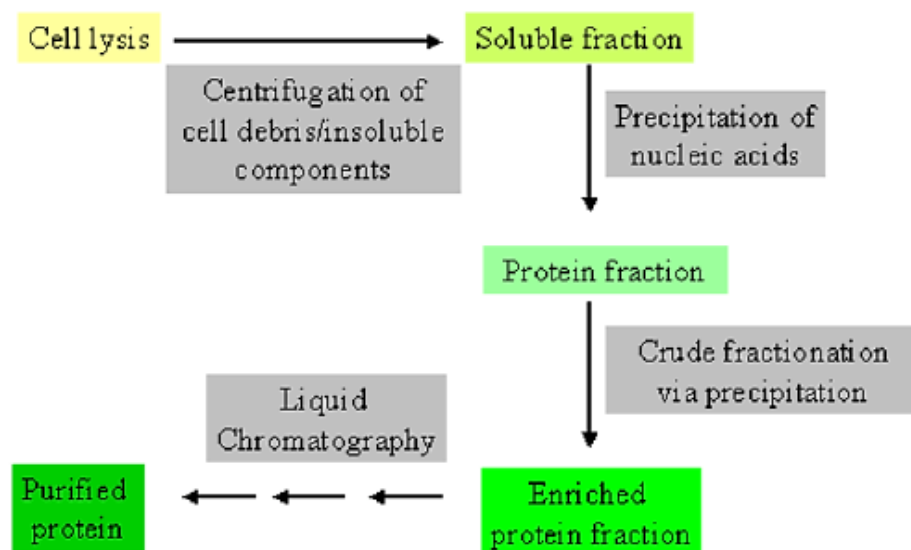


Figure 4.1.1: Purification steps

- It is extremely helpful to have some information not only on the general physical and chemical characteristics of the protein you are trying to purify, but also on the *contaminating* components.
- For example, many *E. coli* proteins are generally *low molecular weight* (<50,000 Da) and somewhat *acidic* in isoelectric point

Usually the initial steps in purification make use of general physical and/or chemical differences between soluble proteins and

other cell components.

- For example, soluble proteins can be separated from general cellular debris, and intact cells, by *centrifugation*.
- Thus, cells are physically disrupted (via homogenization or a cell press) to allow release of cell contents. This is then followed by centrifugation to separate generally soluble components from those which are insoluble.
- It is at this point that data collection begins in order to monitor the purification.

Nucleic acids can sometimes be readily removed from the sample by the addition of large cationic compounds such as *polyethylene imine*, or *streptomycin sulfate*.

- The nucleic acids bind to these compounds via electrostatic interactions and the complex precipitates and can be removed via centrifugation.
- The same general result can be obtained by mixing in ion exchange resins which are *anion exchangers* (i.e. the resins contain cationic groups) and then filtering or centrifuging to remove. As with either method, it should be confirmed that the desired protein is *not* bound as well.

Crude fractionations of proteins can be achieved by adding various quantities of precipitants such as *ammonium sulfate*, or *polyethylene glycol* (PEG).

- For this type of purification step an initial experiment is performed to monitor the fraction of overall protein, as well as desired protein, remaining in solution (and pellet) as a function of precipitant concentration.

Ammonium Sulfate (% saturated)	0	10	20	30	40	50	60	70	80	90
Sample A ₂₈₀	1000	900	600	200	100	75	50	40	25	20
Activity assay(units)	200	200	200	190	170	100	30	5	0	0

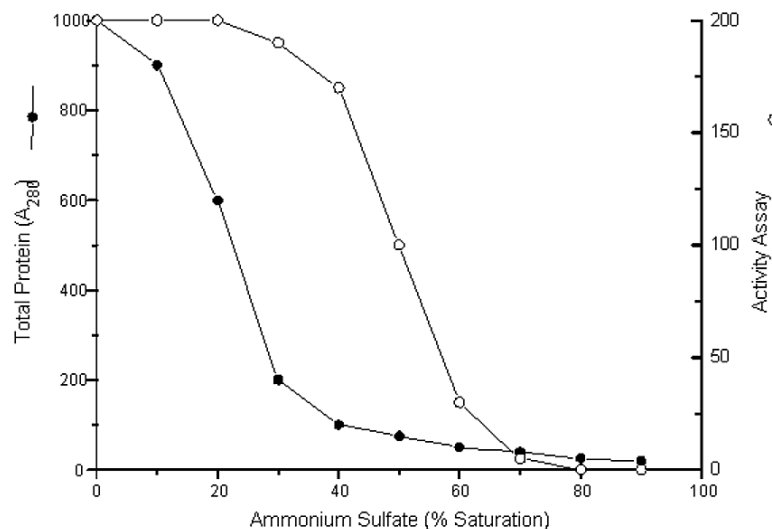


Figure 4.1.2: Protein activity as a function of precipitant concentration

- In this particular example we are in luck: at around 30% ammonium sulfate we can precipitate about 80% of the total protein concentration in our sample, yet our activity assay for our desired protein indicates that about 95% of our desired protein is still soluble.

- At 80% ammonium sulfate all of our desired protein has precipitated. Thus, from these results we would do the following:
 1. Add ammonium sulfate to our sample to a concentration of 30% saturation
 2. Centrifuge and discard the pellet
 3. Add ammonium sulfate to 80% saturation
 4. Centrifuge and keep the pellet. Resuspend the pellet in buffer to solubilize the protein.
- We would expect about a 5-fold purification with about 95% yield.

Column Chromatography - Ion exchange; Dialysis and Concentration

Column chromatography

After initial fractionation steps the typical procedure is to move to *column chromatography*.

- In column chromatography we have a glass tube (column) which is filled with a material ("resin") which has certain physical/chemical characteristics.
- These characteristics allow it to interact in various ways with different proteins.
- Some common types of chromatographic resins include:
 1. Ion exchange
 2. Hydrophobic
 3. Gel filtration
 4. Affinity

Ion exchange

Ion exchange resins contain *charged groups*.

- These may be *acidic* in nature (in which case the resin is a *cation* exchanger)
- or *basic* (in which case it is an *anion* exchanger).
- Cation and anion exchangers may be broken down further into *weak* and *strong* exchangers (reflecting binding affinity).

Type of exchanger	Functional group	Common name
Weak cation exchanger	carboxymethyl	CM cellulose/sephadex
Strong cation exchanger	sulfopropyl	SP sephadex
Weak anion exchanger	diethylaminoethyl	DE cellulose/sephadex
Strong anion exchanger	quaternary amine	QAE sephadex

- Usually, samples are loaded under *low ionic strength* conditions and bound material is eluted using either a step or gradient elution of buffer with *higher ionic strength*.
- Generally speaking, a protein will bind to a cation exchange resin if the buffer pH is *lower than the isoelectric point (pI)* of the protein, and will bind to an anion exchange resin if the pH is *higher than the pI*.

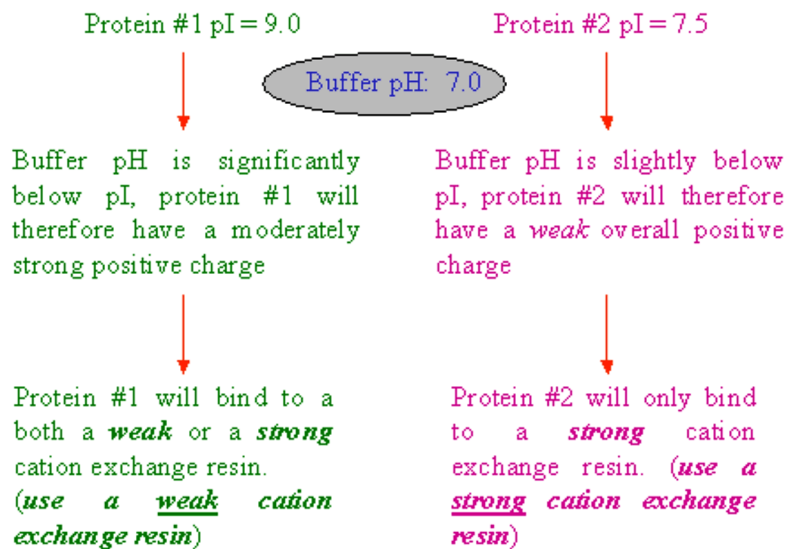


Figure 4.1.3: Protein binding to resins

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

Generally speaking, ion exchange columns are short and fat in dimensions.

Elution of proteins from ion exchange resins

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

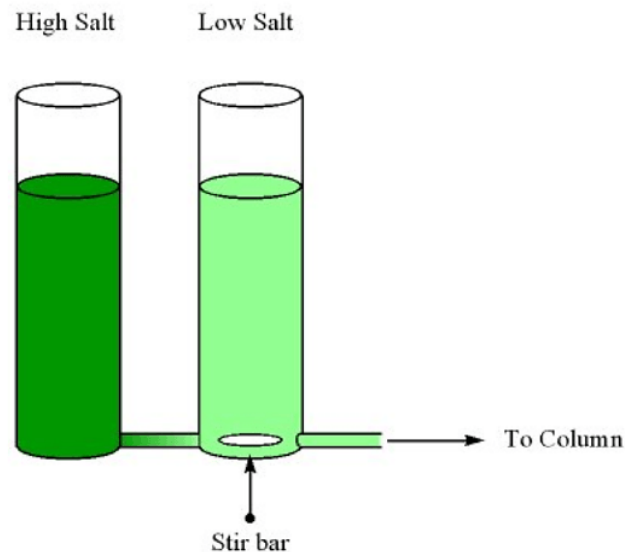


Figure 4.1.4: Gradient maker

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

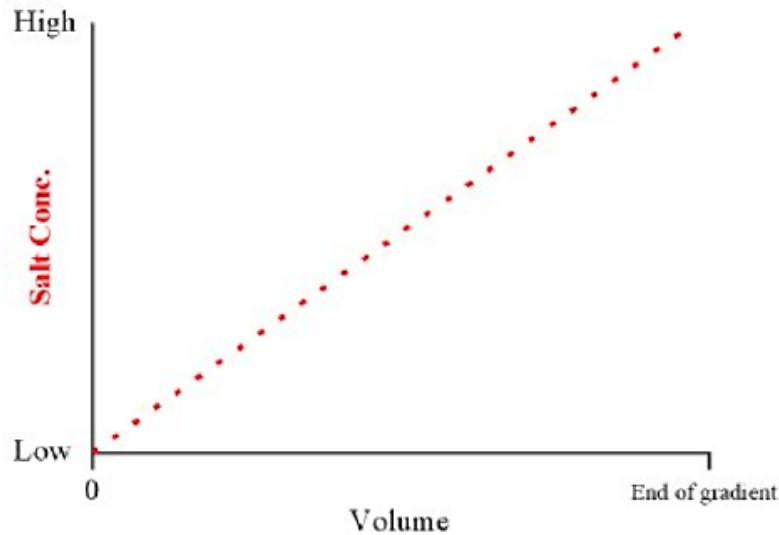


Figure 4.1.5: Salt concentration and volume

- If we know the concentration range of salt over which a protein of interest will elute we can simply elute with a buffer containing that concentration of salt. This is known as a **step elution**.
- Step elutions are generally faster to run, and elute the protein in a smaller overall volume than with gradient elutions. They generally work best when contaminants elute at a significantly different salt concentration than the protein of interest

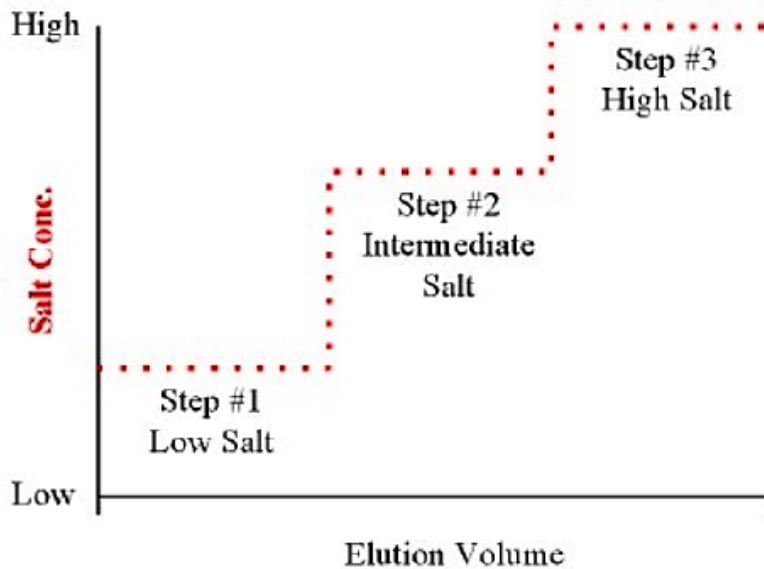


Figure 4.1.6: Step elution

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

Dialysis

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

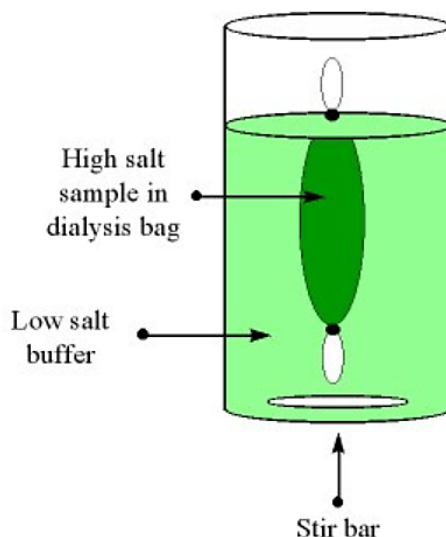


Figure 4.1.7: Dialysis

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

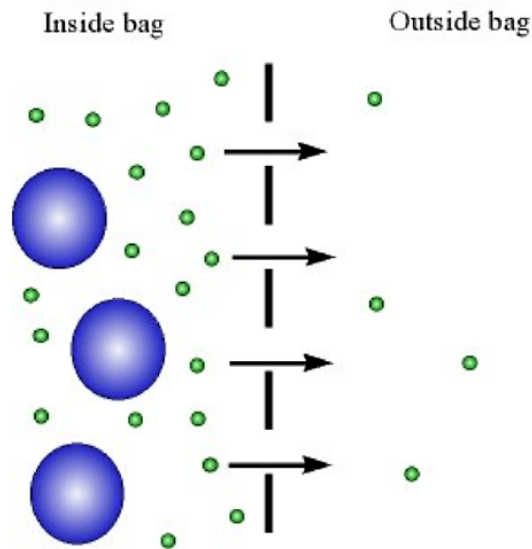


Figure 4.1.8: Salt diffusion

- At equilibrium the salt concentration of the sample can be calculated as follows:

$$\frac{(\text{sample volume}) \times (\text{sample salt concentration}) + (\text{buffer volume}) \times (\text{buffer salt concentration})}{\text{total volume}} = \text{final salt concentration}$$

Note

Often the buffer salt concentration is 0 M

- The buffer volume for the dialysis is a function of the required final concentration of salt in the sample

Example 4.1.1:

Dialysis example

We have a 10ml protein sample from an ion exchange column elution pool which contains 1.0M NaCl. For our next step in the purification we can have no more than 1mM NaCl in the sample.

$$\frac{(\text{sample vol}) * (\text{sample salt conc}) + (\text{buffer vol}) * (\text{buffer salt conc})}{\text{total vol}} = \text{final salt conc}$$

$$\frac{(0.01\text{L}) * (1.0\text{M}) + (\text{buffer vol}) * (0)}{\text{total vol}} = 0.001\text{M}$$

$$\frac{(0.01\text{L}) * (1.0\text{M})}{0.001\text{ M}} = \text{total vol}$$

$$10\text{ L} = \text{total vol}$$

Therefore, the required buffer volume would be (total vol - sample vol) = 9.990 L (or ~ 10 L)

- Thus, if we dialyzed 10mls of sample (with 1.0M NaCl conc) in 10 L of water after equilibrium the NaCl concentration in the sample would be 1.0 mM.

- Note that in the above example this would commonly be referred to as a "1:1,000" dialysis.
- Suppose that we don't want to make up 10 L of buffer? We can actually achieve the same results with two sequential "1:32" dialyses (i.e. the square root of the 1:1,000 dialysis - in other words, two sequential 1:32 dialyses is equivalent to a single 1:1,000 dialysis):

First dialysis versus 310 ml of buffer: sample NaCl conc will be $(10 * 1.0) / (320) = 31\text{ mM}$

Second dialysis versus 310 ml of buffer: sample NaCl conc will be $(10 * 0.031) / (320) = 0.97\text{ mM}$

Thus, instead of making 10 L of buffer, we could make only 620 ml and achieve the same results with two dialysis steps

- In this case, removing the salt would take twice as long, i.e. we need to perform two dialysis steps. How long does dialysis take?

A useful rule of thumb is that for most types of dialysis tubing the dialysis is 80% complete after four hours

- One consequence of dialysis to watch out for is that while salt ions are moving out of the bag, water molecules are moving into the bag. Thus the volume of sample may actually increase (the bag will swell) and, therefore, the protein concentration will decrease
- In the extreme case, the bag may actually swell to the point of rupture. Therefore, *it is a good idea not to fill the bag completely, but leave a void to allow for potential swelling.*

Concentration

- What if our protein sample is too dilute for our needs? How can we concentrate our samples?
- One common method is, again, to use a semi-permeable membrane for this purpose.
- A very simple method is to place our sample in a dialysis bag and coat it with a high molecular weight solute which can readily be dissolved by the buffer.
- For example, polyethylene glycols and polyvinyl pyrrolidones can have very large molecular masses (i.e. 20,000 Da). These compounds are also readily dissolved in water. If our sample in a dialysis bag is coated with dry forms of the above polymers, water will leave the dialysis bag (it can go through the pores) and hydrate the polymers. The result is a decrease in volume of buffer in the dialysis bag (the protein will be concentrated).
- In another variation, the semi-permeable membrane is manufactured into a flat disk and placed at the bottom of a container

which holds our sample. In one method the container is pressurized and forces buffer out of the container (protein is retained and is concentrated). In another method, the vessel is centrifuged and the centripetal force achieves the same goal as pressure in the previous example.

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

- With the pressure type concentrators, **dialysis and concentration can be achieved in tandem**. For example, the sample can be concentrated and then buffer added to the sample. The sample is then concentrated again. Every time buffer is added the salt concentration is reduced. After repeated cycles of this, the salt concentration is at the desired level and the sample is concentrated to the desired final volume.

Gel Filtration, Affinity and Hydrophobic resins; Preparation of Resin, Plumbing

Gel filtration

Gel filtration does not rely on any chemical interaction with the protein, rather it is based on a **physical** property of the protein - that being the *effective molecular radius* (which relates to mass for most typical globular proteins).

- Gel filtration resin can be thought of as beads which contain pores of a defined size range.
- Large proteins which cannot enter these pores pass around the *outside* of the beads.
- Smaller proteins which can enter the pores of the beads have a longer, tortuous path before they exit the bead.
- Thus, a sample of proteins passing through a gel filtration column will *separate based on molecular size*: the big ones will elute first and the smallest ones will elute last (and "middle" sized proteins will elute in the middle).

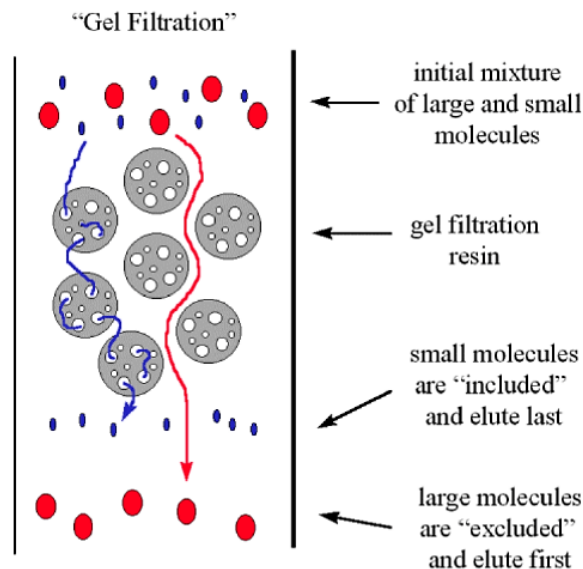


Figure 4.1.9: Gel filtration

- If your protein is unusually "small" or "large" in comparison to contaminating proteins then gel filtration may work quite well.

Where will a protein elute in a gel filtration experiment?

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

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

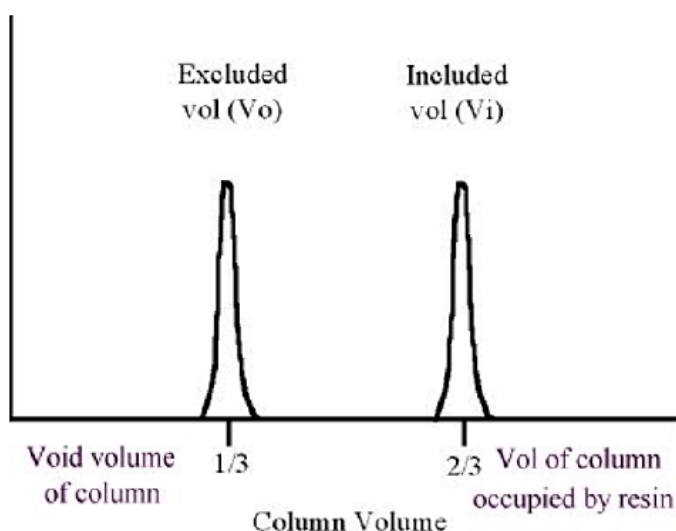


Figure 4.1.10: Excluded vs. included volume

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

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

Affinity chromatography

Affinity chromatography is a general term which applies to a wide range of chromatographic media. It can be basically thought of as some inert resin to which has been attached some compound which has a *specific affinity* for your protein of interest.

- Thus, a specific antibody attached to an inert resin would be a type of affinity chromatography.
- Other examples might include: a protease inhibitor attached to some matrix, designed to bind a specific protease
- a cofactor bound to some matrix, designed to bind to a particular enzyme
- a metal ion bound to a matrix, designed to chelate a protein with a metal binding site, and so on.

In each case, the type of resins used and the method of attachment may vary, as will the method of elution. One generalization regarding method of elution is that the bound ligand can be competed off of the column's functional group by including in the elution buffer a high concentration of the free functional group. For example, if the functional group of the column is a cofactor, then the bound protein can be competed off the column by passing a buffer containing a high concentration of cofactor (or cofactor analog) through the column.

Other methods of elution include changing the buffer conditions such that the protein is no longer in the native state (since it is the native state which confers the structure required for the specific binding interaction). This can be achieved by changing pH

or by adding denaturing agents such as urea or guanidine.

With affinity chromatography, typically the purification achieved in a single step can be dramatic - on the order of several thousand fold. Single step purifications with specific affinity columns are not unheard - in fact it is an ideal goal of purification - *a matrix which recognizes only the protein of interest and none other.*

Hydrophobic resins

Hydrophobic resins contain a non-polar functional group, such as an alkane or aromatic group.

- Many proteins are able to sequester such groups on their surface and this exclusion from solvent provides the basis of the binding energy (i.e. the "hydrophobic effect").
- This interaction is enhanced by increasing ionic strength, such that proteins may *bind under high salt conditions* and *elute under low salt conditions*.
- As such these columns may be used to not only provide purification, but to *desalt* samples (for example after an initial ammonium sulfate precipitation).
- It is usually not possible to predict in advance which particular resin will bind a given protein, this is usually determined empirically. However, the longer the alkane, or the larger the aromatic compound, the stronger the binding typically will be.

Due to the nature of hydrophobic interactions and ionic strength, hydrophobic chromatography and ion exchange chromatography can be conveniently used sequentially. For example, after ion exchange the protein is in high salt conditions, thus it can be loaded directly onto a hydrophobic column. Conversely, a hydrophobic column is eluted in low salt, which is a requirement for binding to an ion exchange resin.

A distinction should be noted between hydrophobic interaction chromatography and reverse phase chromatography

- Hydrophobic interaction chromatography is performed in aqueous solvent conditions and changes in ionic strength are used to elute the column. The protein typically binds in the native state via hydrophobic groups located on the surface of the protein. The native state is retained during the elution conditions
- Reverse phase chromatography utilizes a hydrophobic solvent (typically acetonitrile) and the binding of a ligand is a function of the phase partition between the hydrophobic nature of the solvent and column functional group. Proteins are typically denatured in such solvents and bind due to the hydrophobic nature of the entire polypeptide sequence. Since the majority of hydrophobic groups are located in the core of globular proteins, the binding is related to the denaturation of the protein and the accessibility of these groups to the column functional groups. Proteins can be purified using reverse phase chromatography, but usually must be refolded in some way to regain functionality (i.e. the native state)

Preparation of resins

The steps in preparing a chromatographic resin typically involve:

1. Hydration of resin
 2. Decanting fines
 3. Equilibrating the resin and preparing a slurry
 4. Degassing the slurry
- Resins come either dry or preswollen. If they are dry they need to be hydrated. This is usually accomplished by mixing the dry resin with buffer and letting it hydrate slowly overnight (or faster at higher temperatures)
 - After the resin has hydrated and settled, very fine particles will settle at the top. These "fines" slow the flow rate of the packed resin. The settled resin is therefore carefully decanted to discard these fines.
 - The resin is then equilibrated in the buffer to be used for the analysis. Equilibration usually involves pH'ing the resin, or buffer exchanges. Never use a stir bar when pH'ing the resin (it can mechanically shear the resin and produce fines), rather stir the resin slurry with a stir rod.
 - After the equilibrated resin has settled, an equal volume of buffer is added to produce a 50% *slurry* of resin. This is usually "thin" enough to allow air bubbles to escape when packing the column.
 - Finally, the slurry is degassed prior to packing the column. This will help minimize the formation of air bubbles.

Packing the column

Low pressure columns are typically packed using gravity.

- Add a small amount of buffer to the bottom of the column.
- Place a **packing reservoir** on the top of the column. Since we will be using a 50% slurry we will have a volume which is 2x the column volume and its best to pour the resin in all at one time. Thus, the packing reservoir should have a volume equal to, or greater, than the column volume.
- Carefully pour the resin slurry into the packing reservoir/column, avoiding the introduction of air bubbles as much as possible
- Let the column sit for about 5 minutes to allow large air bubbles to escape
- Open the column valve at the bottom and allow the column to pack under gravity
- Note the top of the resin bed. It will move down as the column packs. When the column is packed the top of the resin bed will no longer move down.

Plumbing

Chromatography systems may be run using only gravity and a beaker to collect the appropriate fraction. Most common systems, however, will include the following:

- A pump. Usually a peristaltic pump with variable flow rate and a communications port for a controller. *The pump is usually set up to push buffer through the column, rather than sucking buffer out of the column (which can cause a low pressure condition with production of air bubbles)*
- A detector. This is typically a UV (A_{280}) detector. Most detectors are of the two-cell type - meaning that you can have plain buffer as a blank in the detector while analyzing your column fractions. The detector sends the absorbance information to a chart recorder to be displayed (printed)
- A fraction collector. This allows you to collect fractions either by number of drops (~30 per ml) or by time. In conjunction with a controllable pump, time collection translates to volume. The fraction collector will typically have an communications port to output a signal when it changes fractions and to receive commands from the detector/controller on some sophisticated systems.
- A chart recorder. This will print a continuous trace of the detector output and the fraction collector event marker (signalling when a fraction changes). Fractions can also be read individually on a UV spectrophotometer if a chart recorder is unavailable.

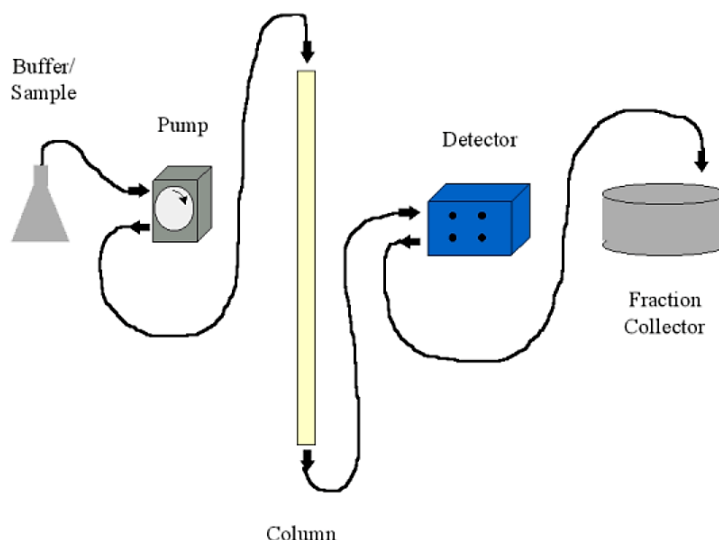


Figure 4.1.11: Chromatography setup

- If a gravity system is used, a safety loop should be installed to prevent the column from drying up if the buffer is used up when the column is unattended

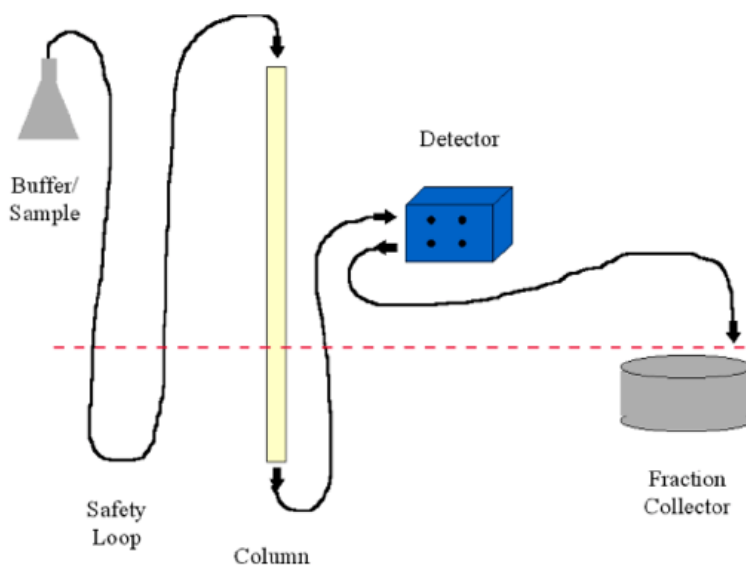


Figure 4.1.12: Safety loop

Note that the bottom of the safety loop is lower than the outlet to the fraction collector.

Running the Experiment, Resolving Peaks

The following represents an example of a low pressure liquid chromatography (ion exchange resin) experiment.

Sample:

- Volume = 90 mls
- $A_{280} = 1.8$
- Total $A_{280} = 162$

Column:

- DE-52 (diethylaminoethyl cellulose; anion exchange)
- Size = 1.0 x 12.7 cm

$$\pi \times (\text{radius})^2 \times \text{height}$$

- Volume = = 40 mL

Fraction Collector:

- 10 mls / fraction (~300 drops/fraction)

The chromatogram for this experiment looked like this:

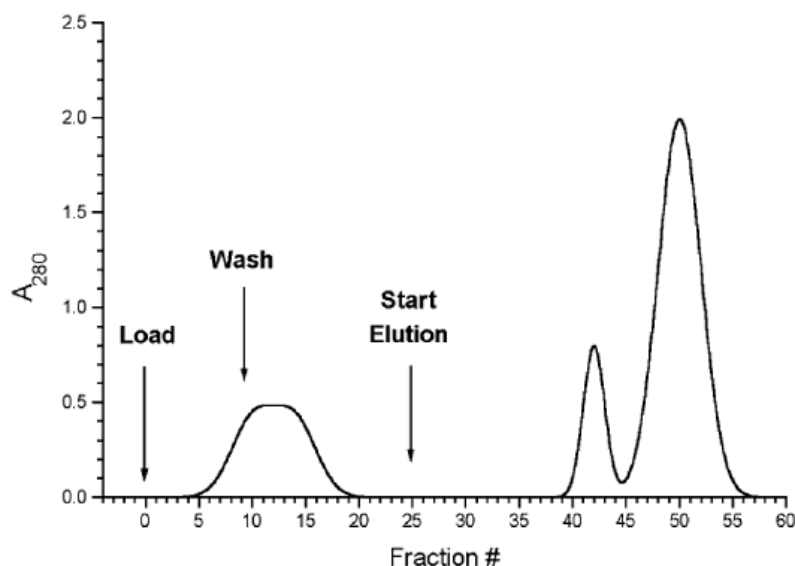


Figure 4.1.13: Chromatogram

The following events took place during this chromatography run:

- Note the tick marks on the chromatogram.
 - The "event" marker from the fraction collector notifies the chart recorder when a tube change takes place.
 - The experiment begins with the tick next to the '0' on the x-axis ("tick 0"); this indicates the **start** of fraction (tube) number 1.
 - The next tick mark ("tick 1") indicates the **end** of fraction number 1, and the **start** of fraction number 2.
 - Thus, **fractions span the gap between the tick marks**.
- The sample loading is begun at tick 0.
- Sometime during fraction 5 we begin to notice the absorbance of the column effluent increasing
 - It has taken about (5 fractions $\times$ 10mls per fraction) or 50 mls from the start of loading until the detector notes any absorbance.
 - This compares well with the fact that the column volume is about 40 mls and there is some volume associated with the tubing going in and out of the column.
 - Thus, this 'delay' from sample load to sample detection is the **dead volume** of the system
- Obviously, some material is not binding to the resin during the loading step. This is the **flow-through**. Is this some component of the sample which does not have affinity for the resin, or, does it represent that we have exceeded the capacity of the resin?
 - If we have exceeded the capacity of the resin, then the flow-through will have an A_{280} similar to the sample being loaded
 - Also, prior to exceeding the capacity, the flow-through will have some characteristic A_{280} which will then transition to another A_{280} (that of the loaded sample), resulting in a **double-plateau** chromatogram.
 - In the above experiment the flow-through plateaus around $A_{280}=0.5$ or about 25% of the absorbance of the load. This would seem to indicate that a component, or component(s), representing one quarter of our sample, does not have affinity for the resin in the column
- Around fraction 9 we begin to wash the column
 - This makes sense because 9 fractions $\times$ 10mls per fraction = 90 mls have loaded and this is equivalent to our original sample volume (i.e. all the sample has loaded)
 - The column is typically washed using the same buffer conditions in the protein sample
- Around fraction 14 we note the A_{280} begins to decrease

- This makes sense given that we determined the dead volume of the system to be approximately 50 mls or 5 fractions. Thus, a wash which was begun at fraction 9 is observed to decrease the absorbance around fraction 14
- We continue washing the column until the A_{280} approaches 0 (*baseline*). In other words, all of the non-binding material in the sample has been washed away

7. After the A_{280} comes back down to baseline we begin our *elution* protocol. In this particular experiment we will use a linear gradient of increasing salt (NaCl) concentration (in wash buffer) to compete off the material bound to the ion exchange resin.

8. Our elution has produced two peaks: a small peak centered around fraction 42 and a larger peak centered around fraction 50

- We will have to assay each peak (and the flow through) to find out where our protein of interest has gone
- The two *elution peaks* are fairly well resolved. We could combine fractions 40-44 and call that "peak 1", and combine fractions 46-55 and call that "peak 2".

9. *Is there any material left on the column?* The integrated areas (i.e. summing the A_{280} 's of each fraction in a pool) of the flow-through, peak 1 and peak 2 are as follows

- Flow through: 4
- Peak 1: 2
- Peak 2: 10
- This gives a total integrated area of 16. Each fraction is 10 mls, so this gives a total $A_{280} = 16 \times 10 = 160$ which is quite close to the total A_{280} of our loaded sample.
- In other words, it looks like our chromatogram is accounting for all the components in our original sample.

10. If our protein of interest was actually peak 1 (and if our yield was 100%), then this column has provided an eight fold purification ($2 \times 10 / 162$).

Resolving peaks

- Contaminating peaks will not necessarily be completely separated from the peak which contains our protein of interest
- In the following picture there are two components being resolved, and they are present in equimolar amounts (thus, the starting purity is 50%). The yield and purity are listed for the situation where we were to pool each peak by splitting at the midpoint between them (in this particular example the yield and purity are identical in each case)

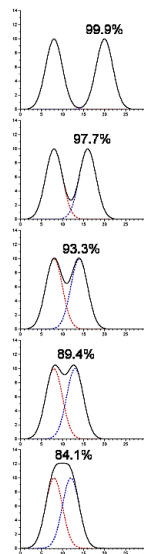


Figure 4.1.14: Contaminating peaks

- This gives you some idea of the amount of cross-contamination in each peak as a function of their separation from one

another.

- Software to fit gaussians to a chromatogram can provide this type of information

Pooling for purity versus yield

Do we try to pool to maximize yield or to maximize purity?

Usually, you will probably be pooling fractions in such a way as to maximize the recovery of your protein of interest. However, you always have the option of pooling to increase purity, and if you have lots of protein to work with this may allow you to achieve the desired purity with fewer steps. Here's an example of how it's done:

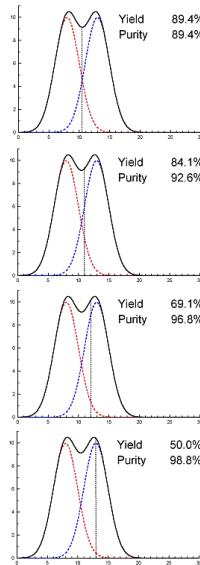


Figure 4.1.15: Yield vs. purity

- These are all the same chromatogram, however, we can pool them differently to get better purity (at the expense of yield)
- The blue peak is the peak of interest and it is not resolved from a contaminating peak (in red).
- The vertical line represents the left-most fraction we use to pool the peak (we pool all fractions to the right of the vertical line to get our protein of interest)
- In the last panel we see that we can achieve about 98.8% purity if we are willing to part with half our protein!

Monitoring the purification

How do you know when you are finished purifying a protein?

There are several criteria. One criteria is that we cannot improve upon the **specific activity** of our sample. This value refers to the **functional activity of our sample in relationship to the total protein concentration of the sample**.

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

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

- In the initial stages of purification we will probably see a variety of bands, of various molecular weights, on our gel.
- After the different purification steps, we should see the disappearance of certain bands concomitant with the increasing concentration of a certain band (or bands) representing our protein.
- If we have successfully purified our protein (and if it is a single polypeptide) we should arrive at a constant specific activity and a *single band on a gel*.
- Analytical methods like HPLC or densitometer scanning of a stained gel can give us a quantitative idea of the purity of our

final sample.

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

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

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

- [8.1: The Two Criteria for Aromaticity](#) is licensed [CC BY-NC-SA 4.0](#).

3.2: Cell Disruption

There are several ways to break open cells.

- Lysis methods include lowering the ionic strength of the medium cells are kept in. This can cause cells to swell and burst. Mild surfactants may be used to enhance the efficiency of lysis. Most bacteria, yeast, and plant tissues, which have cell walls, are resistant to such osmotic shocks, however, and stronger disruption techniques are often required.
- Enzymes may be useful in helping to degrade the cell walls. [Lysozyme](#), for example, is very useful for breaking down bacterial walls. Other enzymes commonly employed include cellulase (plants), glycanases, proteases, mannases, and others.
- Mechanical agitation may be employed in the form of tiny beads that are shaken with a suspension of cells. As the beads bombard the cells at high speed, they break them open. *Sonication* (20-50 kHz sound waves) provides an alternative method for lysing cells. The method is noisy, however, and generates heat that can be problematic for heat-sensitive compounds.
- Another means of disrupting cells involves using a "cell bomb". In this method, cells are placed under very high pressure (up to 25,000 psi). When the pressure is released, the rapid pressure change causes dissolved gases in cells to be released as bubbles which, in turn, break open the cells.
- *Cryopulverization* is often employed for samples having a tough extracellular matrix, such as connective tissue or seeds. In this technique, tissues are flash-frozen using liquid nitrogen and then ground to a fine powder before extraction of cell contents with a buffer.

Whatever method is employed, the crude lysates obtained contain all of the molecules in the cell, and thus, must be further processed to separate the molecules into smaller subsets, or fractions.

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

- [The Effect of Temperature on the \$\text{Co}\(\text{H}_2\text{O}\)_6^{2+}/\text{CoCl}_4^{2-}\$ Equilibrium](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

3.3: Fractionation

Fractionation of samples typically starts with centrifugation. Using a centrifuge, one can remove cell debris, and fractionate organelles, and cytoplasm. For example, nuclei, being relatively large, can be spun down at fairly low speeds. Once nuclei have been sedimented, the remaining solution, or supernatant, can be centrifuged at higher speeds to obtain the smaller organelles, like mitochondria. Each of these fractions will contain a subset of the molecules in the cell. Although every subset contains fewer molecules than does the crude lysate, there are still many hundreds of molecules in each. Separating the molecule of interest from the others is where chromatography comes into play. We will consider several separation techniques.

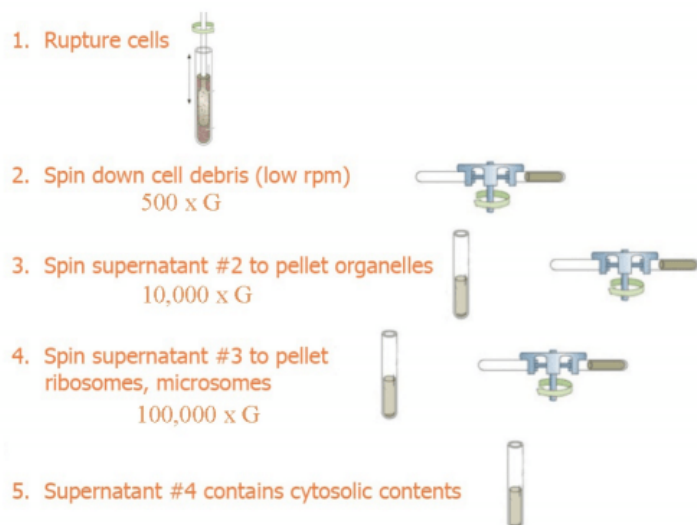


Figure 9.2.1: *Fractionation by centrifugation*

Many chromatographic techniques are performed in “columns.” These are tubes containing the material (called the “support”) used to perform the separation. Supports are designed to exploit the chemical, or size, differences of the many molecules in a mixture. Columns are “packed” (filled) with the support and a buffer or solvent carries the mixture of compounds to be separated through the support. Molecules in the sample interact differentially with the support and consequently, will travel through it with different speeds.

This page titled [3.3: Fractionation](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Kevin Ahern and Indira Rajagopal](#).

- [The Effect of Temperature on the NO₂/N₂O₄ Equilibrium](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

3.4: Chromatography

This page titled [3.4: Chromatography](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

3.4.1: Affinity Chromatography

Affinity chromatography is a very powerful technique that exploits the binding affinities of target molecules (typically proteins) for substances covalently linked to beads. For example, if one wanted to separate all of the proteins in a sample that bound to ATP from proteins that do not bind ATP, one could covalently link ATP to support beads and then pass the sample through column. All proteins that bind ATP will “stick” to the column, whereas those that do not bind ATP will pass quickly through it. The proteins adhering to the column may then released from the column by adding ATP.

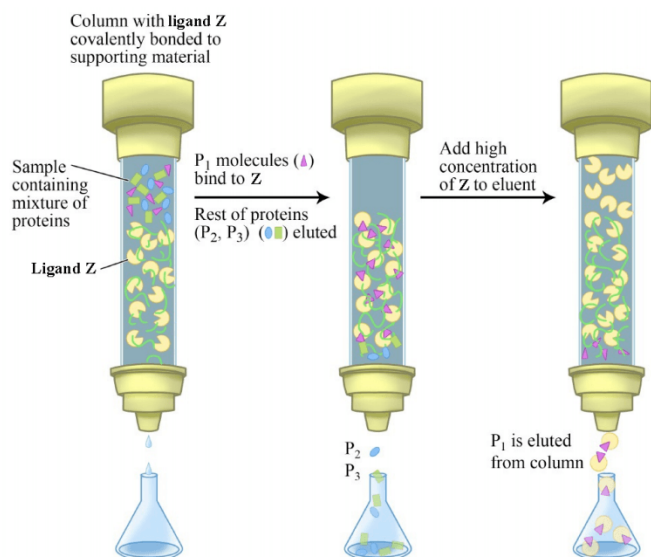


Figure 9.5.1: Affinity Chromatography

This page titled **3.4.1: Affinity Chromatography** is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Kevin Ahern](#) and [Indira Rajagopal](#).

- [The \$\text{Fe}\(\text{SCN}\)_2^{2+} / \text{Fe}\(\text{SCN}\)^{2+}\$ Equilibrium - Expanded Version](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

3.4.2: Gel Exclusion Chromatography

Gel exclusion chromatography (also called molecular exclusion chromatography, size exclusion chromatography, or gel filtration chromatography) is a low resolution isolation method that employs a cool “trick.” This involves the use of beads that have tiny “tunnels” in them that each have a precise size. The size is referred to as an “exclusion limit,” which means that molecules above a certain molecular weight will not fit into the tunnels. Molecules with sizes larger than the exclusion limit do not enter the tunnels and pass through the column relatively quickly by making their way between the beads. Smaller molecules, which can enter the tunnels, do so, and thus, have a longer path that they take in passing through the column. Because of this, molecules larger than the exclusion limit will leave the column earlier, while those that pass through the beads will elute from the column later. This method allows separation of molecules by their size.

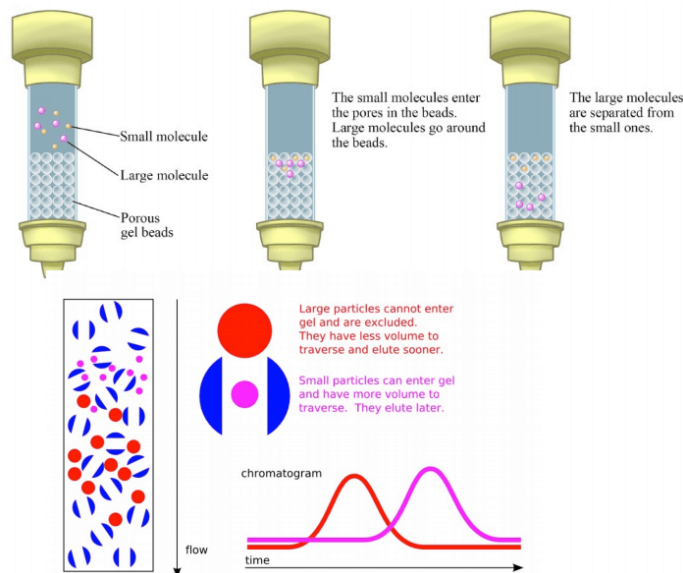


Figure 9.4.1: Gel Exclusion Chromatography

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

- [The \$\text{Fe}\(\text{SCN}\)_2^{\ominus}/\text{Fe}\(\text{SCN}\)_2^{\ominus}\$ Equilibrium](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

3.4.3: Ion Exchange Chromatography

In [ion exchange chromatography](#), the support consists of tiny beads to which are attached chemicals possessing a charge. Each charged molecule has a counter-ion. The figure shows the beads (blue) with negatively charged groups (red) attached. In this example, the counter-ion is sodium, which is positively charged. The negatively charged groups are unable to leave the beads, due to their covalent attachment, but the counter-ions can be “exchanged” for molecules of the same charge. Thus, a cation exchange column will have positively charged counter-ions and positively charged compounds present in a mixture passed through the column will exchange with the counter-ions and “stick” to the negatively charged groups on the beads. Molecules in the sample that are neutral or negatively charged will pass quickly through the column. On the other hand, in anion exchange chromatography, the chemical groups attached to the beads are positively charged and the counter-ions are negatively charged. Molecules in the sample that are negatively charged will “stick” and other molecules will pass through quickly. To remove the molecules “stuck” to a column, one simply needs to add a high concentration of the appropriate counter-ions to displace and release them. This method allows the recovery of all components of the mixture that share the same charge.

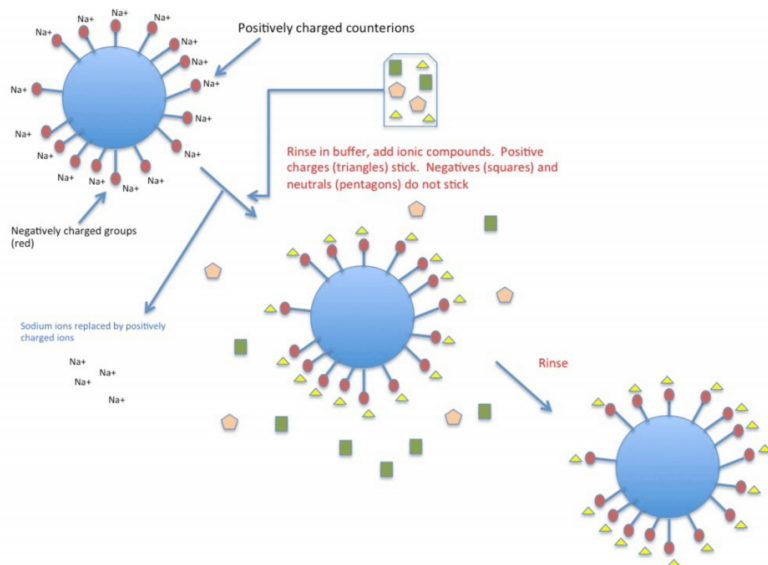


Figure 9.3.1: Cation exchange chromatography

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

- [The Electrolysis of Water](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

3.5: Electrophoresis

DNA molecules are long and loaded with negative charges, thanks to their phosphate backbones. Electrophoretic methods separate large molecules, such as DNA, RNA, and proteins based on their charge and size. For DNA and RNA, the charge of the nucleic acid is proportional to its size (length). For proteins, which do not have a uniform charge, a clever trick is employed to make them mimic nucleic acids.

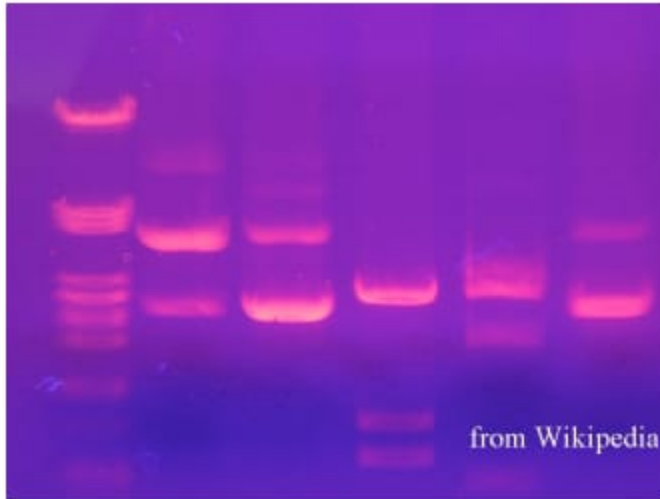


Figure 9.8.1: DNA fragments separated by agarose gel

electrophoresis

Agarose

Agarose gel electrophoresis is a method for separating nucleic acids. It is worth noting that nucleic acids are the largest molecules found in cells, in some cases by orders of magnitude. Agarose provides a matrix which encases a buffer. The matrix provides openings for macromolecules to move through and the largest macromolecules have the most difficult time navigating, whereas the smallest macromolecules slip through it the easiest. Unlike column chromatography, electrophoresis uses an electric current as a force to drive the molecules through the matrix. Since the size to charge ratios for DNA and RNA are constant for all sizes of these nucleic acids, the size per force is also constant (since force is directly proportional to charge), so the molecules simply sort on the basis of their size - the smallest move fastest and the largest move slowest. Visualization of the DNA fragments in the gel is made possible by addition of a dye, such as ethidium bromide that fluoresces under ultraviolet light.

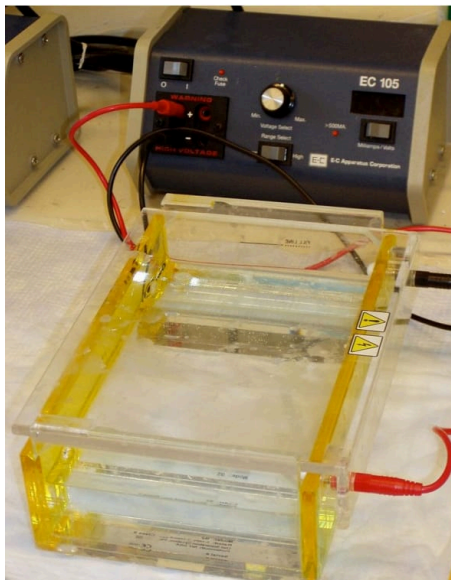


Figure 9.8.2: Agarose gel electrophoresis

SDS-PAGE

Like DNA and RNA, proteins are large macromolecules. Proteins, however, vary tremendously in their charge. Whereas double-stranded DNA is rod-shaped, most proteins are globular (folded up). Further, proteins are considerably smaller than nucleic acids, so the openings of the matrix of the agarose gel are simply too large to effectively provide separation. Consequently, unlike nucleic acids, proteins cannot be effectively separated by electrophoresis on agarose gels. To separate proteins by electrophoresis, one must make several modifications. First, a matrix made by polymerizing and crosslinking acrylamide units is employed. One can adjust the pore size of the matrix readily by changing the percentage of acrylamide in the gel. Higher percentages of acrylamide create smaller pores and are more effective in separating smaller molecules, whereas lower percentages of acrylamide reverse that. Second, proteins must be physically altered to “present” themselves to the matrix like the negatively charged rods of DNA. This is accomplished by treating the proteins with the detergent called SDS (sodium dodecyl sulfate). SDS denatures the proteins so they assume a rod-like shape and the SDS molecules coat the proteins such that the exterior surface is loaded with negative charges proportional to the mass, just like the backbone of DNA. Third, a “stacking gel” may be employed at the top of the gel to provide a way of compressing the samples into a tight band before they enter the main polyacrylamide gel (called the resolving gel). Just as DNA fragments get sorted on the basis of size (largest move slowest and smallest move fastest), the proteins migrate through the gel matrix at rates inversely related to their size. Upon completion of the electrophoresis, there are several means of staining to visualize the proteins on the gel. They include reagents, such as Coomassie Brilliant Blue or silver nitrate (the latter is much more sensitive than Coomassie Blue staining and can be used when there are very small quantities of protein).



Figure 9.8.3: SDS-Page Separation of proteins

Isoelectric Focusing

Proteins vary considerably in their charges and, consequently, in their pI values (pH at which their charge is zero). Separating proteins by isoelectric focusing requires establishment of a pH gradient in an acrylamide gel matrix. The matrix's pores are adjusted to be large to reduce the effect of sieving based on size. Molecules to be focused are applied to the gel with the pH gradient and an electric current is passed through it. Positively charged molecules, for example, move towards the negative electrode, but since they are traveling through a pH gradient, as they pass through it, they reach a region where their charge is zero and, at that point, they stop moving. They are at that point attracted to neither the positive nor the negative electrode and are thus "focused" at their pI. By using isoelectric focusing, it is possible to separate proteins whose pI values differ by as little as 0.01 units.

2-D Gel Electrophoresis

Both SDS-PAGE and isoelectric focusing are powerful techniques, but a clever combination of the two is a powerful tool of proteomics - the science of studying all of the proteins of a cell/tissue simultaneously. In 2D gel electrophoresis, an extract containing the proteins is first prepared. One might, for example, be studying the proteins of liver tissue. The liver cells are lysed and all of the proteins are collected into a sample. Next, the sample is subjected to isoelectric focusing as described earlier, to separate the proteins by their pI values. Next, as shown on the previous page, the isoelectric gel containing the separated proteins is rotated through 90° and placed on top of a regular polyacrylamide gel for SDS-PAGE analysis (to separate them based on size). The proteins in the isoelectric gel matrix are electrophoresed into the polyacrylamide gel and separation on the basis of size is performed. The product of this analysis is a 2D gel, in which proteins are sorted by both mass and charge.

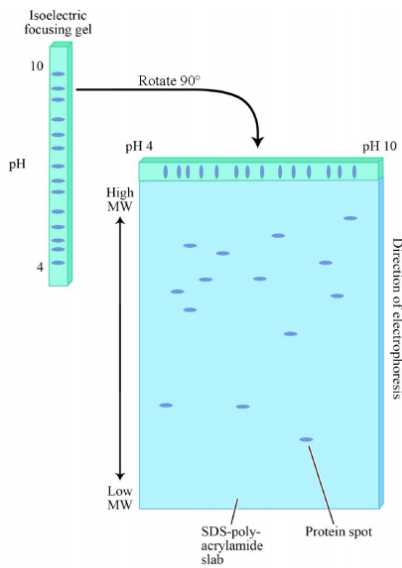


Figure 9.8.4: 2D Gel Electrophoresis

The power of 2D gel electrophoresis is that virtually every protein in a cell can be separated and appear on the gel as a distinct spot. In the figure, spots in the upper left correspond to large positively charged proteins, whereas those in the lower right are small negatively charged ones. It is possible using high-throughput mass spectrometry analysis to identify every spot on a 2D gel. This is particularly powerful when one compares protein profiles between different tissues or between the samples of the same tissue treated or untreated with a particular drug. Comparison of a 2D separation of a non-cancerous tissue with a cancerous tissue of the same type provides a quick identification of proteins whose level of expression differs between them. Information such as this might be useful in designing treatments or in determining the mechanisms by which the cancer arose.

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

- [The Photochemical Blue Bottle Demonstration](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

3.5.1: Blotting and Hybridization

Bands of DNA in an electrophoretic gel form only if most of the DNA molecules are of the same size, such as following a PCR reaction, or restriction digestion of a plasmid. In other situations, such as after restriction digestion of chromosomal (genomic) DNA, there will be a large number of variable size fragments in the digest and it will appear as a continuous smear of DNA, rather than distinct bands. In this case, it is necessary to use additional techniques to detect the presence of a specific DNA sequence within the smear of DNA separated on an electrophoretic gel. This can be done using a “Southern Blot”.

Southern Blots

A **Southern blot** (also called a **Southern Transfer**) is named after Ed Southern, its inventor. In the first step, DNA is digested with restriction enzymes and separated by gel electrophoresis (as discussed above). Then a sheet or **membrane** of nylon or similar material is laid under the gel and the DNA, in its separated position (bands or smear), is transferred to the membrane by drawing the liquid out of the gel, in a process called **blotting** (Figure 3.5.1.1). The blotted DNA is usually covalently attached to the nylon membrane by briefly exposing the blot to UV light. Transferring the DNA to the sturdy membrane is necessary because the fragile gel would fall apart during the next two steps in the process. Next, the membrane is bathed in a solution to **denature** (double stranded made single stranded) the attached DNA. Then a **hybridization** solution containing a small amount of single-stranded **probe** DNA that is complementary in sequence to a target molecule on the membrane. This probe DNA is labeled using fluorescent or radioactive molecules, and if the hybridization is performed properly, the probe DNA will form a stable duplex only with those DNA molecules on the membrane that are exactly complementary to it. Then, the unhybridized probe is washed off and remaining radioactive or fluorescent signal will appear in a distinct band when appropriately detected. The band represents the presence of a particular DNA sequence within the mixture of DNA fragments.

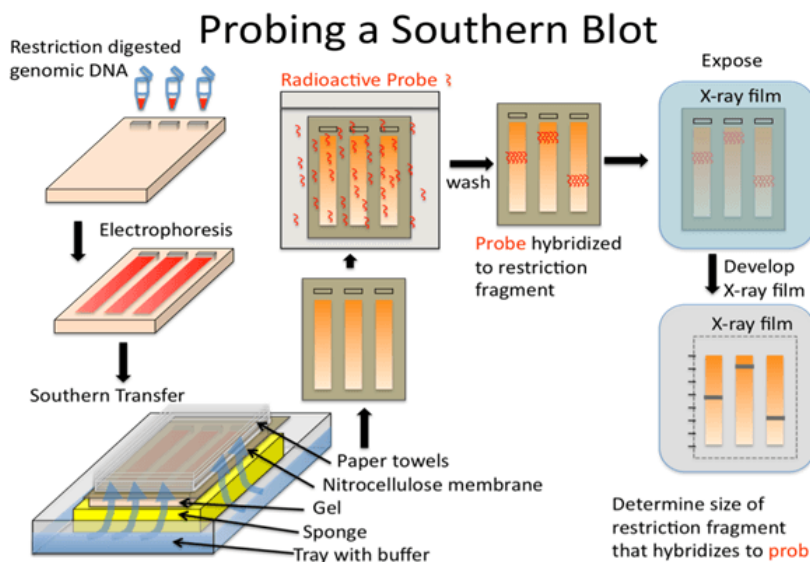


Figure 3.5.1.1: A diagram of Southern blotting.

Genomic DNA that has been digested with a restriction enzyme is separated on an agarose gel, then the DNA is transferred from the gel to a nylon membrane (grey sheet) by blotting. The DNA is immobilized on the membrane, then probed with a **radioactively labeled DNA fragment** that is complementary to a target sequence. After stringent washing, the blot is exposed to X-ray film to detect what size fragment the probe is bound. In this case, the probe bound to different-sized fragments in lanes 1, 2, and 3. In the last image the orange represent the position of the digested DNA, but it is not actually present on the X-ray film. (Original-J. Locke-CC:AN)

The probe is sequence specific (requires complementarity). However, variation in hybridization temperature and washing solutions can alter the **stringency** of the probe. At maximum stringency (higher temperature) hybridization conditions, probes will only hybridize with the exact target sequences that are perfectly complementary (maximum number of hydrogen bonds). At lower temperatures, probes will be able to hybridize to targets to which they do not match exactly, but only are roughly complementary for part of the sequence.

Southern blotting is useful not only for detecting the presence of a DNA sequence within a mixture of DNA molecules, but also for determining the size of a restriction fragment in a DNA sample. Southern blots are useful for detecting fragments larger than those normally amplified by PCR, and when trying to detect fragments that may be only distantly related to a known sequence. Applications of Southern blotting will be discussed further in the context of molecular markers in a subsequent chapter. Southern blotting was invented before PCR, but PCR has replaced blotting in many applications because of its simplicity, speed, and convenience. Following the development of the Southern blot, other types of blotting techniques were invented.

Northern Blots

The **Northern blot** involves the size separation of RNA in gels like that of DNA. Because we wish to determine the native size of the RNA transcript (and because RNA is single stranded) no restriction enzymes are ever used. Because most RNA is single stranded and can fold into various conformations through intra-molecular base pairing, the electrophoresis separation is more haphazard and the bands are often less sharp, compared to that of double stranded DNA.

Western Blots

In a **Western blot**, protein is size separated on a gel (usually an acrylamide gel) before transferring to a membrane, which is then probed with an antibody that specifically binds to an antigenic site on the target protein. This antibody is then detected by other antibodies with some fluorescent or color production marker system. It will also give bands proportional to the amount and size of the target protein (Figure 3.5.1.2).

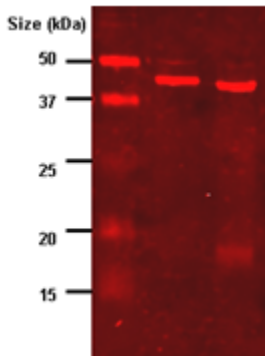


Figure 3.5.1.2: Western Blot using a anti-lipoic acid primary antibody and an IR-dye labelled secondary antibody in *Leishmania major* extracts. (Wikipedia-TimVickers-CC:AS)

A comparison of all three blotting methods is shown in Figure 3.5.1.3.

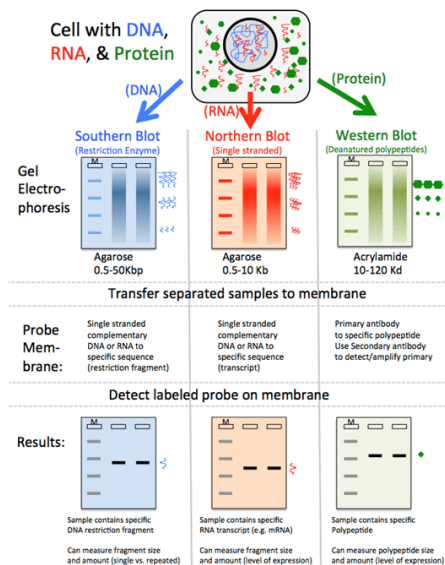


Figure 3.5.1.3: Comparison of Southern, Northern, and Western blots. Size and amount of DNA, RNA, and polypeptides can be determined using similar blotting methods. DNA is in blue, RNA in red, and polypeptides

in **green**. A marker lane is shown in the left of each gel to determine size. A eukaryote cell is shown, but the same methods can be applied to prokaryotes, too. (Original-Locke-CC:AN)

This page titled [3.5.1: Blotting and Hybridization](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Todd Nickle and Isabelle Barrette-Ng](#).

- [MO9. Bonding Between Different Atoms](#) by [Chris Schaller](#) is licensed [CC BY-NC 3.0](#).

3.5.2: DNA Analysis - Gel Electrophoresis

A solution of DNA is colorless, and except for being viscous at high concentrations, is visually indistinguishable from water. Therefore, techniques such as **gel electrophoresis** have been developed to detect and analyze DNA (Figure 3.5.2.11).

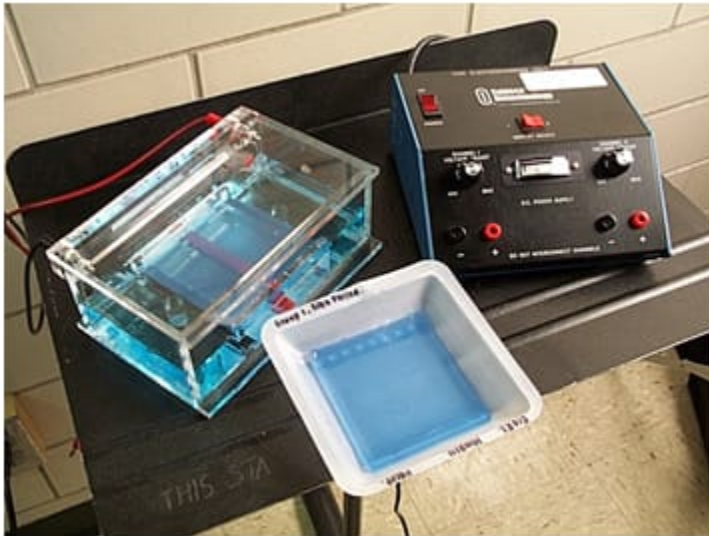


Figure 3.5.2.11: Apparatus for agarose gel electrophoresis. A waterproof tank is used to pass current through a slab gel, which is submerged in a buffer in the tank. The current is supplied by an adjustable power supply. A gel (stained blue by a dye sometimes used when loading DNA on the gel) sits in a tray, awaiting further analysis, such as photography under a UV light source. (Flickr-457088634_585df11af5_o-pending)

This analysis starts when a solution of DNA is deposited at one end of a gel slab. This gel is made from polymers such as **agarose**, which is a polysaccharide isolated from seaweed. The DNA is then forced through the gel by an electrical current, with DNA molecules moving toward the positive electrode (Figure 3.5.2.12).

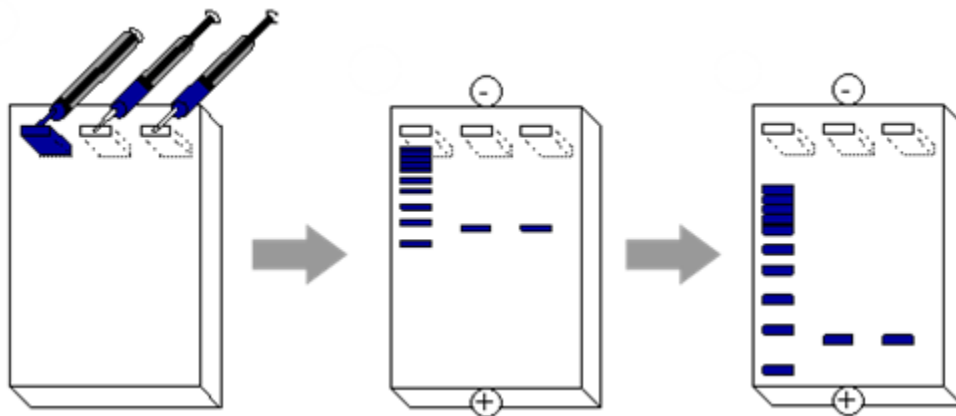


Figure 3.5.2.12: Agarose gel electrophoresis. DNA is loaded into wells at the top of a gel. A current is passed through the gel, pulling DNA towards the positively charged electrode. The DNA fragments are separated by size, with smaller fragments moving fastest towards the electrode. (Wikipedia-Magnus Manske_PD)

As it migrates, each piece of DNA threads its way through the pores, which form between the polymers in the gel. Because shorter pieces can move through these pores faster than longer pieces, gel electrophoresis separates molecules based on their size (length), with smaller DNA pieces moving faster than long ones. DNA molecules of a similar size migrate to a similar location in each gel, called a **band**. This feature makes it easy to see DNA after staining the DNA with a fluorescent dye such as **ethidium bromide** (Figure 3.5.2.13). By separating a mixture of DNA molecules of known size (**size markers**) in adjacent lanes on the same gel, the length of an uncharacterized DNA fragment can be estimated. Gel segments containing the DNA bands can also be cut out of the gel, and the size-selected DNA extracted and used in other types of reactions, such as sequencing and cloning.

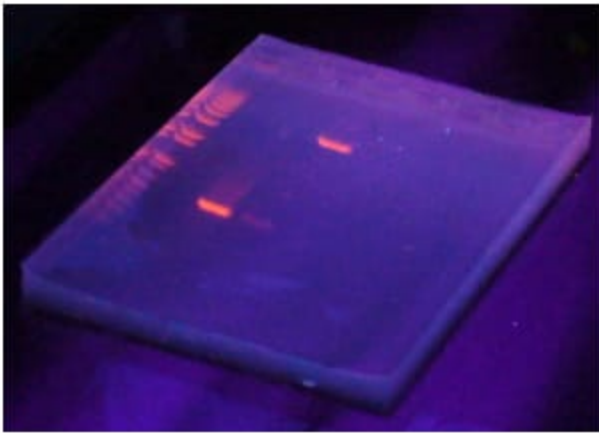


Figure 3.5.2.13: An agarose gel stained with ethidium bromide and illuminated by UV light. The stain associated with DNA is fluorescent. (Wikipedia-Transcontrol-GFDL)

This page titled [3.5.2: DNA Analysis - Gel Electrophoresis](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Todd Nickle and Isabelle Barrette-Ng](#).

- [MO8. Symmetry & Mixing](#) by [Chris Schaller](#) is licensed [CC BY-NC 3.0](#).

3.5.3: Exercises

This page titled [3.5.3: Exercises](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

4: OVERVIEW OF HEMOGLOBIN AND MYOGLOBIN

4.1: Myoglobin, Hemoglobin, and their Ligands

4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin

4.3: Exercises

This page titled [4: Overview of Hemoglobin and Myoglobin](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

4.1: Myoglobin, Hemoglobin, and their Ligands

Almost all biochemistry textbooks start their description of the biological functions of proteins using the myoglobin and hemoglobin as exemplars. These are very rational approaches since they have become model systems to describe the binding of simple ligands, like dioxygen (O_2), CO_2 , and H^+ , and how the structure of the protein determines and is influenced by binding of ligands.

Yet in a way these "ligands" are dissimilar to perhaps the majority of other proteins which bind small ligands such as substrates (for enzymes), inhibitors and activators or large "ligands" such as other proteins, nucleic acids, carbohydrates and lipids. These type of ligands are reversibly bound through classical intermolecular forces (IMFs), such as hydrogen bonds, London dispersion forces, dipole-dipole interactions, and ion-ion interactions. In addition to these are the less commonly discussed pi-pi (aromatic) interactions and cation-pi (aromatic) interactions.

EXTERNAL

YouTube: Less Common Discussed IMFs - pi-pi, cation-pi interactions, and halogen binding

The classic ligands that reversibly bind to hemoglobin, dioxygen, carbon dioxide, and protons, are bound covalently. Dioxygen binds to a heme Fe^{2+} , protons obviously bind to proton acceptors (like His), while CO_2 binds covalently as if forms a carbamate with the N terminus of one of the hemoglobin chains.

This discrepancy in ligand binding mode can be explained easily for dioxygen as it forms a coordinate or dative covalent bond with the transition metal ion Fe^{2+} . In ordinary covalent bonds, each bonded atom contributes to and shares the two electron in the bond. In coordinate or dative covalent bonds, the ligand, a Lewis base, contributes both electrons in the bond. Both electrons are still considered to be "owned" by the ligand and not by the transition metal ion, a Lewis acid. Hence the ligand can readily dissociate from the metal ion, much as a ligand bound through classical IMFs can. This analogy can be extended to protons which are also Lewis acids (with no contributing electrons) as they react with Lewis bases (lone pair donors) on atoms such as N on a His side chain.

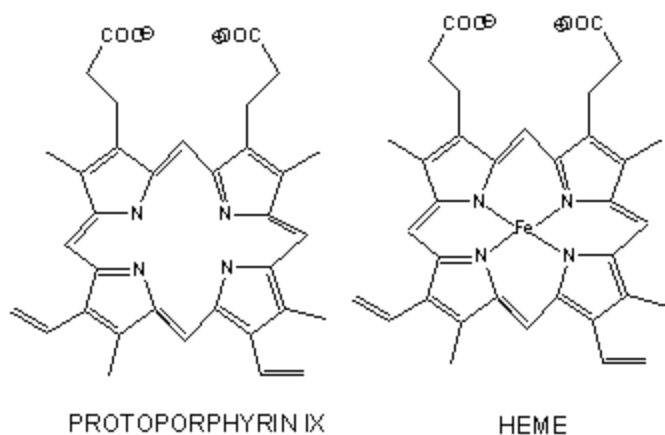
Mb (a monomer containing 8 α -helices, A-H) and Hb (a heterotetramer with two α - and two β -subunits, each which also contains 8 α -helices) are both oxygen binding proteins. Dioxygen is transported from lungs, gills, or skin of an animal to capillaries, where it can be delivered to respiring tissue. It has a low solubility in blood (0.1 mM). Whole blood, which contains 150 g Hb/L, can carry up to 10 mM dioxygen. Invertebrate can have alternative proteins for oxygen binding, including hemocyanin, which contains Cu and hemerythrin, a non-heme protein. On binding dioxygen, solutions of Hb change color to bright red. Solutions of hemocyanin and hemerythrin change to blue and burgundy colored, respectively, on binding dioxygen. Some Antarctic fish don't require Hb since dioxygen is more soluble at low temperature. Mb is found in the muscle, and serves as a storage protein for oxygen transported by Hb. Some information about the proteins is given below:

Hemoglobin:

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

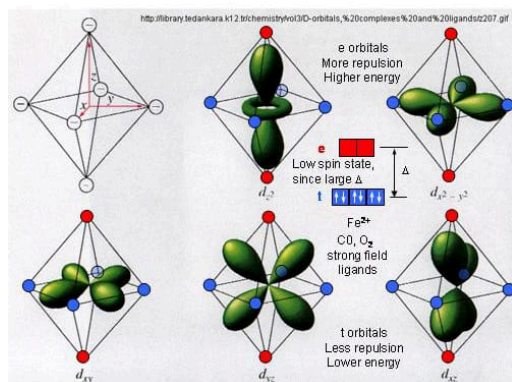
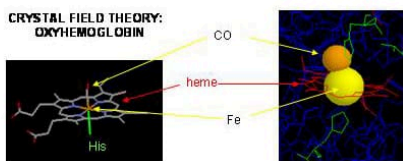
Heme group:

Figure: Heme



- The heme group contains protoporphyrin IX, with four tetrapyrrole rings linked by methene bridges. Attached to the tetrapyrrole structure are four methyl, two vinyl, and two propionate groups. These can be arranged in 15 ways, only one (IX) occurs in biological systems. Ferrous ion bonded to the protoporphyrin IX constitutes the heme group.
- The heme fits into a hydrophobic crevice in the proteins with the propionate groups exposed to solvent.
- The Fe^{2+} ion is coordinated to 4 N's on the 4 pyrrole rings, The 5th ligand is supplied by proximal His (the 8th amino acid on helix F) of the protein. In the absence of dioxygen, the 6th ligand is missing. and the geometry of the complex is square pyramidal with the Fe above the plane of the heme ring. A distal His (E7) is on the opposite side of the heme ring, but too far to coordinate with the Fe. When dioxygen binds, it occupies the 6th coordination site and pulls the Fe into the plane of the ring, leading to octahedral geometry. CO, NO, and H_2S also bind to the 6th site, but with higher affinity than dioxygen, which can lead to CO poisoning. The distal His keeps these ligands (including dioxygen) bound in a bent, non-optimal geometry. This minimizes the chances of CO poisoning.
- When oxygen binds to heme Fe^{2+} , the geometry of the complex becomes octahedral. The Fe^{2+} has 6 electrons in d orbitals. (The electronic configuration of Fe is $3d^6 4s^2$ while the Fe^{2+} ion is $3d^6$). From crystal field theory (review a general chemistry text), the ligand electrons pairs (treated as anions) interact with the d orbital Fe^{2+} electrons. They interact most strongly (in a repulsive fashion) with the d_{z^2} and $d_{x^2-y^2}$ Fe^{2+} electrons, which are oriented at the axial ends and equatorial corners, respectively, of the octahedral complex. This repulsion raises the energy of these two orbitals. The other three d orbitals feel less repulsion by the ligand electrons and are lower in energy. The difference in energy between the three lower energy levels of d electrons (in t orbitals) and the higher energy levels of electrons (in e orbitals) allows photons matching that energy difference to be absorbed as electrons are promoted to the e orbitals. Ligands that interact strongly with the d_{z^2} and $d_{x^2-y^2}$ electrons raise their energy even more, increasing the energy of the photon required to promote an electron from the t to e orbitals. Dioxygen is one of those ligands, leading to the absorbance of high energy photons, causing the solution to appear bright red (since blue light is absorbed). Because of this strong ligand, the significant energy gap between the t and e orbitals causes the electrons to reside mostly in the lower energy t orbitals where they remained paired, giving the complex a low spin state in which all 6 d electrons are paired. Hence oxyhemoglobin is diamagnetic. The heme complex can also exist in a high spin state in which the electrons are easily excited to the e orbitals which are closer in energy to the t orbitals when the ligand is weakly interacting with the Fe^{2+} electrons. In these cases, the heme is paramagnetic since unpaired electrons are possible. (We will discuss paramagnetism later in the course during our exploration of oxygen chemistry).

Figure: Heme- O_2 Octahedral Complex



- When dioxygen binds, the oxidation state of the Fe^{2+} ion does not change, even though dioxygen is a great oxidizing agent. Hence the Fe^{2+} ion is a reversible carrier of dioxygen not of electrons. Free heme in solution is oxidized by dioxygen, forming a complex with water which occupies the 6th position, with the iron in the Fe^{3+} state. An intermediate in this process is the formation of a dimer of 2 hemes linked by 1 dioxygen. This can't occur readily when the heme is in Hb or Mb. Other heme proteins (like Cytochrome C) are designed to be carriers of electrons.
- A small amount of the Fe^{2+} ion gets oxidized to Fe^{3+} ion. Mb and Hb in this state are called metHb and metMb. An enzyme is required to reduce the iron back to the Fe^{2+} state.



Jmol: Updated deoxy-heme and carbon monoxide-heme Jmol14 (Java) | JSMol (HTML5)

Myoglobin

- Mb is extremely compact, and consists of 75% alpha helical structure.
- 8 α -helices are present, labeled A-H.
- 4 alpha helices are terminated by Pro
- The interior amino acids are almost entirely nonpolar. The only polar amino acids found completely buried are the two His (proximal and distal) found at the active site of dioxygen binding.



Jmol: Updated Met-Myoglobin Jmol14 (Java) | JSMol (HTML5)

Difference between Hb and Mb

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

The α and β chains are similar to Mb, which is unexpected since only 24 of 141 residues in the α and β chains of Hb are identical to amino acids in Mb. This suggests that different sequences can fold to similar structures. The globin fold of Mb and each chain of Hb is common to vertebrates and must be nature's design for dioxygen carriers. A comparison of the sequence of Hb from 60 species show much variability of amino acids, with only 9 identical amino acids found. These must be important for structure/function. All internal changes are conservative (e.g. changing a nonpolar for a nonpolar amino acid).

Not even Pro's are conserved, suggesting there are different ways to break helices. The two active site His are conserved, as is Gly B6 (required for a reverse turn). <http://www.umass.edu/molvis/tutorials/hemoglobin/>



Jmol: Hemoglobin

Contributors and Attributions

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

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

4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin

Oxygen Transport

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

Three different chemical solutions to the problem of oxygen transport have developed independently in the course of evolution, as indicated in Table 26.8.13. Mammals, birds, reptiles, fish, and some insects use a heme protein called hemoglobin to transport oxygen from the lungs to the cells, and they use a related protein called myoglobin to temporarily store oxygen in the tissues. Several classes of invertebrates, including marine worms, use an iron-containing protein called hemerythrin to transport oxygen, whereas other classes of invertebrates (arthropods and mollusks) use a copper-containing protein called hemocyanin. Despite the presence of the hem- prefix, hemerythrin and hemocyanin do not contain a metal-porphyrin complex.

Table 26.8.13 Some Properties of the Three Classes of Oxygen-Transport Proteins

Protein	Source	M per Subunit	M per O_2 Bound	Color (deoxy form)	Color (oxy form)
hemoglobin	mammals, birds, fish, reptiles, some insects	1 Fe	1 Fe	red-purple	red
hemerythrin	marine worms	2 Fe	2 Fe	colorless	red
hemocyanin	mollusks, crustaceans, spiders	2 Cu	2 Cu	colorless	blue

Myoglobin and Hemoglobin

Myoglobin is a relatively small protein that contains 150 amino acids. The functional unit of myoglobin is an iron-porphyrin complex that is embedded in the protein (Figure 26.8.1). In myoglobin, the heme iron is five-coordinate, with only a single histidine imidazole ligand from the protein (called the proximal histidine because it is near the iron) in addition to the four nitrogen atoms of the porphyrin. A second histidine imidazole (the distal histidine because it is more distant from the iron) is located on the other side of the heme group, too far from the iron to be bonded to it. Consequently, the iron atom has a vacant coordination site, which is where O_2 binds.

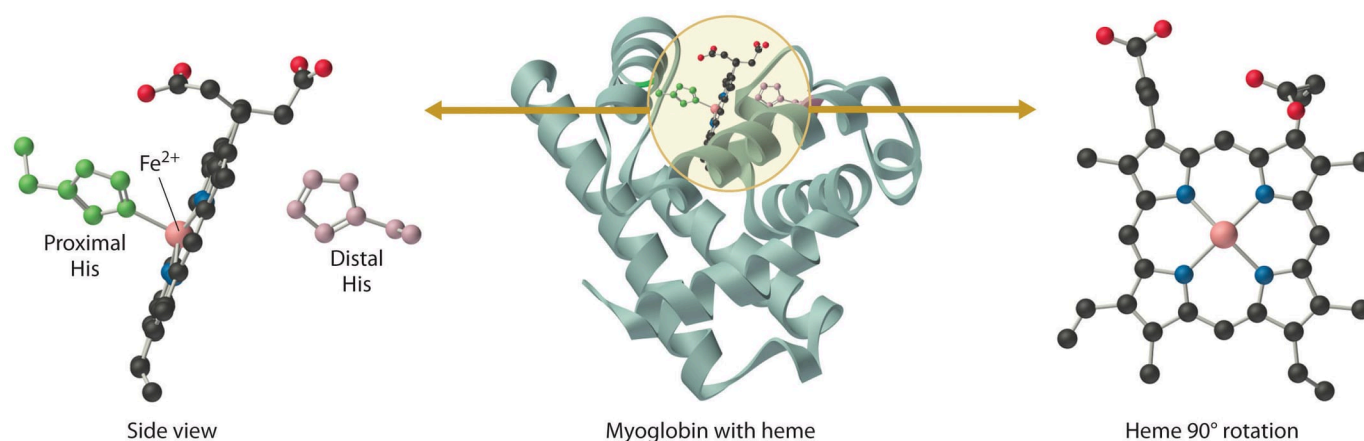


Figure 26.8.1: The Structure of Deoxymyoglobin, Showing the Heme Group. The iron in deoxymyoglobin is five-coordinate, with one histidine imidazole ligand from the protein. Oxygen binds at the vacant site on iron.

In the ferrous form (deoxymyoglobin), the iron is five-coordinate and high spin. Because high-spin Fe^{2+} is too large to fit into the “hole” in the center of the porphyrin, it is about 60 pm above the plane of the porphyrin. When O_2 binds to deoxymyoglobin to form oxymyoglobin, the iron is converted from five-coordinate (high spin) to six-coordinate (low spin; Figure 26.8.2). Because low-spin Fe^{2+} and Fe^{3+} are smaller than high-spin Fe^{2+} , the iron atom moves into the plane of the porphyrin ring to form an octahedral complex. The O_2 pressure at which half of the molecules in a solution of myoglobin are bound to O_2 ($P_{1/2}$) is about 1 mm Hg (1.3×10^{-3} atm).

affinity unchanged, which is important because carbon monoxide is produced continuously in the body by degradation of the porphyrin ligand (even in nonsmokers). Under normal conditions, CO occupies approximately 1% of the heme sites in hemoglobin and myoglobin. If the affinity of hemoglobin and myoglobin for CO were 100 times greater (due to the absence of the distal histidine), essentially 100% of the heme sites would be occupied by CO, and no oxygen could be transported to the tissues. Severe carbon-monoxide poisoning, which is frequently fatal, has exactly the same effect. Thus the primary function of the distal histidine appears to be to decrease the CO affinity of hemoglobin and myoglobin to avoid self-poisoning by CO.

4.3: Exercises

This page titled [4.3: Exercises](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

5: MICHAELIS-MENTEN ENZYME KINETICS

5.1: Catalytic Efficiency of Enzymes

5.2: Michaelis-Menten Kinetics

5.3: Enzyme Parameters

5.4: Enzyme Inhibition

5.4.1: Enzyme Inhibition

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

5.6: Multi-Substrate Sequential Mechanisms

5.7: Double displacement reaction

5.1: Catalytic Efficiency of Enzymes

Enzymes exist in all biological systems in abundant numbers, but not all of their functions are fully understood. Enzymes are important for a variety of reasons, most significantly because they are involved in many vital biochemical reactions. Increasing the reaction rate of a chemical reaction allows the reaction to become more efficient, and hence more products are generated at a faster rate. These products then become involved in some other biological pathway that initiates certain functions of the human body. This is known as the catalytic efficiency of enzymes, which, by increasing the rates, results in a more efficient chemical reaction within a biological system.

Introduction

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

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

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

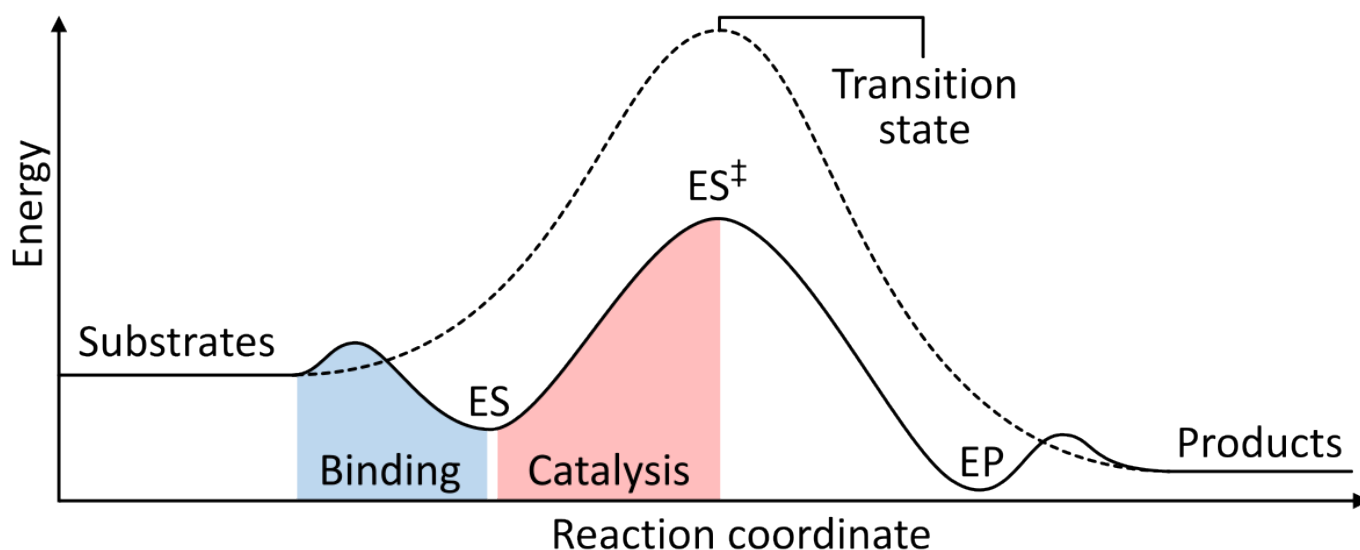


Figure 1: The energies of the stages of a chemical reaction. Uncatalyzed (dashed line), substrates need a lot of activation energy to reach a transition state, which then decays into lower-energy products. When enzyme catalyzed (solid line), the enzyme binds the substrates (ES), then stabilizes the transition state ($ES^\ddagger$) to reduce the activation energy required to produce products (EP) which are finally released. from Wikipedia.

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

reaction via many mechanisms, with one of the more prominent explanations being that it adopts a more favorable conformation.

The efficiency of the enzyme can be determined as follows: consider a simple enzymatic reaction:

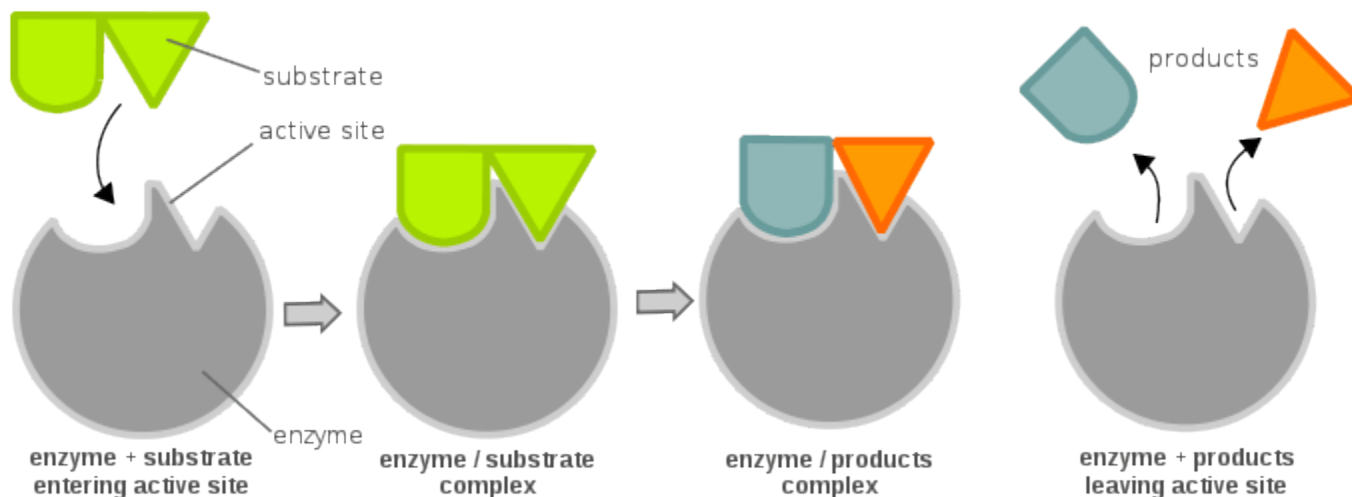


Figure 5.1.2: An enzyme catalyzes the reaction of two substrates and to form one product. (CC BY-SA 4.0; Wikipedia).

German biochemist Leonor Michaelis and Canadian biochemist Maud Menten derived an equation describing this system, later known as the "Michaelis-Menten Equation", shown below:

$$v_0 = \frac{V_{max}[S]}{K_M + [S]} \quad (1)$$

This equation gives the rate of the reaction at a given substrate concentration, assuming a known V_{max} , which is the maximum rate the reaction can proceed at, and K_M , the Michaelis constant. However, in a practical application of the Michaelis-Menten, V_0 is often measured, and V_{max} is observed as a saturation or plateau in a data plot. Because the substrate concentration is known, K_M is usually the calculated value of interest.

For K_M , assume $V_0 = \frac{V_{max}}{2}$:

$$\begin{aligned} \frac{V_{max}}{2} &= \frac{V_{max}[S]}{K_M + [S]} \\ K_M + [S] \frac{V_{max}}{2} &= V_{max}[S] \\ K_M + [S] &= \frac{V_{max}[S]}{\frac{V_{max}}{2}} \\ K_M + [S] &= 2[S] \\ K_M &= [S] \end{aligned}$$

The Michaelis constant can be thought of as the rate at which the substrate becomes unbound from the enzyme, which can either occur in the events of substrate-enzyme complex becoming the product, or the substrate becomes unbound to the enzyme. K_M can be shown as an equation.

$$K_M = \frac{k_{-1} + k_2}{k_1} \quad (7)$$

Whereas k_{-1} is the rate constant at which the substrate becomes unbound to the enzyme, resulting in the dissociation of the enzyme-substrate complex, k_2 is the rate constant where the substrate-enzyme complex disappears and turns into product, and k_1 is the rate constant for the formation of the the substrate-enzyme complex formation. Therefore, K_M can be viewed as the rate of substrate-enzyme complex disappearance divided by the rate of substrate-enzyme complex formation, which is the level at which half of the substrate is bound to the enzyme. K_M is a useful indicator for the presence of an inhibitor because we can look

for changes in K_M and compare to our control (biological systems that we know have zero inhibitor presence). K_M is a dependent variable, and its value can change due to many reasons, including the pH level of the system, temperature, or any other condition that might affect a chemical reaction. A small K_M indicates that the substrate has a high affinity for the enzyme.

The Michaelis-Menten equation is most useful in measuring enzyme efficiency if v_0 is plotted against $[S]$, as follows:

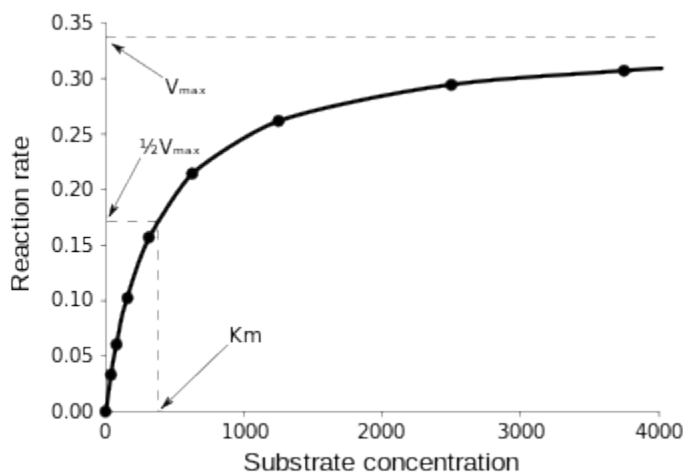


Figure 3: Diagram of reaction speed and Michaelis-Menten constant. from Wikipedia.

V_{max} is the maximum rate at which the reaction can run, regardless of $[S]$, meaning that even if you add more substrate, the reaction cannot go any faster. That is because at V_{max} all of the active sites on the enzyme are occupied. After all the explanations on various forms of enzyme kinetic equations, we arrive at our conclusion of catalytic efficiency. Referring back to Fig 3, we have:

$$V_o = k_2 \left(\frac{[E]_o[S]}{\frac{k_{-1} + k_2}{k_1} + [S]} \right) \quad (8)$$

Notice k_2 describes an irreversible reaction as opposed to an equilibrium expression, when compared to k_{-1} and k_1 . k_2 here is also known as k_{cat} , the catalytic efficiency of enzyme. From the previous discussion, v_0 is the measured reaction rate, which is the product formation over time, so it can be concluded that an equation would look like the following:

$$v_0 = \frac{d[P]}{dt} = k_2[E]_0 \quad (9)$$

where $[E]_0$ is the total enzyme concentration.

It is also known that V_{max} is observed when all of the enzyme-substrate complex disappear and turn into products, so we can make the following assumption:

$$V_{max} = k_2[E]_0 \quad (10)$$

and after rearrangement, we have this equation:

$$k_{cat} = k_2 = \frac{V_{max}}{[E]_0} \quad (11)$$

That is the equation for calculating catalytic efficiency, to be used after we obtain data from experiments and after using the Michaelis-Menten equation. With a larger k_{cat} , the enzyme is efficient because less enzyme is needed.

Exercise 5.1.1

What is the rate of product formation if k_2 is 4.3 min^{-1} and $[ES]$ is $2.3 \times 10^{-2} \text{ M}$

Exercise 5.1.2

Find the formation of ES if $k_1 = 3.3 \times 10^3 \text{ min}^{-1}$, $k_{-1} = 1.1 \times 10^{-10} \text{ min}^{-1}$ and $[E]$ and $[S]$ are $1.2 \times 10^{-6} \text{ M}$ and $6.3 \times 10^{-2} \text{ M}$ respectively

Exercise 5.1.3

Given that $k_{-1} = 8 \times 10^4 \text{ s}^{-1}$, $k_2 = 9 \times 10^5 \text{ s}^{-1}$, $k_1 = 7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ find K_M . What does the answer tell us about the affinity of the substrate

Exercise 5.1.4

Using the K_M value from above, find v_0 if we determine experimentally the $V_{\max}$ value to be $1.5 \times 10^{-4} \text{ M min}^{-1}$ and the $[S]$ is $5.1 \times 10^{-4} \text{ M}$

Exercise 5.1.5

Judging from K_M from problem 1, we hypothesize that $[E]_0$ is half of $[S]$, find k_2 using the $V_{\max}$ value given above

5.2: Michaelis-Menten Kinetics

Two 20th century scientists, Leonor Michaelis and Maud Leonora Menten, proposed the model known as Michaelis-Menten Kinetics to account for enzymatic dynamics. The model serves to explain how an enzyme can cause kinetic rate enhancement of a reaction and explains how reaction rates depends on the concentration of enzyme and substrate.

Introduction

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



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

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

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

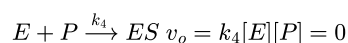
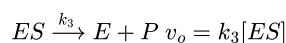
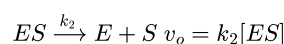
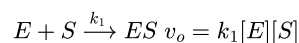


The table below defines each of the rate constants in the above scheme.

Table 1: Model parameters

Rate Constant	Reaction
k_1	The binding of the enzyme to the substrate forming the enzyme substrate complex.
k_2	Catalytic rate; the catalysis reaction producing the final reaction product and regenerating the free enzyme. This is the rate limiting step.
k_3	The dissociation of the enzyme-substrate complex to free enzyme and substrate .
k_4	The reverse reaction of catalysis.

Substrate Complex



The ES complex is formed by combining enzyme E with substrate S at rate constant k_1 . The ES complex can either dissociate to form E_F (free enzyme) and S, or form product P at rate constant k_2 and k_3 , respectively. The velocity equation can be derived in either of the 2 methods that follow.

Method 1: The Rapid Equilibrium Approximation

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

$$v_o = k_2[ES] \quad (5.2.4)$$

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

$$[E]_T = [E_F] + [ES] \quad (5.2.5)$$

$$K_s = \frac{k_2}{k_1} = \frac{[E][S]}{[ES]} \quad (5.2.6)$$

$$K_s = \frac{([E_o] - [ES])[S]}{[ES]} \quad (5.2.7)$$

$$[ES] = \frac{[E_o][S]}{K_s + [S]} \quad (5.2.8)$$

$$v_o = \left(\frac{dP}{dt} \right)_o = k_3[ES] \quad (5.2.9)$$

$$v_o = \left(\frac{dP}{dt} \right)_o = \frac{k_3[E_o][S]}{K_s + [S]} \quad (5.2.10)$$

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

$$v_o = \left(\frac{dP}{dt} \right)_o = k_3[E_o] = V_{max} \quad (1)$$

Method 2: The Steady-State Approximation

Figure 5.2.1 below shows the relatively low and constant concentration of the enzyme-substrate complex due to the complex's slow formation and rapid consumption. Note the falling substrate concentration and the rising product concentration.

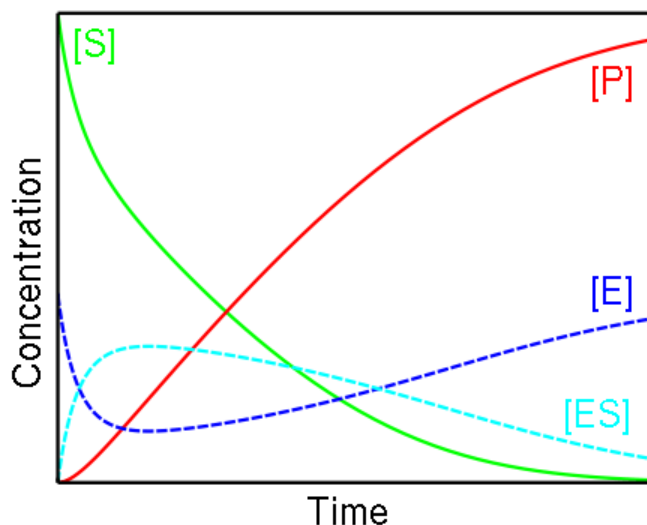


Figure 5.2.1: Change in concentrations over time for enzyme E, substrate S, complex ES and product P. from Wikipedia.

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

- The rate of formation of E-S = $k_1[E][S]$
(with the assumption that $[P] = 0$)
- The rate of breakdown of E-S = $k_2[ES] + k_3[ES] = (K_2 + K_3)[ES]$

At steady state,

$$\frac{d[ES]}{dt} = k_1[E][S] + k_2[ES] + k_3[ES] = 0 \quad (5.2.11)$$

Therefore, rate of formation of E-S is equal to the rate of breakdown of E-S

So,

$$k_1[E][S] = (k_2 + k_3)[ES] \quad (5.2.12)$$

Dividing through by k_1 :

$$[E][S] = \frac{(k_2 + k_3)}{k_1}[E - S] \quad (5.2.13)$$

Substituting $\frac{(k_2 + k_3)}{k_1}$ with K_M :

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

$$K_M = \frac{\text{breakdown}[ES]}{\text{formation}[ES]} \quad (5.2.14)$$

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

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

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

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

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

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

$$[E_T] [S] = [ES] ([S] + K_M)$$

Solving for $[ES]$:

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

The rate equation from the rate limiting step is:

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

Multiplying both sides of the equation by k_2 :

$$k_2[ES] = k_2 \left(\frac{([E_T][S])}{(K_M + [S])} \right) \quad (5.2.15)$$

$$V_o = k_2 \left(\frac{([E_T][S])}{(K_M + [S])} \right) \quad (5.2.16)$$

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

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

$$v_o = \frac{(v_{max}[S])}{(K_M + [S])}$$

Reaction Order Note

When $[S] \ll K_M$,

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

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

When $[S] \gg K_M$,

$$v = V_{max} \quad (5.2.17)$$

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

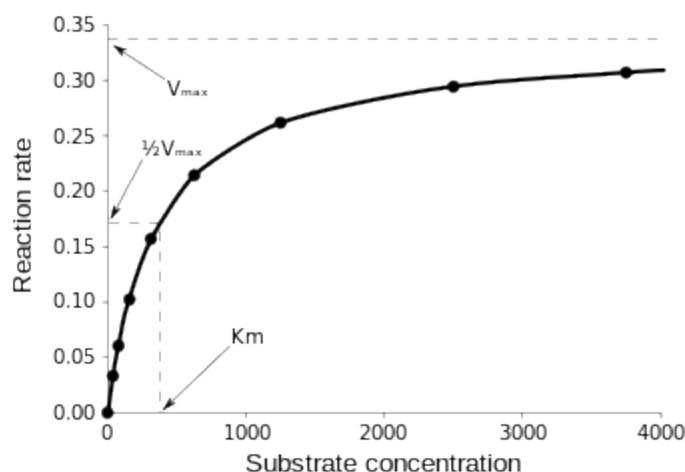


Figure 5.2.2: Diagram of reaction speed and Michaelis-Menten kinetics. from Wikipedia.

Then, at

$$v = \frac{V_{max}}{2}, K_M = [S]$$

$$v = \frac{V_{max}}{2} = \frac{V_{max}[S]}{K_M + [S]} \quad (5.2.18)$$

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

Lineweaver-Burk Plot

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

$$v_o = \frac{(V_{max}[S])}{(K_M + [S])} \quad (5.2.19)$$

$$\frac{1}{v} = \frac{(K_M + [S])}{v_{max}[S]} \quad (5.2.20)$$

$$\frac{1}{v} = \left(\frac{K_M}{V_{max}} \right) \left(\frac{1}{[S]} \right) + \frac{1}{V_{max}} \quad (5.2.21)$$

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

$$y = \frac{1}{v} \quad (5.2.22)$$

$$x = \frac{1}{[S]} \quad (5.2.23)$$

$$m = \text{slope} = \frac{K_M}{V_{max}} \quad (5.2.24)$$

$$b = y - \text{intercept} = \frac{1}{V_{max}} \quad (5.2.25)$$

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

$$x - \text{intercept} = \frac{-1}{K_M} \quad (5.2.26)$$

$$y - \text{intercept} = \frac{1}{V_{max}} \quad (5.2.27)$$

$$\text{slope} = \frac{K_M}{V_{max}} \quad (5.2.28)$$

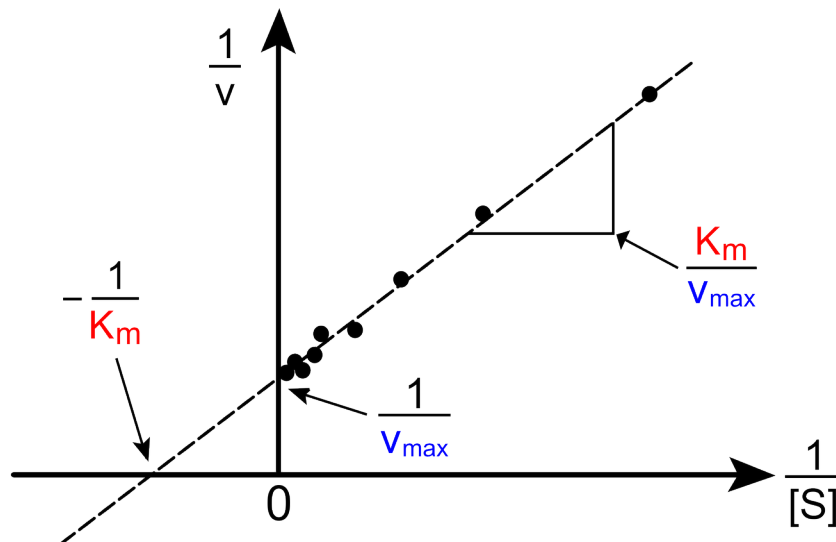


Figure 3. An example of a Lineweaver-Burke plot. from Wikipedia.

Eadie-Hofstee

Another way to calculate these values (k_M , V_{max}) and represent enzyme kinetics:

$$V_o = \frac{(V_{max}[S])}{(K_M + [S])} \quad (5.2.29)$$

$$v_o (k_M + [S]) = v_{max}[S]$$

$$v_o k_M + v_o[S] = v_{max} [S]$$

$$v_o [S] = -v_o k_M = v_{max} [S]$$

Dividing through by $[S]$ $v_o = -k_m \frac{v_o}{[S]} + v_{max}$

Figure 4

References

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

Contributors and Attributions

- Han Le, Sandy Algaze, Eva Tan

This page titled [5.2: Michaelis-Menten Kinetics](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [Michaelis-Menten Kinetics](#) is licensed [CC BY-NC-SA 4.0](#).

5.3: Enzyme Parameters

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

To perform this analysis, I would perform the following experiment. Into 20 different tubes, I would put enzyme buffer (to keep the enzyme stable), the same amount of enzyme, and then a different amount of substrate in each tube, ranging from tiny amounts in the first tubes to very large amounts in the last tubes. I would let the reaction proceed for a fixed, short amount of time and then I would measure the amount of product contained in each tube. For each reaction, I would determine the velocity of the reaction as the concentration of product found in each tube divided by the time. I would then plot the data on a graph using velocity on the Y-axis and the concentration of substrate on the X-axis.

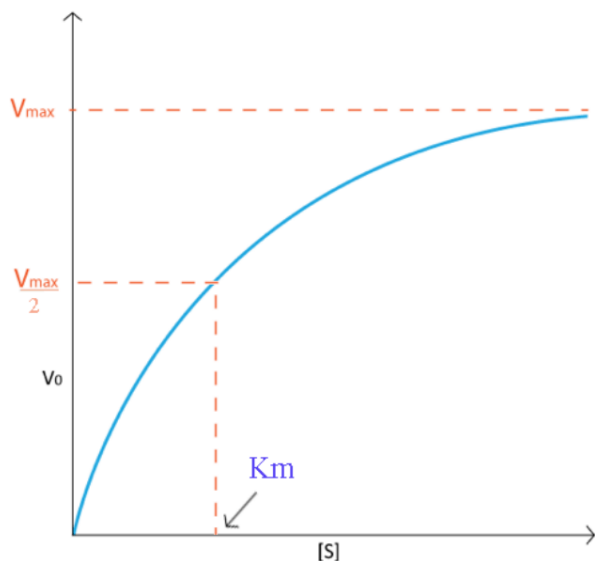


Figure 5.3.1: Velocity vs Substrate concentration plot

Typically, I would generate a curve like that shown on Figure 5.3.1. Notice how the velocity increase is almost linear in the tubes with the lowest amounts of substrate. This indicates that substrate is limiting and the enzyme converts it into product as soon as it can bind it. As the substrate concentration increases, however, the velocity of the reaction in tubes with higher substrate concentration ceases to increase linearly and instead begins to flatten out, indicating that as the substrate concentration gets higher and higher, the enzyme has a harder time keeping up to convert the substrate to product. What is happening is the enzyme is becoming saturated with substrate at higher concentrations of the latter. Not surprisingly, when the enzyme becomes completely saturated with substrate, it will not have to wait for substrate to diffuse to it and will therefore be operating at maximum velocity.

Turnover Number

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

$$K_{\text{cat}} = \frac{V_{\max}}{[\text{Enzyme}]} \quad (4.7.1)$$

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

Enzyme	Turnover Number (per second)
Carbonic anhydrase	600,000
3-Ketoesteroid isomerase	280,000
Acetylcholinesterase	25,000
Penicillinase	2,000
Lactate dehydrogenase	1,000
Chymotrypsin	100
DNA Polymerase I	15
Tryptophan synthetase	2
Lysozyme	0.5

Figure 5.3.2: Common turnover numbers

Michaelis constant

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

5.4: Enzyme Inhibition

Inhibition of specific enzymes by drugs can be medically useful. Understanding the mechanisms of enzyme inhibition is therefore of considerable importance. We will discuss four types of enzyme inhibition – competitive, non- competitive, uncompetitive, and suicide. Of these, the first three types are reversible. The last one 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. When the drug methotrexate is present, some of the enzyme binds to it instead of to folate and during the time methotrexate is bound, the enzyme is inactive and unable to bind folate. Thus, the enzyme is inhibited. Notably, the binding site on DHFR for methotrexate is the active site, the same place that folate would normally bind. As a result, methotrexate ‘competes’ with folate for binding to the enzyme. The more methotrexate there is, the more effectively it competes with folate for the enzyme’s active site. Conversely, the more folate there is, the less of an effect methotrexate has on the enzyme because folate outcompetes it.

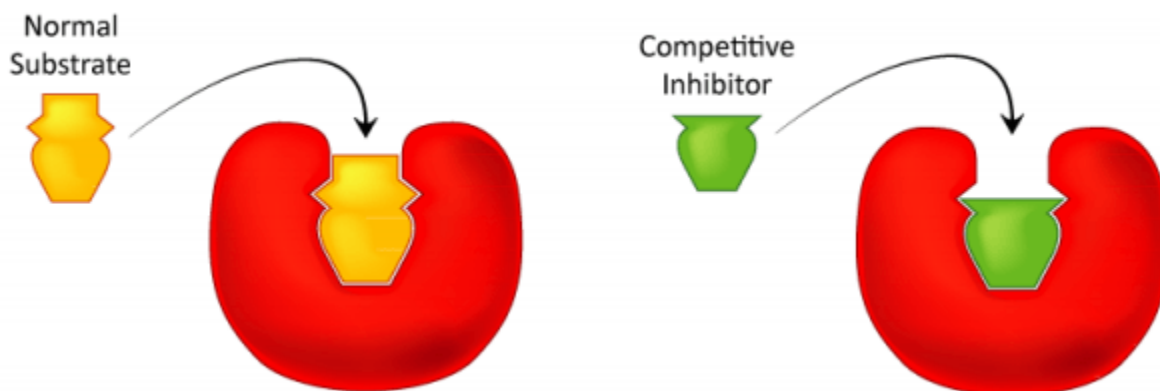


Figure 4.10.1:

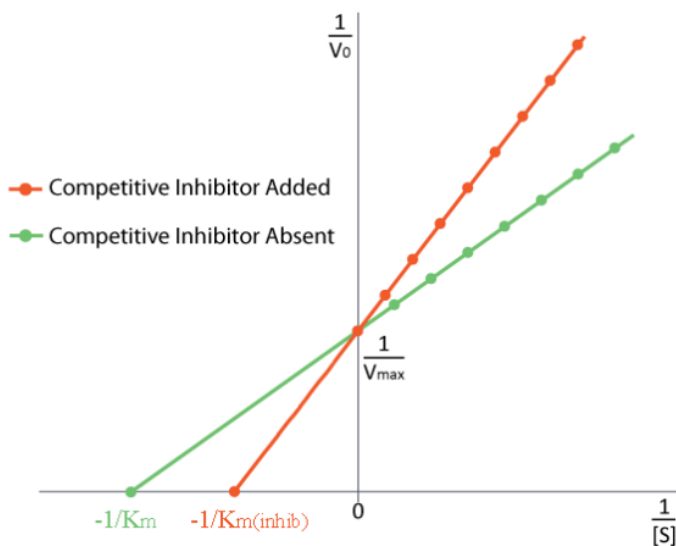


Figure 4.10.2: Line-Weaver Burk Plot

Competitive Inhibition
of competitive inhibition

No Effect On V_{MAX}

How do we study competitive inhibition. It is typically done as follows. First one performs a set of V vs. $[S]$ reactions without inhibitor (20 or so tubes, with buffer and constant amounts of enzyme, varying amounts of substrate, equal reaction times). V vs. $[S]$ is plotted, as well as $1/V$ vs. $1/[S]$, if desired. Next, a second set of reactions is performed in the same manner as before, except that a fixed amount of the methotrexate inhibitor is added to each tube. At low concentrations of substrate, the inhibitor competes for the enzyme effectively, but at high concentrations of substrate, the inhibitor will have a much reduced effect, since the substrate outcompetes it, due to its higher concentration (remember that the inhibitor is at fixed concentration). Graphically, the results of these experiments are shown above. Notice that at high substrate concentrations, the competitive inhibitor has essentially no effect, causing the 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.

Increased K_M

Note that the apparent K_M of the enzyme for the substrate increases ($-1/K_M$ gets closer to zero - red line above) when the inhibitor is present, thus illustrating the better competition of the inhibitor at lower substrate concentrations. It may not be obvious why we call the changed K_M the apparent K_M 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 K_M 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 K_M values are unchanged. Why then, does K_M appear higher in the presence of a competitive inhibitor. The reason is that the competitive inhibitor is reducing the amount of active enzyme at lower concentrations of substrate. When the amount of enzyme is reduced, one must have more substrate to supply the reduced amount of enzyme sufficiently to get to $V_{max}/2$.

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 (rectangular site below). The effect of binding a non-competitive inhibitor is significantly different from binding a competitive inhibitor because there is no competition. In the case of competitive inhibition, the effect of the inhibitor could be reduced and eventually overwhelmed with increasing amounts of substrate. This was because increasing substrate made increasing percentages of the enzyme active. With non-competitive inhibition, increasing the amount of substrate has no effect on the percentage of enzyme that is active. Indeed, in non-competitive inhibition, the percentage of enzyme inhibited remains the same through all ranges of $[S]$.

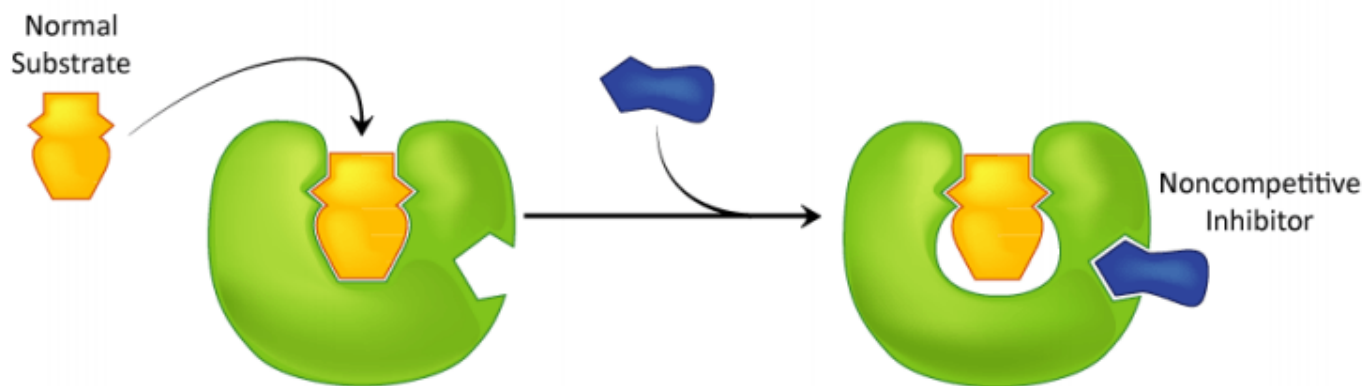


Figure 4.10.3: *Non-competitive Inhibition*

This means, then, that non-competitive inhibition effectively reduces the amount of enzyme by the same fixed amount in a typical experiment at every substrate concentration used. The effect of this inhibition is shown above. As you can see, V_{max} is reduced in non-competitive inhibition compared to uninhibited reactions. This makes sense if we remember that V_{max} 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, K_M 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.

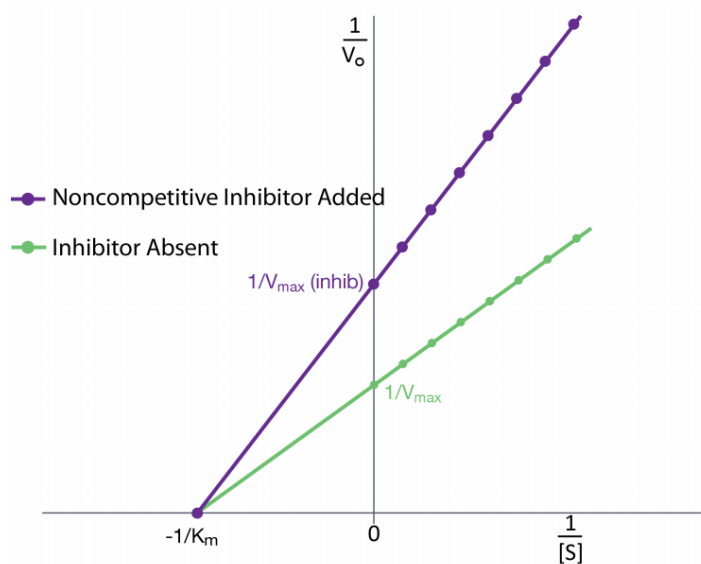


Figure 4.10.4: *Line-Weaver Burk Plot of noncompetitive*

inhibition

Uncompetitive Inhibition

A third type of enzymatic inhibition is that of uncompetitive inhibition, which has the odd property of a reduced V_{max} as well as a reduced K_M . The explanation for these seemingly odd results is rooted in the fact that the uncompetitive inhibitor binds only to the enzyme-substrate (ES) complex. The inhibitor-bound complex forms mostly under concentrations of high substrate and the ES-I complex cannot release product while the inhibitor is bound, thus explaining the reduced V_{max} .

The reduced K_M is a bit harder to conceptualize. The answer lies in the fact that the inhibitor-bound complex effectively reduces the concentration of the ES complex. By Le Chatelier's Principle, a shift occurs to form additional ES complex, resulting in less free enzyme and more enzyme in the forms ES and ESI (ES with inhibitor). Decreases in free enzyme correspond to an enzyme with greater affinity for its substrate. Thus, paradoxically, uncompetitive inhibition both decreases V_{max} and increases an enzyme's affinity for its substrate.

Suicide Inhibition

In contrast to the first three types of inhibition, which involve reversible binding of the inhibitor to the enzyme, suicide inhibition is irreversible because the inhibitor becomes covalently bound to the enzyme during the inhibition and thus cannot be removed. Suicide inhibition rather closely resembles competitive inhibition because the inhibitor generally resembles the substrate and binds to the active site of the enzyme. The primary difference is that the suicide inhibitor is chemically reactive in the active site and makes a bond with it that precludes its removal. Such a mechanism is that employed by penicillin (Figure 4.10.5), which covalently links to the bacterial enzyme, D-D transpeptidase and stops it from functioning. Since the normal function of the enzyme is to make a bond necessary for the peptido-glycan complex of the bacterial cell wall, the cell wall cannot properly form and bacteria cannot reproduce. If one were to measure the kinetics of suicide inhibitors under conditions where there was more enzyme than inhibitor, they would resemble non-competitive inhibition's kinetics because both involve reducing the amount of active enzyme by a fixed amount in a set of reactions.

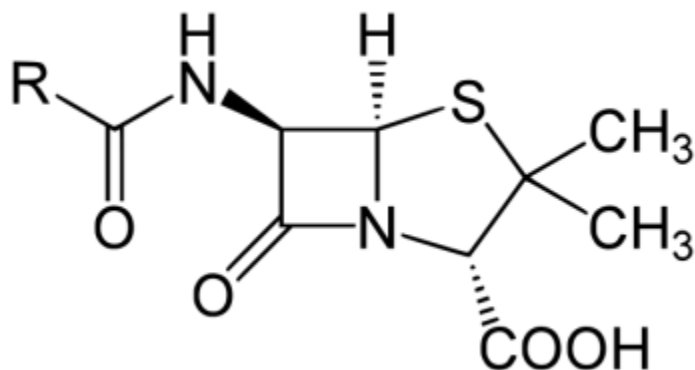


Figure 4.10.5: *Penicillin*

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

- [Solubility Demonstration II -- Extraction](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

5.4.1: Enzyme Inhibition

Inhibition of specific enzymes by drugs can be medically useful. Understanding the mechanisms of enzyme inhibition is therefore of considerable importance. We will discuss four types of enzyme inhibition – competitive, non- competitive, uncompetitive, and suicide. Of these, the first three types are reversible. The last one 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. When the drug methotrexate is present, some of the enzyme binds to it instead of to folate and during the time methotrexate is bound, the enzyme is inactive and unable to bind folate. Thus, the enzyme is inhibited. Notably, the binding site on DHFR for methotrexate is the active site, the same place that folate would normally bind. As a result, methotrexate ‘competes’ with folate for binding to the enzyme. The more methotrexate there is, the more effectively it competes with folate for the enzyme’s active site. Conversely, the more folate there is, the less of an effect methotrexate has on the enzyme because folate outcompetes it.

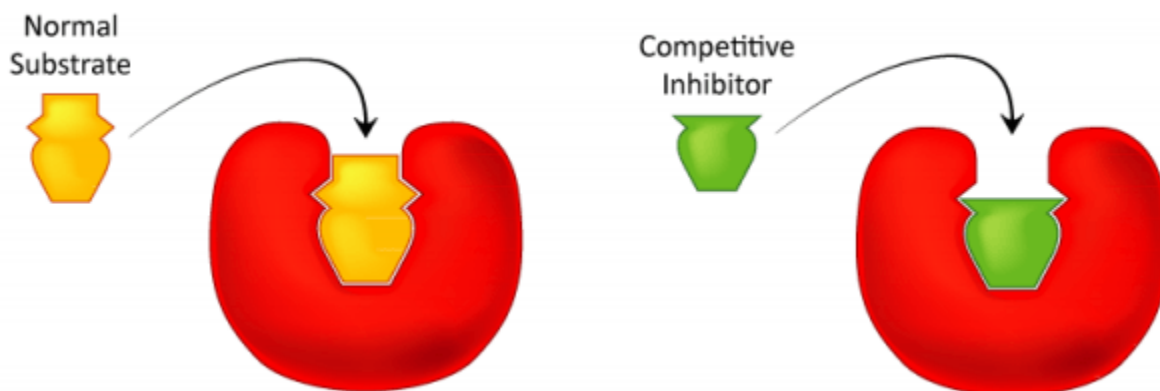
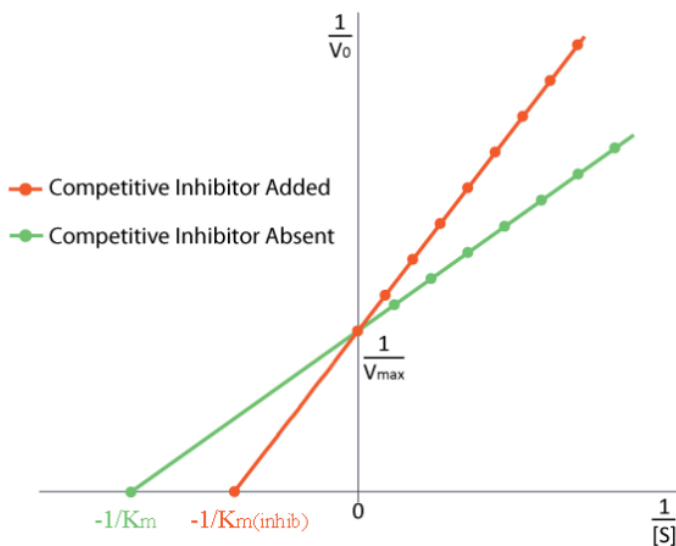


Figure 4.10.1:



Competitive Inhibition
of competitive inhibition

Figure 4.10.2: Line-Weaver Burk Plot

No Effect On V_{MAX}

How do we study competitive inhibition. It is typically done as follows. First one performs a set of V vs. $[S]$ reactions without inhibitor (20 or so tubes, with buffer and constant amounts of enzyme, varying amounts of substrate, equal reaction times). V vs. $[S]$ is plotted, as well as $1/V$ vs. $1/[S]$, if desired. Next, a second set of reactions is performed in the same manner as before, except that a fixed amount of the methotrexate inhibitor is added to each tube. At low concentrations of substrate, the inhibitor competes for the enzyme effectively, but at high concentrations of substrate, the inhibitor will have a much reduced effect, since the substrate outcompetes it, due to its higher concentration (remember that the inhibitor is at fixed concentration). Graphically, the results of these experiments are shown above. Notice that at high substrate concentrations, the competitive inhibitor has essentially no effect, causing the 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.

Increased K_M

Note that the apparent K_M of the enzyme for the substrate increases ($-1/K_M$ gets closer to zero - red line above) when the inhibitor is present, thus illustrating the better competition of the inhibitor at lower substrate concentrations. It may not be obvious why we call the changed K_M the apparent K_M 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 K_M 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 K_M values are unchanged. Why then, does K_M appear higher in the presence of a competitive inhibitor. The reason is that the competitive inhibitor is reducing the amount of active enzyme at lower concentrations of substrate. When the amount of enzyme is reduced, one must have more substrate to supply the reduced amount of enzyme sufficiently to get to $V_{max}/2$.

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 (rectangular site below). The effect of binding a non-competitive inhibitor is significantly different from binding a competitive inhibitor because there is no competition. In the case of competitive inhibition, the effect of the inhibitor could be reduced and eventually overwhelmed with increasing amounts of substrate. This was because increasing substrate made increasing percentages of the enzyme active. With non-competitive inhibition, increasing the amount of substrate has no effect on the percentage of enzyme that is active. Indeed, in non-competitive inhibition, the percentage of enzyme inhibited remains the same through all ranges of $[S]$.

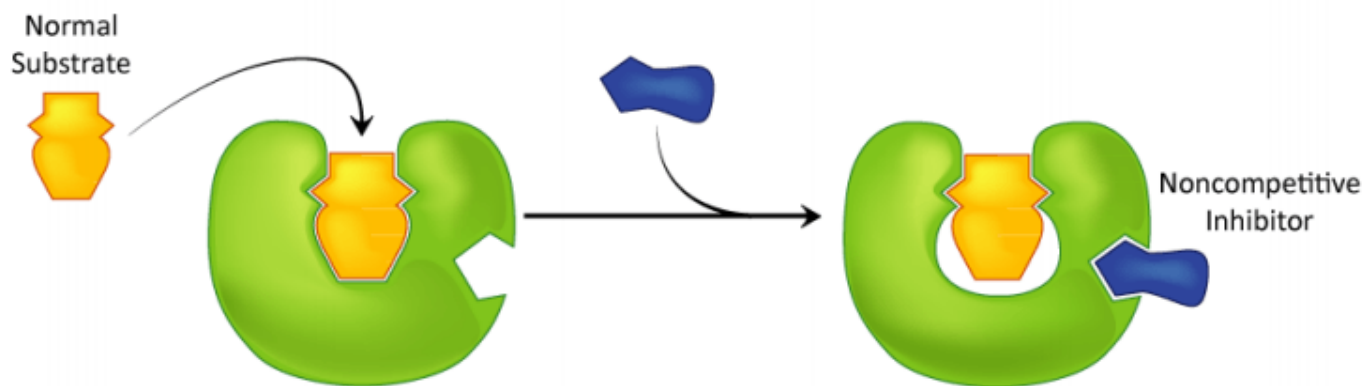


Figure 4.10.3: *Non-competitive Inhibition*

This means, then, that non-competitive inhibition effectively reduces the amount of enzyme by the same fixed amount in a typical experiment at every substrate concentration used. The effect of this inhibition is shown above. As you can see, V_{max} is reduced in non-competitive inhibition compared to uninhibited reactions. This makes sense if we remember that V_{max} 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, K_M 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.

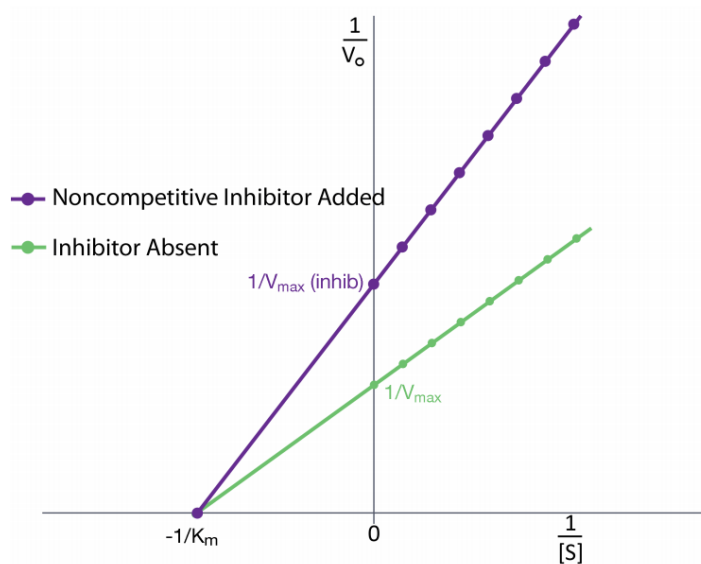


Figure 4.10.4: *Line-Weaver Burk Plot of noncompetitive*

inhibition

Uncompetitive Inhibition

A third type of enzymatic inhibition is that of uncompetitive inhibition, which has the odd property of a reduced V_{max} as well as a reduced K_M . The explanation for these seemingly odd results is rooted in the fact that the uncompetitive inhibitor binds only to the enzyme-substrate (ES) complex. The inhibitor-bound complex forms mostly under concentrations of high substrate and the ES-I complex cannot release product while the inhibitor is bound, thus explaining the reduced V_{max} .

The reduced K_M is a bit harder to conceptualize. The answer lies in the fact that the inhibitor-bound complex effectively reduces the concentration of the ES complex. By Le Chatelier's Principle, a shift occurs to form additional ES complex, resulting in less free enzyme and more enzyme in the forms ES and ESI (ES with inhibitor). Decreases in free enzyme correspond to an enzyme with greater affinity for its substrate. Thus, paradoxically, uncompetitive inhibition both decreases V_{max} and increases an enzyme's affinity for its substrate.

Suicide Inhibition

In contrast to the first three types of inhibition, which involve reversible binding of the inhibitor to the enzyme, suicide inhibition is irreversible because the inhibitor becomes covalently bound to the enzyme during the inhibition and thus cannot be removed. Suicide inhibition rather closely resembles competitive inhibition because the inhibitor generally resembles the substrate and binds to the active site of the enzyme. The primary difference is that the suicide inhibitor is chemically reactive in the active site and makes a bond with it that precludes its removal. Such a mechanism is that employed by penicillin (Figure 4.10.5), which covalently links to the bacterial enzyme, D-D transpeptidase and stops it from functioning. Since the normal function of the enzyme is to make a bond necessary for the peptido-glycan complex of the bacterial cell wall, the cell wall cannot properly form and bacteria cannot reproduce. If one were to measure the kinetics of suicide inhibitors under conditions where there was more enzyme than inhibitor, they would resemble non-competitive inhibition's kinetics because both involve reducing the amount of active enzyme by a fixed amount in a set of reactions.

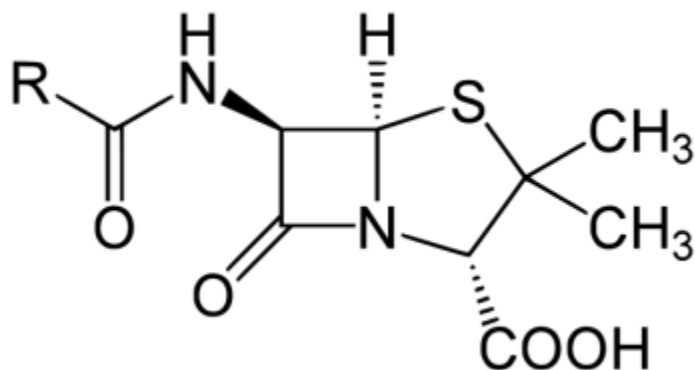


Figure 4.10.5: *Penicillin*

This page titled [5.4.1: Enzyme Inhibition](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Kevin Ahern](#) and [Indira Rajagopal](#).

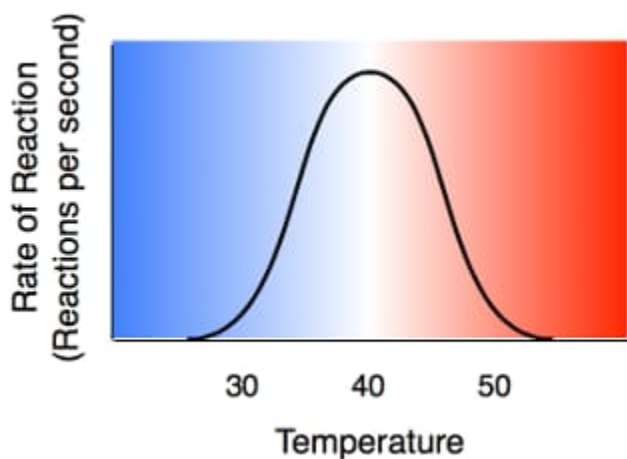
- [Solubility Demonstration II -- Extraction](#) by George Bodner is licensed [CC BY-NC-SA 4.0](#).

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

This laboratory exercise will explore the effect of **temperature**, **pH**, and **enzyme concentration** on the rate of a reaction.

Part A- Temperature

Higher temperature generally causes more collisions among the molecules and therefore increases the rate of a reaction. More collisions increase the likelihood that substrate will collide with the active site of the enzyme, thus increasing the rate of an enzyme-catalyzed reaction. Above a certain temperature, activity begins to decline because the enzyme begins to **denature**. The rate of chemical reactions therefore increases with temperature but then decreases as enzymes denature. This part of the exercise will explore the rate of enzyme activity at two different temperatures. It will also explore the rate of enzyme activity using an enzyme that has been denatured.



Rennin is an enzyme found in the stomach of mammals where it functions to solidify milk. You will observe the activity of this enzyme by mixing it with milk in a test tube. The presence of a reaction is indicated by milk becoming solid.

A1. Predict what will happen in each of the tubes described in the two experiments described below and write your two hypotheses in your notebook. The word "milk" should appear in each of your hypotheses. You should have two hypotheses- one that addresses the effect of temperature on rate of reaction (tubes 1 and 3) and one that addresses the effect of denaturing the enzyme on reaction time (tubes 3 and 5). Your hypotheses should be specific for this experiment; it should state the expected outcome of this experiment.

Be sure to write your hypotheses as statements, not as questions. At this point, you do not know what will happen; your hypotheses may be correct or incorrect. An incorrect hypothesis is perfectly acceptable. You will test the hypotheses by performing the experiment below.

Effect of Temperature

A1. Write a hypothesis for this experiment in your lab notebook. The word "milk" should be used in your hypothesis. For example: "The milk in tube #1 will...."

A2. Obtain five test tubes and put a mark 2 cm from the bottom of each using a wax pencil. Four of these test tubes will be used in this experiment (Effect of Temperature) and two of them will be used in the next experiment (Effect of Denaturing the Enzyme, below).

A3. Mark the tubes with your initials or some other means to identify them. The tubes should also be numbered 1 through 5.

A4. Fill all of the tubes to the 2 cm mark with milk.

A5. Add 3 drops of rennin to the milk in tube #1 and keep it at room temperature for 15 minutes. Tube #2 without rennin should also be kept at room temperature to serve as a control.

A6. Add 3 drops of rennin to the milk in tube #3 and place it in a 37 degree incubator for 15 minutes. Tube #4 without rennin should also be placed in the incubator for 15 minutes to serve as a control.

Effect of Denaturing the Enzyme

A7. Place a mark at 1 cm on a test tube and add rennin up to the mark. Place the tube in a boiling water bath for 2 minutes. After boiling, pour the rennin in a small beaker.

A8. Add 3 drops of boiled rennin to the milk in tube #5 and place it in the 37 degree incubator for 15 minutes along with the tubes prepared in step #A6 above. Tube #4 prepared earlier will serve as a control in this experiment as well as a control in the previous experiment (step A6 above).

The two experiments that you conducted in Part A are summarized below. Tubes 1 through 4 are used to investigate the effect of temperature on enzyme activity. Tubes 4 and 5 are used to investigate the effect of denaturing an enzyme with heat and then using the enzyme at its normal temperature (body temperature).

Effect of Temperature Room Temperature: 37 Degrees:	Tube 1, Tube 2 (control) Tube 3, Tube 4 (control)
Effect of Denaturing the Enzyme:	Tube 5, Tube 6 (control)

A9. Observe the milk in each of the five tubes after the 15 minute period and record your observations in your notebook.

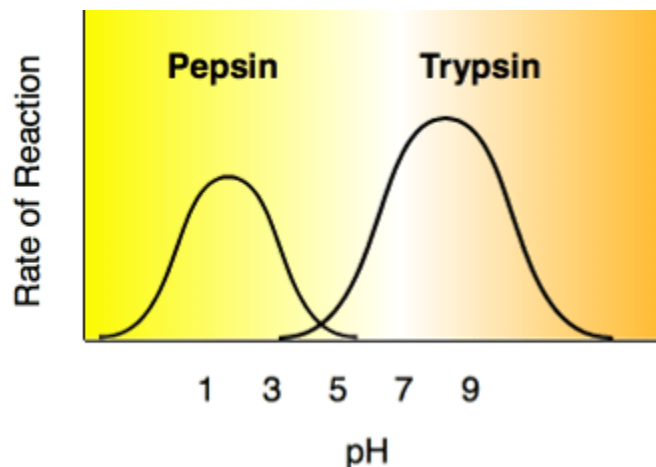
A10. Do your results support either of your hypotheses? Explain.

Part B- pH

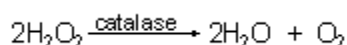
Each enzyme has an optimal pH.

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

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



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



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

In the experiment below, bubbling will be used as an indication that a reaction is occurring. A faster reaction will have more bubbling.

B1. Predict how pH will affect the rate of reaction in the three different situations described below and write your hypothesis or hypotheses in your notebook. At this point, you do not know if your hypothesis is true or false. It is acceptable to create a hypothesis which will be shown to be false.

B2. Mark three test tubes with a wax pencil 3 cm from the bottom and 6 cm from the bottom.

B3. Add 3 cm of hydrogen peroxide to each tube.

B4. Fill one of the tubes to the 6 cm mark with 5 m HCl.

B5. Fill another tube to the 6 cm mark with 5 m NaOH.

B6. Fill the third tube to the 6 cm mark with distilled water.

B7. Cut three strips of potato that are approximately 3 cm long. The strips should be cut thin enough to easily fit into a test tube.

B8. Place one of the strips of potato into each of the three tubes that you prepared in steps 2-6 above.

B9. Write your observation of the amount of bubbling in your notebook. We are only interested in the amount of bubbling. We are not interested in any change in color or whether the potato floats or sinks.

B10. Use pH paper to measure the pH of each tube and record your measurements in your notebook.

B11. At what pH did catalase function best in this experiment?

B12. Do your results support your hypothesis? Explain.

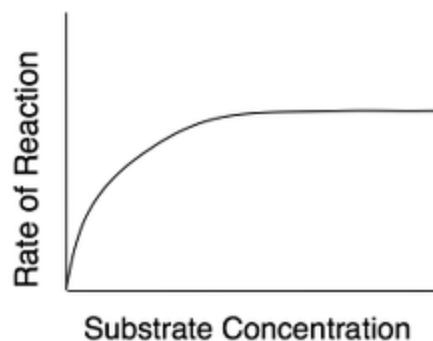
B13. In part A of this lab, you worked with rennin, an enzyme found in the stomach of mammals. The pH of the stomach is normally about 2.0. At what pH do you predict rennin works best?

B14. What can you say about pH and enzyme functioning? Is there a single pH that enzymes function best at or does it depend on the enzyme?

Part C- Substrate Concentration

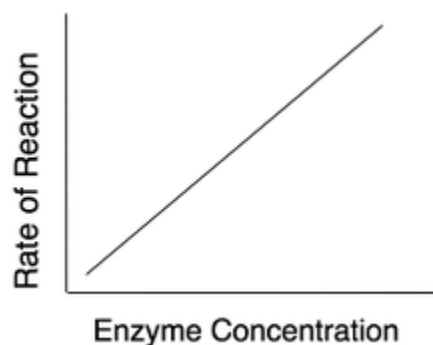
At lower concentrations, the active sites on most of the enzyme molecules are not filled because there is not much substrate. Higher concentrations cause more collisions between the molecules. With more molecules and more collisions, enzymes are more likely to encounter molecules of reactant.

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



Enzyme Concentration

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



In this investigation, we will examine what happens to the rate of a reaction when the amount of enzyme is reduced. We will use urease, an enzyme that converts urea to ammonia. The ammonia causes the pH of the water to increase (it becomes more basic). You will be able to tell when a reaction occurs because the urea solution also contains a pH indicator that is yellow in acid but turns red when the solution becomes basic.

The object of this experiment is to measure the amount of time it takes for the solution to turn red if less enzyme is used.

- C1. Create a hypothesis regarding the amount of urease and the rate of reaction of Urea.
- C2. Obtain four test tubes and rinse them thoroughly using distilled water.
- C3. Add 2 cm of urea to each.
- C4. Label three of these tubes 1 through 3; the remaining tube will not be used; it will serve as a control.
- C5. Have your lab partner start timing as you add 20 drops of urease to tube #1. Swirl the tube, then place it in a test tube rack. Record the amount of time that it takes for the solution to change to a red color.
- C6. Add 10 drops of urease to tube #2, swirl the tube, then place it in a rack. Record the amount of time that it takes for the solution to change to a red color.
- C7. Add 5 drop of urease to tube #3, swirl the tube, swirl the tube, then place it in a rack. Record the amount of time that it takes for the solution to change to a red color.
- C8. Record your results in your notebook.
- C9. Did using less enzyme produce a reaction?

C10. What was the effect of using less enzyme in your experiment? If your experiment did not work as expected, what should have happened?

C11. In general, what happens to the rate of reaction as the amount of enzyme is decreased?

C12. Do your results support your hypothesis? Explain.

Contributors and Attributions

- [Michael J. Gregory, Ph.D. \(Clinton Community College\)](#)

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

- [Homework 10: Nuclear Chemistry](#) is licensed [CC BY-NC-SA 4.0](#).

5.6: Multi-Substrate Sequential Mechanisms

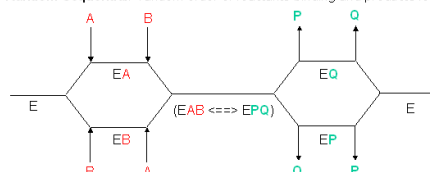
In reality, many enzymes have more than one substrate (A, B) and more than one product (P, Q). For example, the enzyme alcohol dehydrogenase catalyzes the oxidation of ethanol with NAD (a biological oxidizing agent) to form acetaldehyde and NADH. How do you do enzymes kinetics on these more complicated systems? The answer is fairly straightforward. You keep one of the substrates (B, for example) fixed, and vary the other substrate (A) and obtain a series of hyperbolic plots of v_o vs. A at different fixed B concentrations. This would give a series of linear $1/v$ vs $1/A$ double-reciprocal plots (Lineweaver-Burk plots) as well. The pattern of Lineweaver-Burk plots depends on how the reactants and products interact with the enzyme.

Sequential Mechanism

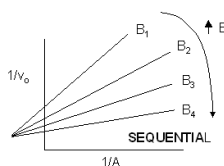
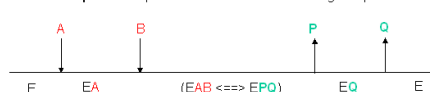
In this mechanism, both substrates must bind to the enzyme before any products are made and released. The substrates might bind to the enzyme in a random fashion (A first then B or vice-versa) or in an ordered fashion (A first followed by B). An abbreviated notation scheme developed by W.W. Cleland is shown below for the sequential random and sequential ordered mechanisms. For both mechanisms, Lineweaver-Burk plots at varying A and different fixed values of B give a series of intersecting lines. Derivative curves can be solved to obtain appropriate kinetic constants.

Sequential: Reactants (A,B) both bind before both products (P,Q) are released

A. **Random Sequential:** random order of reactants binding and products leaving



B. **Ordered Sequential:** specific order of reactants binding and products leaving



Contributors and Attributions

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

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

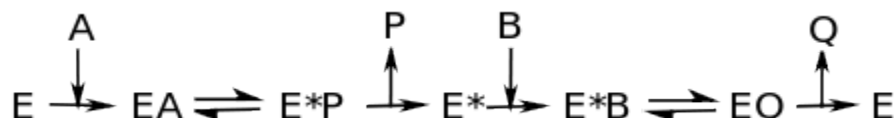
- [Template:ContribSpinney](#) has no license indicated.

5.7: Double displacement reaction

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

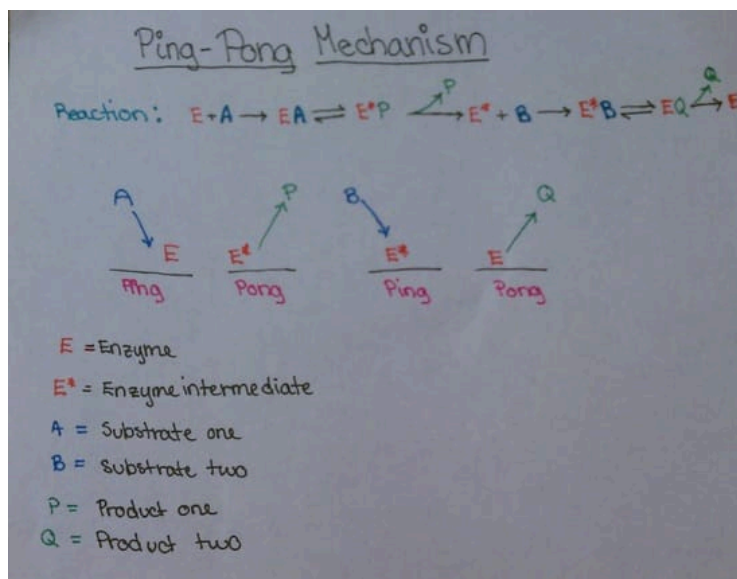
The Mechanism

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



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

Here is another diagram showing this same reaction:



Example 5.7.1: Chymotrypsin

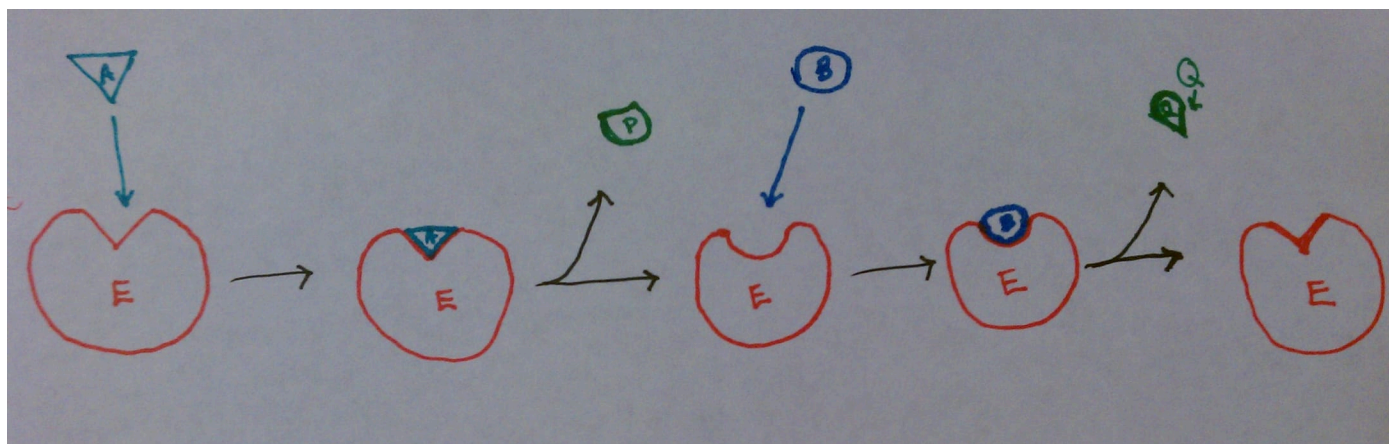
An example of the ping-pong mechanism would be the action of chymotrypsin. When reacted with p-nitrophenyl acetate (A), the reaction of chymotrypsin is seen to occur in two steps. In the first step, the substrate reacts extremely fast with the enzyme, leading to the formation of a small amount of p-nitrophenolate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

Example 5.7.2: Pyruvate Carboxylase

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

Further Reading

An important factor to understand about the ping-pong mechanism is that when plotting a $1/v$ and $1/[A]$ plot at varying concentrations of B, a series of parallel lines are seen. In this case A is the first substrate and B being the second substrate.



Refer to these sections on enzyme kinetics and michaelis-menten kinetics to get a better understanding of what this type of plot means.

Sample Questions

1. The form in which the ping pong mechanism binds substrates is identified as which type of mechanism?
2. What are two characteristics of an enzyme that catalyzes a reaction through the ping-pong mechanism?
3. The following diagram shows the mechanism of glutamate-aspartate aminotransferase: Would this mechanism be considered a ping-pong/double-displacement reaction? Why or why not?

Answers

1. The ping-pong mechanism is a non-sequential mechanism. A product is released after the first substrate is bound.
2. One, a product is seen before the second substrate is bound. Two, binding of the first substrate causes the enzyme to change

into an intermediate form that will bind the second substrate. Three, the plot of $1/v$ vs. $1/[A]$ as $[B]$ changes will be parallel lines.

3. Yes! This would definitely be considered a ping-pong mechanism. First, we can see that there is an E' state which is indicative of a ping-pong mechanism. Pyridoxal is a coenzyme bound to glutamate-aspartate aminotransferase that accepts an amino group from glutamate and becomes pyridoxamine while releasing α -ketoglutarate. Pyridoxamine bound to the enzyme will then donate its amino group to oxaloacetate to regenerate pyridoxal as well as aspartate.

References

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

Contributors and Attributions

- Melissa Hill, Laura Lan, Andrew Jilwan

This page titled [5.7: Double displacement reaction](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [The "Ping-Pong" Mechanism](#) is licensed [CC BY 4.0](#).

6: CATALYTIC STRATEGIES OF ENZYMES

In living systems, speed is everything. Providing the reaction speeds necessary to support life are the catalysts, mostly in the form of enzymes.



-  Topic hierarchy 

6.1: Serine Proteases

6.2: Chymotrypsin

6.3: Trypsin

Contributors

Template:ContribAhern

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

6.1: Serine Proteases

Three protein-digesting enzymes secreted by the pancreas are serine proteases: [chymotrypsin](#), [trypsin](#) and elastase. These three share closely-similar structures (tertiary as well as primary). In fact, their active serine residue is at the same position (**Ser-195**) in all three. Despite their similarities, they have different substrate specificities; that is, they cleave different peptide bonds during protein digestion.

6.2: Chymotrypsin

Consider the mechanism of catalysis of the enzyme known as chymotrypsin. Found in our digestive system, chymotrypsin's catalytic action is cleaving peptide bonds in proteins and it uses the side chain of a serine in its mechanism of catalysis. Many other protein- cutting enzymes employ a very similar mechanism and they are known collectively as serine proteases. As a protease, it acts fairly specifically, cutting not all peptide bonds, but only those that are adjacent to specific amino acids in the protein. One of the amino acids it cuts adjacent to is phenylalanine. The enzyme's action occurs in two phases – a fast phase that occurs first and a slower phase that follows. The enzyme has a substrate binding site that includes a region of the enzyme known as the S1 pocket. Let us step through the mechanism by which chymotrypsin cuts adjacent to phenylalanine.

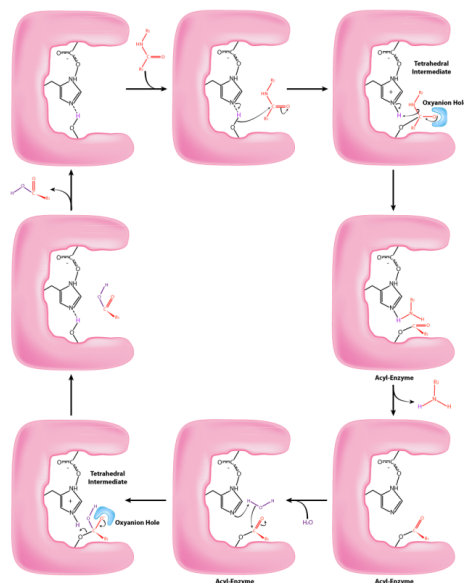


Figure 6.2.1: Serine protease mechanism

The process starts with the binding of the substrate in the S1 pocket. The S1 pocket in chymotrypsin has a hydrophobic hole in which the substrate is bound. Preferred substrates will include amino acid side chains that are hydrophobic, like phenylalanine. If an ionized side chain, like that of glutamic acid binds in the S1 pocket, it will quickly exit, much like water would avoid an oily interior. When the proper substrate binds, it stays and its presence induces an ever so slight shift in the shape of the enzyme. This subtle shape change on the binding of the proper substrate starts the steps of the catalysis and is the reason that the enzyme shows specificity for cutting at specific enzyme positions in the target protein. Only amino acids with the side chains that interact well with the S1 pocket start the catalytic wheels turning.

The slight changes in shape of the enzyme upon binding of the proper substrate cause changes in the positioning of three amino acids (aspartic acid, histidine, and serine) in the active site known as the catalytic triad, during the second step of the catalytic action. The shift of the negatively charged aspartic acid towards the electron rich histidine ring favors the abstraction of a proton by the histidine from the hydroxyl group on the side chain of serine, resulting in production of a very reactive alkoxide ion in the active site. Since the active site at this point also contains the polypeptide chain positioned with the phenylalanine side chain embedded in the S1 pocket, the alkoxide ion performs a nucleophilic attack on the peptide bond on the carboxyl side of phenylalanine sitting in the active site. This reaction, which is the third step of catalysis, breaks the bond and causes two things to happen. First, one end of the original polypeptide is freed and exits the active site. The second is that the end containing the phenylalanine is covalently linked to the oxygen of the serine side chain. At this point we have completed the first (fast) phase of the catalysis.

The second phase of the catalysis by chymotrypsin is slower. It requires that the covalent bond between phenylalanine and serine's oxygen be broken so the peptide can be released and the enzyme can return to its original state. The process starts with entry of water into the active site. Water is attacked in a fashion similar to that of the serine side chain in the first phase, creating a reactive hydroxyl group that performs a nucleophilic attack on the phenylalanine-serine bond, releasing it and replacing the proton on serine. The second peptide is released in the process and the reaction is complete with the enzyme back in its original

state.

Contributors

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

This page titled [6.2: Chymotrypsin](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

6.3: Trypsin

This page titled [6.3: Trypsin](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

7: Enzyme Active Site and Substrate Specificity

Learning Objectives

- Describe models of substrate binding to an enzyme's active site.

Enzyme Active Site and Substrate Specificity

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

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

Active Sites and Environmental Conditions

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

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

Induced Fit and Enzyme Function

For many years, scientists thought that enzyme-substrate binding took place in a simple "lock-and-key" fashion. This model asserted that the enzyme and substrate fit together perfectly in one instantaneous step. However, current research supports a more refined view called induced fit. As the enzyme and substrate come together, their interaction causes a mild shift in the enzyme's structure that confirms an ideal binding arrangement between the enzyme and the substrate. This dynamic binding maximizes the enzyme's ability to catalyze its reaction.

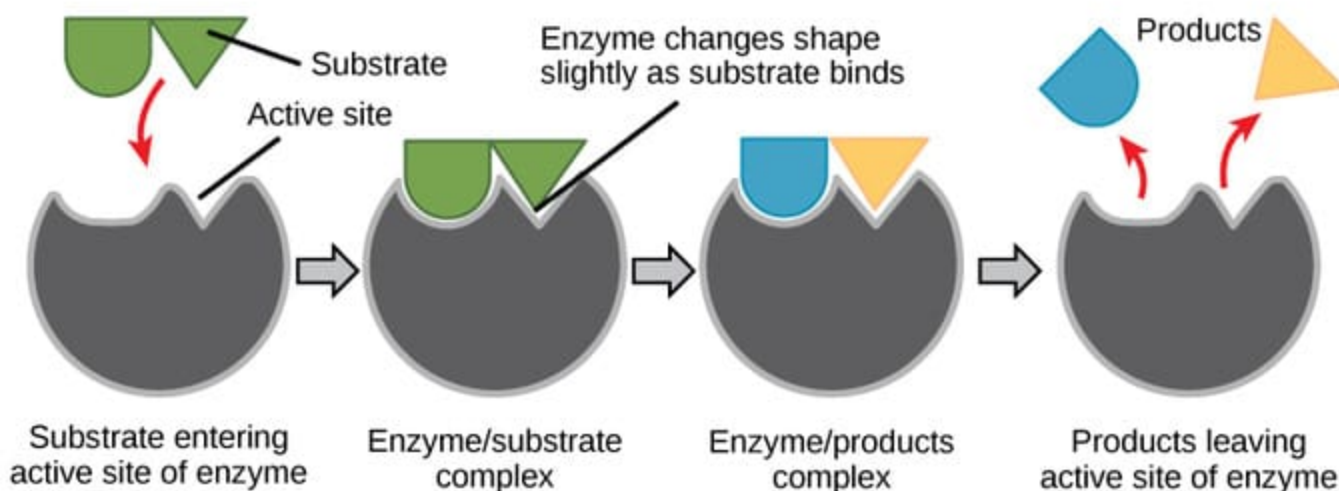


Figure: **Induced Fit:** According to the induced fit model, both enzyme and substrate undergo dynamic conformational changes upon binding. The enzyme contorts the substrate into its transition state, thereby increasing the rate of the reaction.

Enzyme-Substrate Complex

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

LICENSES AND ATTRIBUTIONS

CC LICENSED CONTENT, SPECIFIC ATTRIBUTION

- OpenStax College, Biology. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** <http://cnx.org/content/m44429/latest...ol11448/latest>. **License:** CC BY: Attribution
- Boundless. **Provided by:** Boundless Learning. **Located at:** www.boundless.com/biology/de...llosteric-site. **License:** CC BY-SA: Attribution-ShareAlike
- cofactor. **Provided by:** Wiktionary. **Located at:** en.wiktionary.org/wiki/cofactor. **License:** CC BY-SA: Attribution-ShareAlike
- coenzyme. **Provided by:** Wiktionary. **Located at:** en.wiktionary.org/wiki/coenzyme. **License:** CC BY-SA: Attribution-ShareAlike
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_04.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_06.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_05.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_07.jpg. **License:** CC BY: Attribution
- OpenStax College, Biology. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** <http://cnx.org/content/m44429/latest...ol11448/latest>. **License:** CC BY: Attribution
- active site. **Provided by:** Wikipedia. **Located at:** en.Wikipedia.org/wiki/active%20site. **License:** CC BY-SA: Attribution-ShareAlike
- substrate. **Provided by:** Wiktionary. **Located at:** en.wiktionary.org/wiki/substrate. **License:** CC BY-SA: Attribution-ShareAlike
- induced fit. **Provided by:** Wikipedia. **Located at:** en.Wikipedia.org/wiki/induced%20fit. **License:** CC BY-SA: Attribution-ShareAlike

- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_04.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_06.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_05.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_07.jpg. **License:** CC BY: Attribution
- OpenStax College, Enzymes. October 16, 2013. **Provided by:** OpenStax CNX. **Located at:** http://cnx.org/content/m44429/latest...e_06_05_03.jpg. **License:** CC BY: Attribution

Key Points

- The enzyme 's active site binds to the substrate.
- Increasing the temperature generally increases the rate of a reaction, but dramatic changes in temperature and pH can denature an enzyme, thereby abolishing its action as a catalyst.
- The induced fit model states an substrate binds to an active site and both change shape slightly, creating an ideal fit for catalysis.
- When an enzyme binds its substrate it forms an enzyme-substrate complex.
- Enzymes promote chemical reactions by bringing substrates together in an optimal orientation, thus creating an ideal chemical environment for the reaction to occur.
- The enzyme will always return to its original state at the completion of the reaction.

Key Terms

- **substrate:** A reactant in a chemical reaction is called a substrate when acted upon by an enzyme.
- **induced fit:** Proposes that the initial interaction between enzyme and substrate is relatively weak, but that these weak interactions rapidly induce conformational changes in the enzyme that strengthen binding.
- **active site:** The active site is the part of an enzyme to which substrates bind and where a reaction is catalyzed.

8: Enzyme Regulation

Learning Objectives

1. Compare and contrast the genetic control of enzyme activity (enzyme synthesis) in bacteria with the control of enzyme activity through feedback inhibition.
2. Compare and contrast an inducible operon with a repressible operon and give an example of each.
3. Compare how the presence or absence of tryptophan affects the *trp* operon.
4. Compare how the presence or absence of lactose affects the *lac* operon.
5. Compare how the presence or absence of an inducer affects activators.
6. Briefly describe how small RNAs can regulate enzyme activity.
7. Define the following:
 - a. repressor
 - b. inducer
 - c. activator
 - d. enhancer
 - e. small RNAs
8. Compare and contrast competitive inhibition with noncompetitive inhibition.

In living cells, there are hundreds of different enzymes working together in a coordinated manner. Living cells neither synthesize nor break down more material than is required for normal metabolism and growth. All of this necessitates precise control mechanisms for turning metabolic reactions on and off. There is tremendous diversity in the mechanisms bacteria use to regulate enzyme synthesis and enzyme activity. For pretty much every step between the activation of a gene and the final enzyme reaction from that gene product there is some bacterial mechanism for regulation that step. Here we will look at several well studied examples.

Genetic Control of Enzyme Synthesis through Repression, Induction, or Enhancement of Transcription

Genetic control of enzyme activity refers to controlling transcription of the mRNA needed for an enzyme's synthesis. In prokaryotic cells, this involves the induction, repression, or enhancement of enzyme synthesis by regulatory proteins that can bind to DNA and either induce, block, or enhance the function of RNA polymerase, the enzyme required for transcription. The regulatory proteins are often part of either an operon or a regulon. An operon is a set of genes transcribed as a polycistronic message that is collectively controlled by a regulatory protein. A regulon is a set of related genes controlled by the same regulatory protein but transcribed as monocistronic units. Regulatory proteins may function either as repressors, activators, or enhancers.

a. Repressors

Repressors are regulatory proteins that block transcription of mRNA. They do this by binding to a portion of DNA called the operator (operators are often called boxes now) that lies downstream of a promoter. The binding of the regulatory protein to the operator prevents RNA polymerase from binding to the promoter and transcribing the coding sequence for the enzymes. This is called negative control and is mostly in biosynthetic reactions where a bacterium only makes a molecule like a particular amino acid when that amino acid is not present in the cell.

Repressors are allosteric proteins that have a binding site for a specific molecule. Binding of that molecule to the allosteric site of the repressor can alter the repressor's shape that, in turn affects its ability to bind to DNA. This can work in one of two ways:

1. Some repressors are synthesized in a form that cannot by itself bind to the operator. This is referred to as a repressible system. The binding of a molecule called a corepressor, however, alters the shape of the regulatory protein to a form that can bind to the operator and subsequently block transcription. An example of this type of **repressible system** is the *trp* operon in *Escherichia coli* that encodes the five enzymes in the pathway for the biosynthesis of the amino acid tryptophan. In this case, the repressor protein coded for by the *trp* regulatory gene, normally does not bind to the operator region of the *trp* operon and the five enzymes

needed to synthesize the amino acid tryptophan are made (Figure 8.1A and Figure 8.1B).

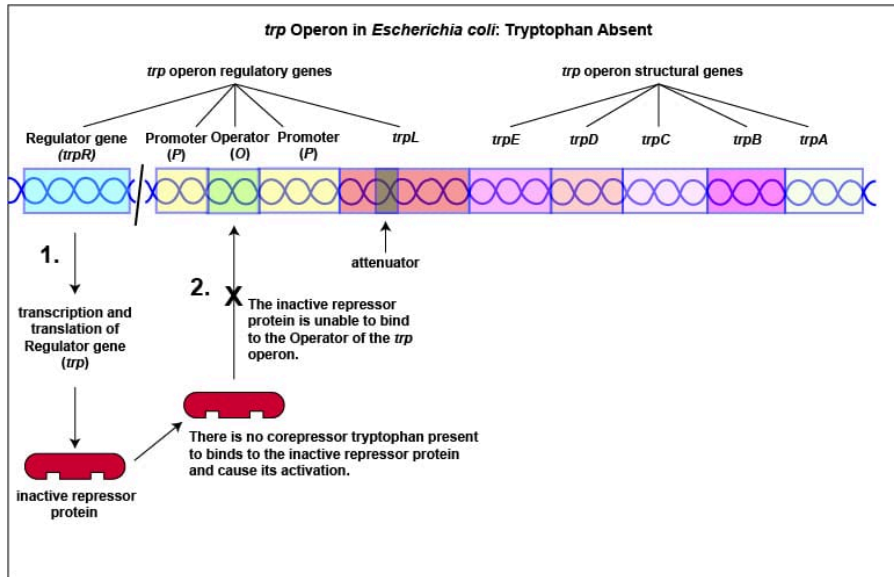


Figure 8.1A: A Repressible Operon in the Absence of a Corepressor (The Tryptophan Operon). Step 1: The Regulator gene codes for an inactive repressor protein. Step 2: The inactive repressor protein is unable to bind to the Operator region of the *trp* operon.

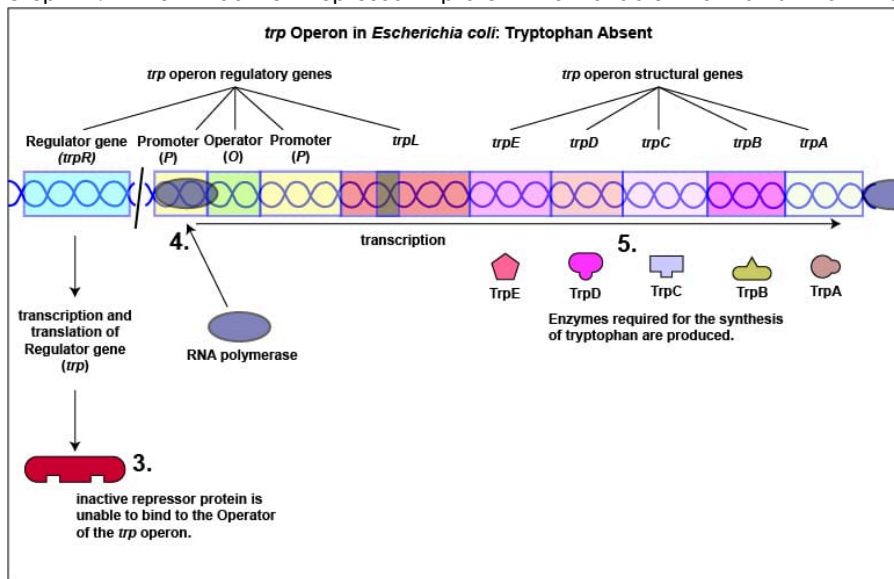


Figure 8.1B: A Repressible Operon in the Absence of a Corepressor (The Tryptophan Operon). Step 3: Since the inactive repressor protein is unable to bind to the Operator region, RNA polymerase (the enzyme responsible for the transcription of genes) is able to bind to the Promoter region of the *trp* operon. Step 4: RNA polymerase is now able to transcribe the five *trp* operon structural genes (*trpE*, *trpD*, *trpC*, *trpB*, and *trpA*) into mRNA. Step 5: With the transcription of these genes, the five enzymes needed for the bacterium to synthesize the amino acid tryptophan are now made. TrpE and TrpD are the two subunits for making anthranilate synthetase, the enzyme that catalyzes the first two reactions in the tryptophan pathway. TrpC is indole glycerolphosphate synthetase, the enzyme that catalyzes the next two steps in the pathway. TrpB and TrpA are subunits for making tryptophan synthetase, the enzyme that catalyzes the synthesis of tryptophan from indole-glycerol phosphate and serine.

Tryptophan, the end product of these enzyme reactions, however, functions as a corepressor. Once sufficient tryptophan has been synthesized, the cell needs to terminate its synthesis. The tryptophan is able to bind to a site on the allosteric repressor protein, changing its shape and enabling it to interact with the *trp* operator region. Once the repressor binds to the operator, RNA polymerase is unable to bind to the promoter and transcribe the genes for tryptophan biosynthesis. Therefore, when sufficient tryptophan is present, transcription of the enzymes that allows for its biosynthesis are turned off (Figure 8.2A and Figure 8.2B).

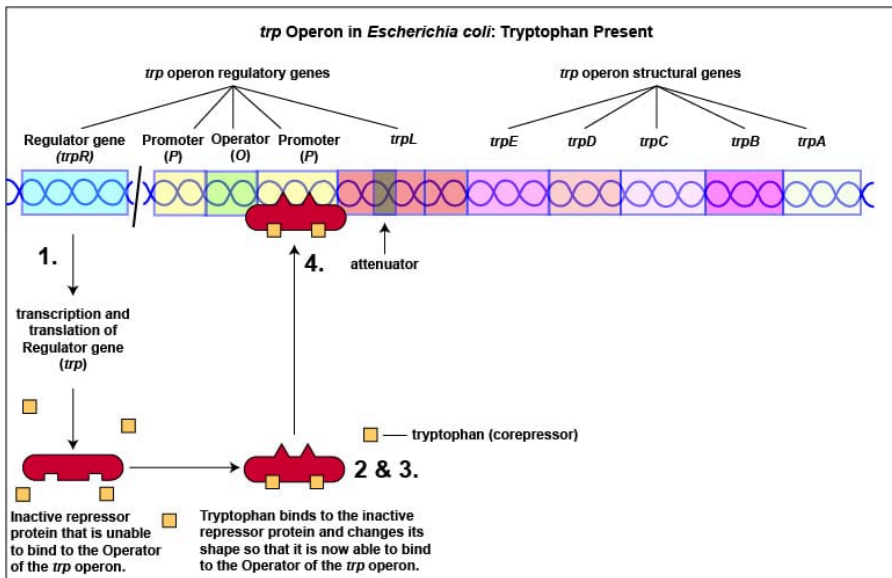
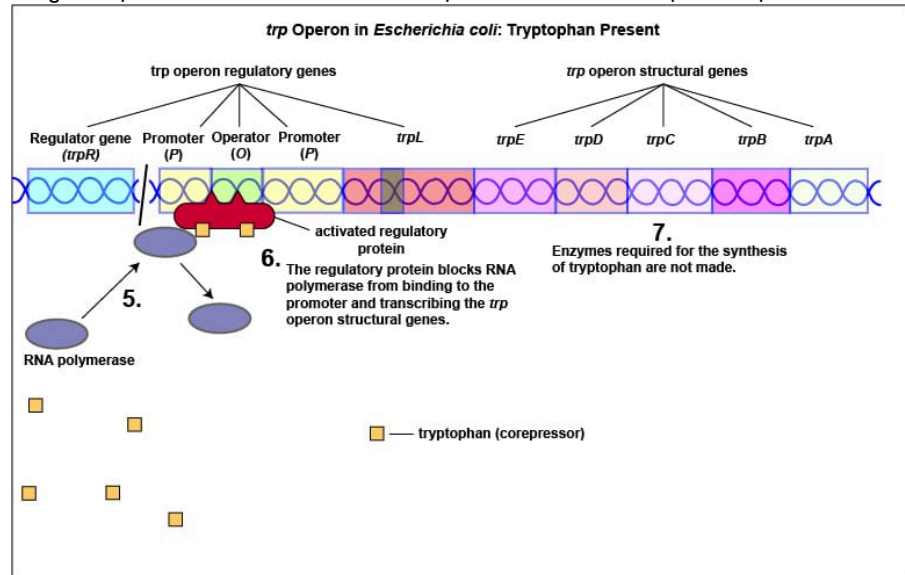


Figure 8.2A: A Repressible Operon in the

Presence of a Corepressor (The Tryptophan Operon). Step 1: The Regulator gene codes for an inactive repressor protein. Step 2: If the corepressor tryptophan is present, it binds to to the inactive repressor protein. Step 3: The binding of the corepressor causes inactive repressor protein to change shape and become activated. Step 4: The activated repressor protein then binds to



the Operator region of the *trp* operon.

Figure

8.2B: A Repressible Operon in the Presence of a Corepressor (The Tryptophan Operon). Step 5: With the active repressor protein bound to the Operator region, RNA polymerase (the enzyme responsible for the transcription of genes) is unable to bind to the Promoter region of the *trp* operon. Step 6: If RNA polymerase does not bind to the Promoter region, the five *trp* operon structural genes are not transcribed into mRNA. Step 7: Without the transcription of the five genes, the five enzymes needed for the bacterium to synthesize the amino acid tryptophan are not made.

In addition to repression, the expression of the *trp* operon is also regulated by attenuation. The *trpL* gene codes for a mRNA leader sequence that controls operon expression through attenuation. This leader sequence mRNA consists of domains 1, 2, 3, and 4. Domain 3 can base pair with either domain 2 or domain 4.

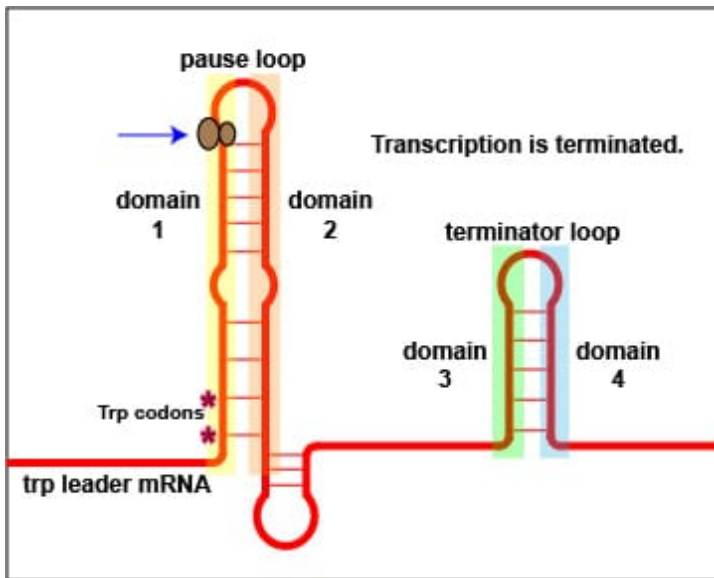


Figure 8.3A: Attenuation in the Trp Operon of *Escherichia coli*:

Excess Tryptophan. When excess tryptophan is available, there is a rapid translation of the early trp leader mRNA enabling domain 2 to pair with domain 1 and form a pause loop. The ribosome pauses at a stop codon (arrow) causing domain 3 to pair with domain 4 and form a terminator loop. Transcription of the remainder of the trp operon is terminated. Rapid initial translation is able to occur with excess tryptophan present because there is a sufficient quantity of Trp tRNA available to translate the two Trp

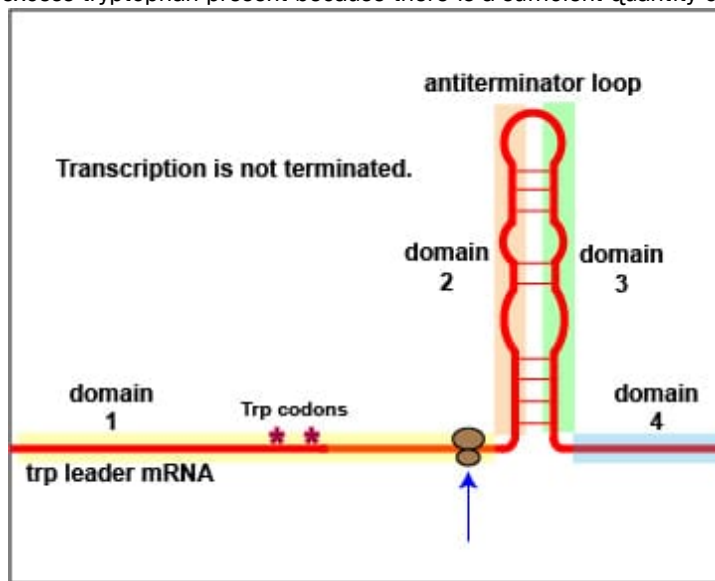


Figure 8.3B: Attenuation in the Trp Operon of *Escherichia coli*: Low Levels of Tryptophan. When tryptophan is limited, there is a slow translation of the early trp leader mRNA which enables domain 2 to pair with domain 3 and form an antiterminator loop. Transcription of the remainder of the trp operon continues and the enzymes required for tryptophan synthesis are made. Slow initial translation is able to occur with low levels of tryptophan present because there is limited Trp tRNA available to translate the two Trp codons (asterisks) causing the ribosome to stall at the Trp codons and enabling domain 2 to pair with domain 3 rather than domain 1.

codons (asterisks). At high tryptophan concentrations, domains 3 and 4 pair in such a way as to form stem and loop structures that block the transcription of the remainder of the leader sequence mRNA and subsequently, the transcription of the structural genes for tryptophan biosynthesis (Figure 8.3A). However, at low concentrations of tryptophan, domains 3 and 2 pair. This pairing allows for the full transcription of the leader sequence mRNA, as well as that of the structural genes for tryptophan biosynthesis (Figure 8.3 B).

At high tryptophan concentrations, domains 3 and 4 pair in such a way as to form stem and loop structures that block the transcription of the remainder of the leader sequence mRNA and subsequently, the transcription of the structural genes for tryptophan biosynthesis (Figure 8.3A). However, at low concentrations of tryptophan, domains 3 and 2 pair. This pairing allows for the full transcription of the leader sequence mRNA, as well as that of the structural genes for tryptophan biosynthesis (Figure 8.3 B).

2. Other repressors are synthesized in a form that readily binds to the operator and blocks transcription. However, the binding of a molecule called an inducer alters the shape of the regulatory protein in a way that now blocks its binding to the operator and thus permits transcription. This is referred to as an **inducible system**.

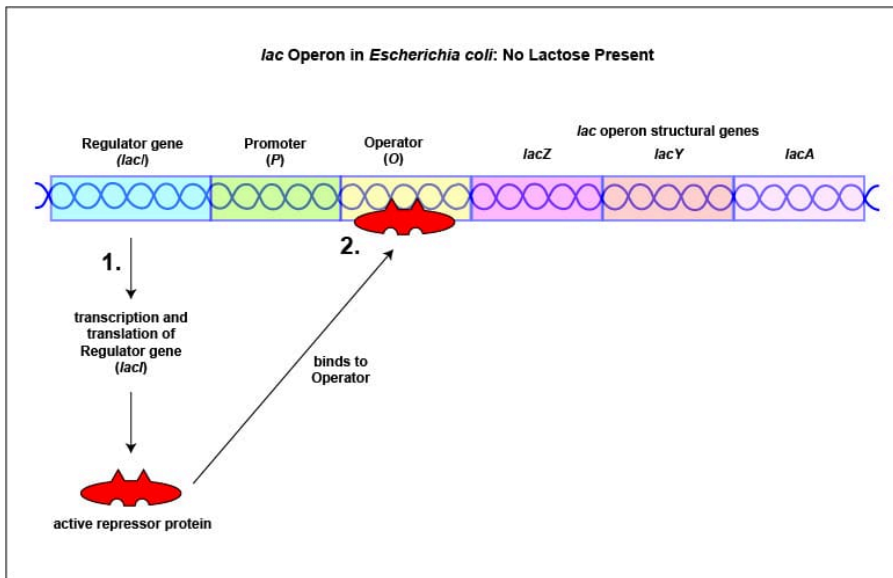


Figure 8.4A: An Inducible Operon in the Absence of an Inducer (The Lactose Operon of *Escherichia coli*). Step 1: The Regulator gene (*lacI*) codes for an active repressor protein. Step 2: The repressor protein then binds to the Operator region of the *lac* operon.

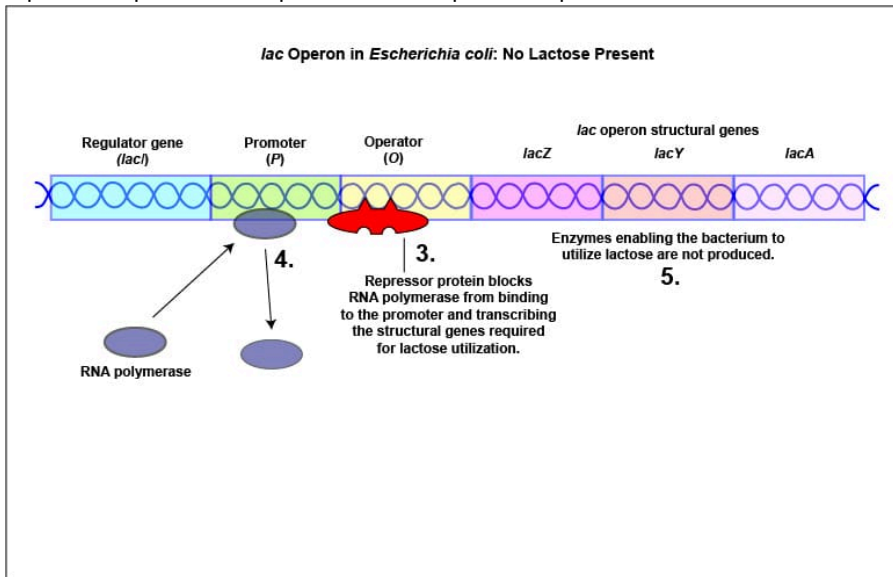


Figure 8.4B: An Inducible Operon in the Absence of an Inducer (The Lactose Operon of *Escherichia coli*). Step 3: With the active repressor protein bound to the Operator region, RNA polymerase (the enzyme responsible for the transcription of genes) is unable to bind to the Promoter region of the operon. Step 4: If RNA polymerase does not bind to the Promoter region, the 3 *lac* operon structural genes (*lacZ*, *lacY*, and *lacA*) are not transcribed into mRNA. Step 5: Without the transcription of these genes, the enzymes needed for the utilization of the sugar lactose by the bacterium are not synthesized.

An example of an inducible system is the *lac* operon that encodes for the three enzymes needed for the degradation of lactose by *E. coli*. *E. coli* will only synthesize the enzymes it requires to utilize lactose if that sugar is present in the surrounding environment. In this case, lactose functions as an inducer. In the absence of lactose, the active repressor protein binds to the operator and RNA polymerase is unable to bind to the promoter and transcribe the genes for utilization of lactose. As a result, the enzymes needed for the utilization of lactose are not synthesized (Figure 8.4A and Figure 8.4B). When lactose, the inducer, is present, a metabolite of lactose called allolactose binds to the allosteric repressor protein and causes it to change shape in such a way that it is no longer able to bind to the operator. Now RNA polymerase is able to transcribe the three *lac* operon structural genes and the bacterium is able to synthesize the enzymes required for the utilization of lactose (Figure 8.5A and Figure 8.5B).

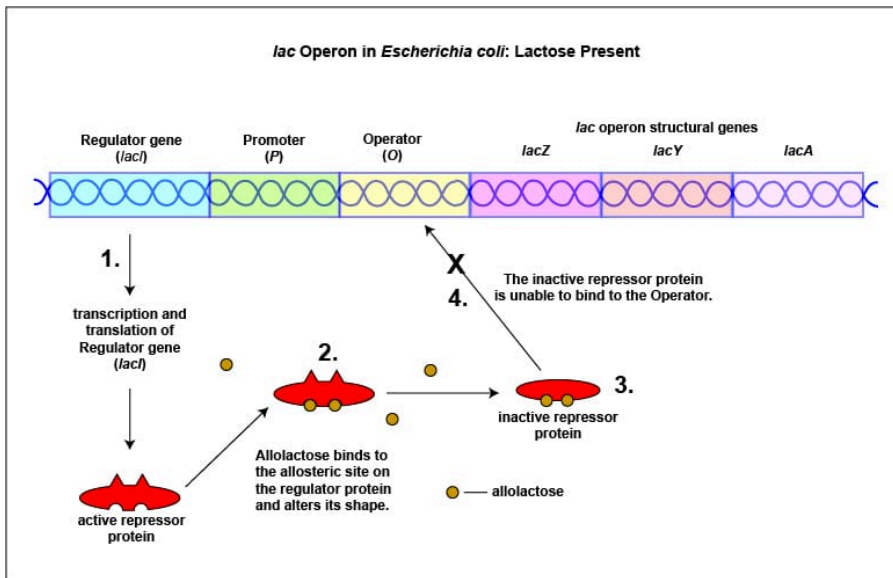


Figure 8.5A: An Inducible Operon in the

Presence of an Inducer(The *Escherichia coli* Lactose Operon) Step 1: The Regulator gene codes for an active repressor protein. Step 2: Allolactose (consisting of glucose and galactose), a metabolite of the inducer molecule lactose, binds to the active repressor protein. Step 3: The binding of the inducer alters the shape of the repressor protein making it inactive. Step 4: The inactive repressor protein is no longer able to bind to the Operator region of the *lac* operon.

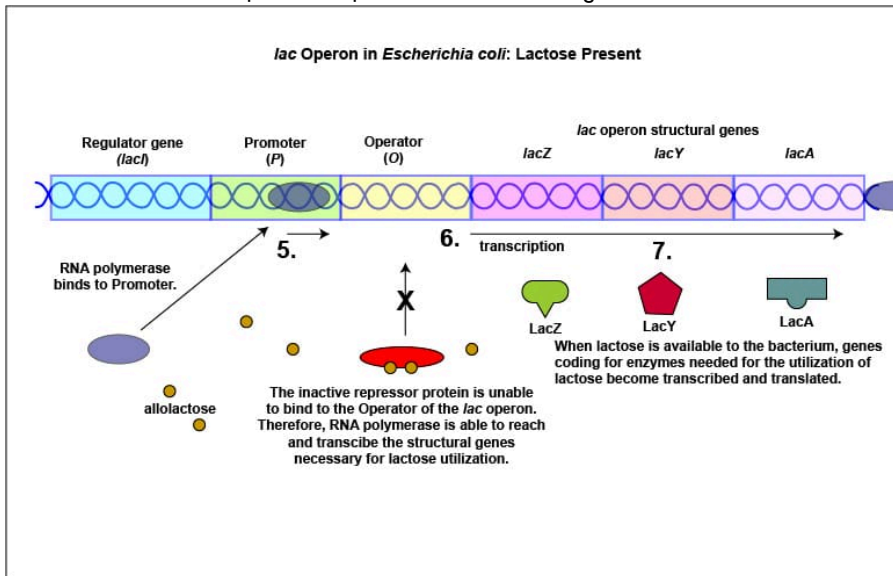


Figure 8.5B: An Inducible Operon in the

Presence of an Inducer (The *Escherichia coli* Lactose Operon) Step 5: Since the inactive repressor protein is unable to bind to the Operator region, RNA polymerase (the enzyme responsible for the transcription of genes) is now able to bind to the Promoter region of the *lac* operon. Step 6: RNA polymerase is now able to transcribe the three *lac* operon structural genes (*lacZ*, *lacY*, and *lacA*) into mRNA. Step 7: With the transcription of these genes, the enzymes needed for the bacterium to utilize the sugar lactose are now synthesized. The *lacZ* gene codes for LacZ (beta-galactosidase), an enzyme that breaks down lactose into glucose and galactose. The *lacY* gene codes for LacY (beta-galactosidase permease), an enzyme which transports lactose into the bacterium. The *lacA* gene codes for LacA (transacetylase), of uncertain function in lactose catabolism.

b. Activators

Activators are **regulatory proteins** that promote transcription of mRNA. Activators control genes that have a promoter to which RNA polymerase cannot bind. The promoter lies adjacent to a segment of DNA called the activator-binding site. The activator is an allosteric protein synthesized in a form that cannot normally bind to the activator-binding site. As a result, RNA polymerase is unable to bind to the promoter and transcribe the genes ([Figure 8.6](#)). However, binding of a molecule called an inducer to the activator alters the shape of the activator in a way that now allows it to bind to the activator-binding site. The binding of the activator to the activator-binding site, in turn, enables RNA polymerase to bind to the promoter and initiate transcription ([Figure](#)

8.7A and Figure 8.7B). This is called positive control and is mostly n in catabolic reactions where a bacterium only makes enzymes for the catabolism of a substrate when that substrate is available to the cell.

c. Enhancers

Enhancers are regulatory proteins that bind to DNA located some distance from the operon they control by working with DNA-bending proteins. The DNA-binding proteins bend the DNA in a way that now allows the enhancer to interact with the promoter in such a way that RNA polymerase can now bind and initiate transcription (Figure 8.8).

2. Genetic Control of Enzyme Synthesis through Promoter Recognition and through DNA Supercoiling

- a. Promoter Recognition: The specific sigma factors that bind to RNA polymerase determine which operon will be transcribed.
- b. DNA Supercoiling: DNA supercoiling can change the tertiary shape of a DNA molecule from its normal form to one that has a left-handed twist called Z-DNA. The activities of some promoters are decreased with Z-DNA while others are increased.

3. Genetic Control of Enzyme Synthesis through the Translational Control of Enzyme Synthesis

- a. RNA interference (RNAi)

RNA interference (RNAi) is a process whereby small non-coding regulatory RNAs (ncRNAs) such as microRNAs (miRNAs) regulate gene expression. These ncRNAs are regulatory molecules that are complementary to an early portion of the 5' end of the mRNA coding for the enzyme. When the small RNA binds to the mRNA by complementary base pairing, ribosomes cannot attach to the mRNA blocking its translation. As a result, the enzyme is not made (Figure 8.9). In bacteria these ncRNAs are often called small RNAs (sRNAs); in animal cells, plant cells, and viruses they are often called microRNAs (miRNA).

- b. Ribosomal Proteins (r-proteins)

Ribosomal proteins bind to rRNA to form ribosomal subunits. Because the nucleotide base sequence for the mRNA coding for the r-proteins has similarities to that of the rRNA to which that r-protein binds during subunit formation, r-proteins not yet incorporated into ribosomal subunits can bind to that mRNA and block translation

4. Controlling the Enzyme's Activity (Feedback Inhibition).

Enzyme activity can be controlled by competitive inhibition and non-competitive inhibition.

- a. With what is termed non-competitive inhibition, the inhibitor is the end product of a metabolic pathway that is able to bind to a second site (the allosteric site) on the enzyme. Binding of the inhibitor to the allosteric site alters the shape of the enzyme's active site thus preventing binding of the first substrate in the metabolic pathway. In this way, the pathway is turned off (Figure 8.10).
- b. In the case of what is called competitive inhibition, the inhibitor is the end product of an enzymatic reaction. That end product is also capable of reacting with the enzyme's active site and prevents the enzyme from binding its normal substrate. As a result, the end product is no longer synthesized (Figure 8.11).

Summary

1. In living cells there are hundreds of different enzymes working together in a coordinated manner, and since cells neither synthesize nor break down more material than is required for normal metabolism and growth, precise enzyme regulation is required for turning metabolic reactions on and off.
2. There is tremendous diversity in the mechanisms bacteria use to regulate enzyme synthesis and enzyme activity.
3. Ways in which enzymes can be controlled or regulated include controlling the synthesis of the enzyme (genetic control) and

controlling the activity of the enzyme (feedback inhibition).

4. In prokaryotes, genetic control of enzyme activity includes the induction or repression of enzyme synthesis by regulatory proteins that can bind to DNA and either block or enhance the function of RNA polymerase, the enzyme required for transcription.
5. An operon is a set of genes collectively controlled by a regulatory protein.
6. Regulatory proteins may function either as repressors or activators.
7. Repressors are regulatory proteins that block transcription of mRNA by preventing RNA polymerase from transcribing the coding sequence for the enzymes.
8. Some repressors, as in the case of the trp operon, are synthesized in a form that cannot by itself bind to the operator. This is referred to as a repressible system. The binding of a molecule called a corepressor, however, alters the shape of the regulatory protein to a form that can bind to the operator and subsequently block transcription.
9. Some repressors, as in the case of the lac operon, are synthesized in a form that readily binds to the operator and blocks transcription. However, the binding of a molecule called an inducer alters the shape of the regulatory protein in a way that now blocks its binding to the operator and thus permits transcription. This is referred to as an inducible system.
10. Activators are regulatory proteins that promote transcription of mRNA by enabling RNA polymerase to transcribing the coding sequence for the enzymes.
11. Enhancers are regulatory proteins that bind to DNA located some distance from the operon they control by working with DNA-bending proteins. The DNA-bending proteins bend the DNA in a way that now allows the enhancer to interact with the promoter in such a way that RNA polymerase can now bind and initiate transcription
12. Bacteria also use translational control of enzyme synthesis. One method is for the bacteria to produce noncoding RNA (ncRNA) molecules that are complementary to the mRNA coding for the enzyme, and when the small RNA binds to the mRNA by complementary base pairing, ribosomes cannot attach to the mRNA, the mRNA is not transcribed and translated into protein, and the enzyme is not made. In bacteria, these ncRNAs are often called small RNAs (sRNAs).
13. Feedback inhibition controls the activity of the enzyme rather than its synthesis and can be noncompetitive or competitive.
14. In the case of non-competitive inhibition, the inhibitor is the end product of a metabolic pathway that is able to bind the allosteric site on the enzyme. Binding of the inhibitor to the allosteric site alters the shape of the enzyme's active site thus preventing binding of the first substrate in the metabolic pathway. In this way, the pathway is turned off.
15. In the case of what is called competitive inhibition, the inhibitor is the end product of an enzymatic reaction. That end product is also capable of reacting with the enzyme's active site and prevents the enzyme from binding its normal substrate. As a result, the end product is no longer synthesized.

This page titled [8: Enzyme Regulation](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Gary Kaiser](#).

- [Ions in Solution 2 \(Worksheet\)](#) by [Mark Draganjac](#) is licensed [CC BY-NC-SA 4.0](#). Original source: <http://myweb.astate.edu/mdraganj/worksheetlist.html>.

8.1: Regulation of Enzyme Activity

Figure 8.1.7 (and 9) also illustrates the effects of two different types of inhibition on the different components of enzyme kinetics. Enzymes can be slowed down or even prevented from catalyzing reactions in many ways including preventing the substrate from entering the active site or preventing the enzyme from altering conformation to catalyze the reaction. The inhibitors that do this can do so either reversibly or irreversibly. The irreversible inhibitors are also called **inactivators**, and either bind to the enzyme with such high affinity as to be virtually irreversible, or they actually form covalent bonds with the enzyme. Reversible inhibitors are generally grouped into two basic types: competitive and non-competitive.

Finasteride (trade names include Propecia and Proscar) is an irreversible inhibitor that binds very tightly to the enzyme 5- α -reductase, used in converting testosterone to dihydrotestosterone. It is used in the treatment of male pattern baldness, benign prostatic hyperplasia, and prostate cancer.

Aspirin is an example of an irreversible inhibitor that actually forms a covalent bond with the enzyme. The aspirin (acetylsalicylic acid) transfers its acetyl group onto a serine residue on cyclooxygenase-2 (COX-2). This stops the production of inflammation-producing prostaglandins and thromboxanes by COX-2.

Methotrexate is a competitive inhibitor of dihydrofolate reductase (DHFR), an enzyme that synthesizes tetrahydrofolate, which is a precursor for purine synthesis, and therefore for DNA and RNA. It has a very similar molecular structure for folic acid, the natural substrate of DHFR. Methotrexate is used as an anti-cancer drug because it affects rapidly reproducing cells (which need to make DNA sooner than other cells) more than non-cancerous cells.

Competitive inhibition is perhaps the simplest to understand. The inhibitor molecule competes directly with the substrate for the active site of an unbound enzyme. If an inhibitor binds to the active site, the substrate is unable to do so until the inhibitor has vacated the site. Thus, one could potentially overwhelm competitive inhibition with sufficiently larger concentrations of substrate so that the probability that the enzyme bumps into a substrate to bind becomes exceeding large compared to the probability of bumping into an inhibitor. Normal, uninhibited V_{max} is then achieved despite the presence of the competitive inhibitor, which has only affected the K_m , that is, the concentration of substrate needed to reach $V_{max}/2$. This is the kinetic signature of competitive inhibitors: with increasing inhibitor concentrations, K_M is increased but V_{max} is unaffected.

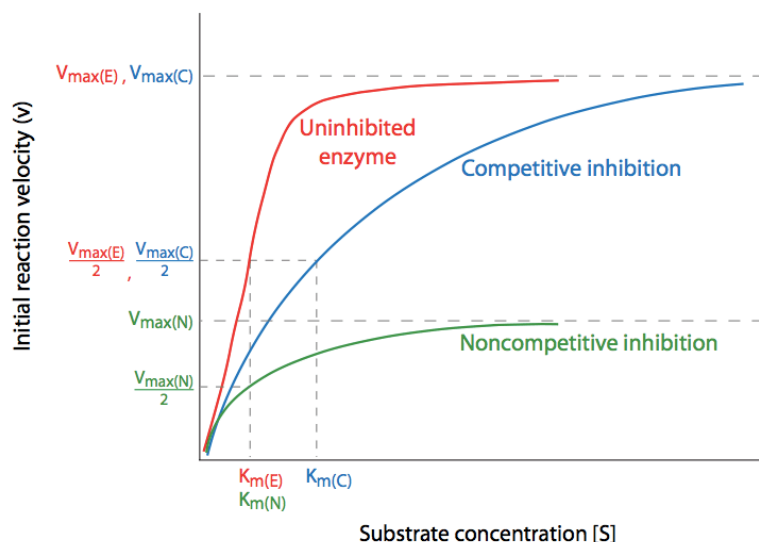


Figure 8.1.9. Saturation curves for enzyme-catalyzed reactions without inhibitor (red) with competitive inhibitor (blue) at a constant concentration, and with noncompetitive inhibitor (green) at a constant concentration.

Non-competitive inhibition involves inhibiting the enzyme by altering its ability to complete the catalyzed reaction through

binding of the enzyme at a position that is not the active site. When the inhibitor binds to the enzyme, it causes a change, usually conformational, that may either prevent the enzyme from binding the substrate, or prevent the enzyme from acting upon a bound substrate. In either case, increasing the availability of substrate will not ultimately overcome the effect of the inhibitor. Thus, V_{max} is reduced not because some proportion of the enzymes are no longer usable, but because the enzymes that are available have the same access to substrate as it would without inhibitor (that is, it is not in competition with an inhibitor), the K_m is not affected.

Allostery

Non-competitive regulation is one example of *allosteric regulation* of enzymes. Allosteric interactions occur when the binding of a ligand (not necessarily a substrate) to a protein influences the binding of another ligand to the protein at a separate binding site. These kinds of interactions can be either positive (activating) or negative (inhibitory), and either homotropic (both ligands are identical) or heterotropic (ligands are different). Interestingly, sometimes the regulator ligand may actually be a product of the catalyzed reaction. In this kind of feedback mechanism, the progress of a reaction is self-regulating.



Figure 8.1.10. Feedback control loops. (Left) Negative feedback is a common biological mechanism for self-regulating processes. The product of the reaction, either by itself or in conjunction with other molecules, acts as an allosteric inhibitor of the enzyme. (Right) Positive feedback is less common because it can lead to rapid expansion of the scope of a reaction, and requires an external (relative to the enzymatic reaction) mechanism for slowing or stopping the reaction.

Many of the enzymes in metabolic pathways (chapters 5 and 6) are regulated by a naturally occurring non-competitive inhibitor. One example is phosphofructokinase (PFK), which is involved in glycolysis, which produces ATP for the cell. However, if there are high enough levels of ATP in the cell that other cellular processes aren't using, then it can bind to phosphofructokinase outside of its active site (it binds fructose-6-phosphate) and turn it off. This blocks glycolysis and production of excess ATP when the cell does not need it. As ATP is used up, there is less available to inhibit PFK and glycolysis starts back up.

Modeling Allosteric Behavior

There are two models for such allosteric interactions.

- The **symmetry model**, also known as the concerted model, or MWC model (Monod, Wyman, and Changeux, 1965), proposes that the allosteric enzyme is an oligomer of several subunits, each of which are symmetrically related, and can be in either a "tensed" or "relaxed" state, but all of the subunits are in the same state and at equilibrium. When a ligand binds, it changes the state of the subunit(s) to which it binds, and to maintain equilibrium, that in turn causes the state of the other subunits to match, thus altering binding properties for a subsequent ligand.
- The **sequential model**, or the KNF model (Koshland, Nemethy, and Filmer, 1966), proposes something quite different: although it also supposes subunits in either tensed or relaxed states, it does not require all subunits to be connected in such a way as to mandate that all subunits be in the same state, and therefore conformational changes in one subunit need not be propagated to all. Instead, using an induced-fit model of ligand binding rather than the more rigid basic lock and key mechanism, it suggests that when a ligand binds to the enzyme, it induces a slight conformational change

in the active site that increases its affinity for the ligand. The conformational change may slightly alter the conformation of other subunits of the enzyme, but may not constitute a state change between relaxed and tensed depending on how tightly the subunits are interacting. However, the change is enough to increase substrate affinity in adjacent subunits.

Advanced: Uncompetitive and Mixed inhibition

Most cell/molecular courses stop the discussion of enzyme inhibitors at competitive vs non-competitive based on their kinetic profiles. However, it should be noted that if you take a biochemistry course, you may encounter the terms *uncompetitive* inhibitor and *mixed* inhibitor. These terms are defined not just by the enzyme kinetics, but the mechanism of interaction:

- *Uncompetitive inhibitors* bind only to the enzyme-substrate (ES) complex, and not to the enzyme before it has encountered substrate. This leads to decreased $V_{\max}$ and decreased K_m .
- *Mixed inhibition* means that the inhibitor can bind to either enzyme alone or the enzyme-substrate complex. Because the affinities of the inhibitor for the two forms of the enzyme are different, and because part of it depends on substrate concentration while the other kind of binding does not, generally, $V_{\max}$ decreases and K_m increases. Non-competitive inhibition is a special case of mixed inhibition in which the catalytic activity of the enzyme is diminished or abolished, but the ability to bind substrate is unaltered.

Other Mechanisms of Inhibition

While important, especially pharmaceutically, the use of enzyme inhibitors is not the only way to regulate enzymes. There are numerous examples of one type of enzyme activating or inhibiting another. The most common general example are the protein kinases. These enzymes phosphorylate (transfer phosphate group to) other enzymes and thereby activate them. Kinases are generally fast and very specific, and this is an efficient method for activating large numbers of particular enzymes quickly. Conversely, protein phosphatases are enzymes (also quite fast, but much less specific than kinases) that remove the phosphate groups from phosphorylated proteins, thereby turning off those enzymes. Keep in mind that this is a generalization, and that not all phosphorylations are activating. In addition to enzymatic inhibition of enzymes, there is also inhibition by binding and sequestration of the substrates. In fact, the antibiotic vancomycin works just this way, binding to the substrate peptide for transpeptidase and preventing the enzyme from recognizing it. Transpeptidase normally helps stabilize the cell wall of certain bacteria by altering some of the proteins, and without its activity, the protection of the cell wall is compromised and the bacteria may be more easily killed.

Well over half of the enzymes discovered so far do not act in the simplistic Michaelis-Menten one-substrate-one-product mechanism, but rather operate with two substrates and two products, usually with the transfer of an active group. These types of reactions are sometimes known as Bi Bi reactions. There are two major classes of these reactions: the *sequential reactions*, in which all substrates bind with the enzyme before the reaction proceeds, and the *ping pong reactions*, in which one or more products are created and released before all of the substrates have been bound. In fact, unlike sequential reactions, the two substrates do not interact with one another while bound to the enzyme.

Optimal Conditions

The activity of enzymes is greatly influenced by both pH and temperature, as expected from the discussion of protein structure in the previous chapter. Activity profiles of most enzymes show a peak of activity that tails off on either side, whether it is pH or temperature. This is an innate characteristic of the enzyme. For example, pepsin, a digestive enzyme secreted into the stomach (pH 2) does not function when the pH > 5. On the other hand, another digestive enzyme, trypsin, which is secreted into the duodenum (proximal small intestine) where the pH is ~8, does not work in acidic environments. Changes in pH can change ionization of amino acid side chains that can thereby alter interaction with the substrate, or lead to changes in tertiary structure.

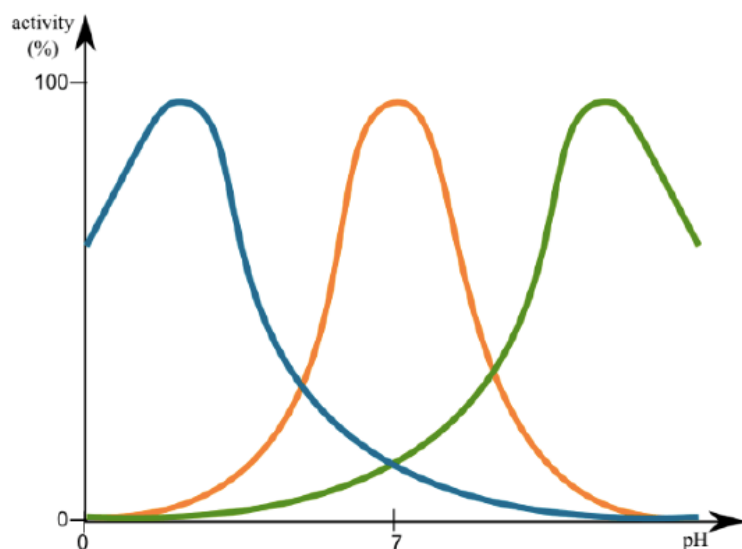


Figure 8.1.11. pH dependence of enzymatic activity. This

graph depicts three hypothetical enzymes with acidic, neutral, and basic pH optima.

Similarly, at suboptimal temperatures, the likelihood of protein-substrate interaction is low but above the optimal temperature, the increased energy can lead to breaking of hydrogen bonds within the structure of the enzyme, resulting in changes that inactivate the catalytic ability of enzyme or prevent it from binding substrate with sufficient affinity. The temperature optimum of most enzymes is very close to its typical environment. Thus, a human enzyme would operate optimally around 37°C, while an enzyme from bacteria that live in deep-sea volcanic vents (e.g. *Thermophilus aquaticus*) might have temperature optima over 90°C. This is one of the reasons that refrigeration can slow down growth of microorganisms (which obviously have no ability to regulate their temperature), and why most microorganisms are killed (enzymes permanently denatured) when put into sustained high temperature environments. Interestingly, the DNA polymerase from the *T. aquaticus* bacteria, also commonly called Taq polymerase, is used in a rapid DNA-amplifying lab technique known as PCR (polymerase chain reaction, see Methods chapter) in which samples are repeatedly heated to high temperatures to separate DNA strands in preparation for making copies of them. DNA polymerases from most prokaryotic or eukaryotic species would be denatured and inactivated by the high heat, but Taq has evolved (with respect to its tertiary structure) for extraordinary structural stability even in heat extremes.

Finally, many enzymes require a molecular partner that has no catalytic activity of its own, but like a catalyst, is not permanently altered by the chemical reaction. These molecules are *cofactors*. Some are simple: elemental, in fact, including metal ions such as Zn^{2+} or Ca^{2+} . Others are slightly more complex: small organic cofactors are called *coenzymes*, and accomplish the same thing, acting as a required partner to the enzyme in catalyzing a reaction. The interaction with the enzyme itself varies and may be only transient, as in $NAD^+/NADH$ which are coenzymes used in redox reactions, or permanently bound to the enzyme by covalent bond like the heme group of hemoglobin. Often the function of the coenzyme is to provide an active group to facilitate the catalyzed reaction. Coenzyme A, in various metabolic pathways such as glycolysis or the tricarboxylic acid cycle, can be bound to a substrate to form a stable product that then acts as an intermediate. The Co-A is released from the molecule as it undergoes the next step in a series of reactions in the metabolic pathway (see Chapter 5).

From a human health standpoint, it is interesting to note that many coenzymes are vitamins, or derived from vitamins. These are the B vitamins biotin (B_7), cobalamin (B_{12}), folic acid (B_9), niacin/nicotinamide (B_3), pantothenic acid (B_5), pyridoxine (B_6), riboflavin (B_2), and thiamine (B_1). Vitamins are small organic compounds that are not synthesized by an organism and must therefore be ingested. They are generally needed only in small quantities, but necessary nonetheless. Naturally, the vitamins we are familiar with are those required by humans. The specific roles of these vitamins and the consequences of not having enough of them are discussed later in this textbook, as the enzymes that they work with are introduced in detail.

This page titled [8.1: Regulation of Enzyme Activity](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [E. V. Wong](#).

- [8.4: Gas Laws](#) by Anonymous is licensed [CC BY-NC-SA 3.0](#). Original source: <https://2012books.lardbucket.org/books/introduction-to-chemistry-general-organic-and-biological>.

8.2: Control of Metabolism Through Enzyme Regulation

Learning Objectives

- Explain the effect of an enzyme on chemical equilibrium

Control of Metabolism Through Enzyme Regulation

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

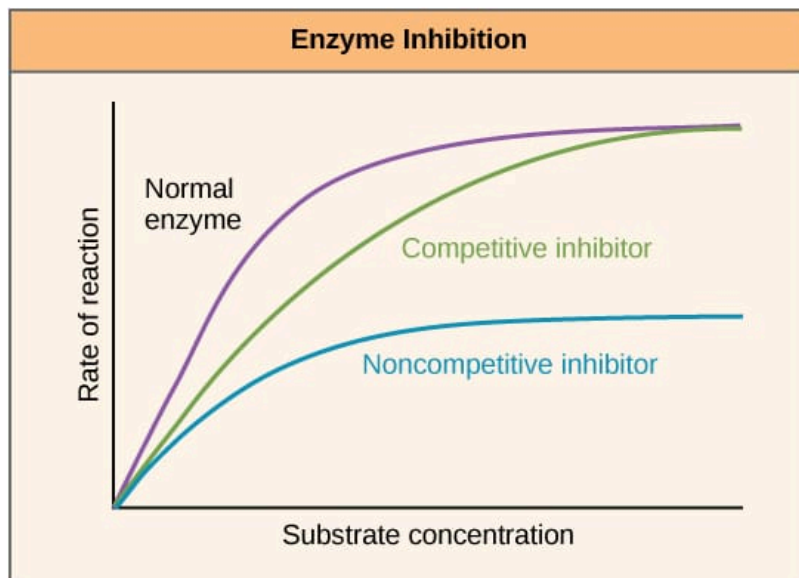


Figure: **Enzyme inhibition:** Competitive and noncompetitive inhibition affect the rate of reaction differently. Competitive inhibitors affect the initial rate, but do not affect the maximal rate, whereas noncompetitive inhibitors affect the maximal rate.

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

Competitive and Noncompetitive Inhibition

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

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

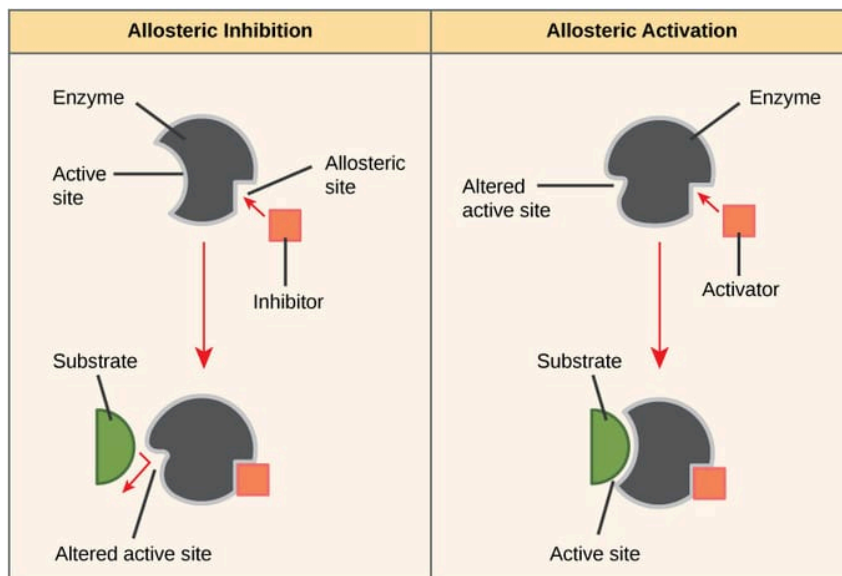


Figure: **Allosteric inhibitors and activators:**

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

Allosteric Inhibition and Activation

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

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

Cofactors and Coenzymes

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

Cofactors are inorganic ions such as iron (Fe^{2+}) and magnesium (Mg^{2+}). For example, DNA polymerase requires a zinc ion (Zn^{2+}) to build DNA molecules. Coenzymes are organic helper molecules with a basic atomic structure made up of carbon and hydrogen. The most common coenzymes are dietary vitamins. Vitamin C is a coenzyme for multiple enzymes that take part in building collagen, an important component of connective tissue. Pyruvate dehydrogenase is a complex of several enzymes that requires one cofactor and five different organic coenzymes to catalyze its chemical reaction. The availability of various cofactors and coenzymes regulates enzyme function.

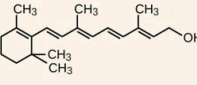
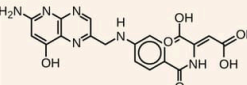
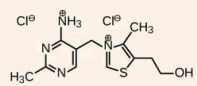
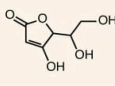
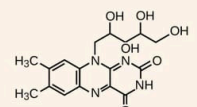
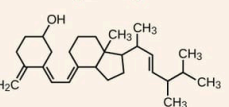
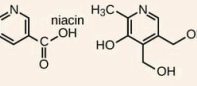
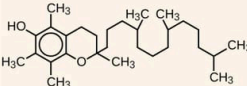
Dietary Vitamins	
Vitamin A 	Folic acid 
Vitamin B₁ 	Vitamin C 
Vitamin B₂ 	Vitamin D₂ (calciferol) 
Vitamin B₆ (pyridoxine) 	Vitamin E (alpha-tocopherol) 

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

Enzyme Compartmentalization

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

Feedback Inhibition in Metabolic Pathways

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

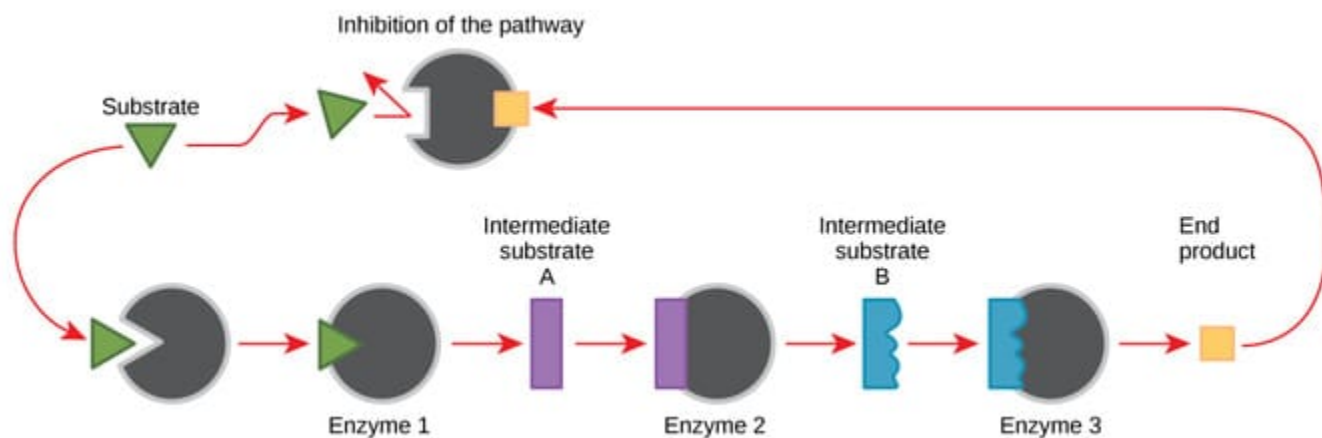


Figure: **Feedback inhibition:** Metabolic pathways are a series of reactions catalyzed by multiple enzymes. Feedback inhibition, where the end product of the pathway inhibits an earlier step, is an important regulatory mechanism in cells.

The production of both amino acids and nucleotides is controlled through feedback inhibition. For an example of feedback

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

Key Points

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

Key Terms

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

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

To ensure that you understand the material in this chapter, you should review the meanings of the bold terms in the following summary and ask yourself how they relate to the topics in the chapter.

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

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

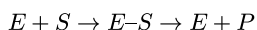
Protein molecules can have as many as four levels of structure. The **primary structure** is the sequence of amino acids in the chain. The **secondary structure** is the arrangement of adjacent atoms in the peptide chain; the most common arrangements are α -helices or β -pleated sheets. The **tertiary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

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

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

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

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



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

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

Substances that interfere with enzyme function are called inhibitors. An **irreversible inhibitor** inactivates enzymes by forming covalent bonds to the enzyme, while a **reversible inhibitor** inactivates an enzyme by a weaker, noncovalent interaction that is easier to disrupt. A **competitive inhibitor** is a reversible inhibitor that is structurally similar to the substrate and binds to the active site. When the inhibitor is bound, the substrate is blocked from the active site and no reaction occurs. Because the binding of such an inhibitor is reversible, a high substrate concentration will overcome the inhibition because it increases the

likelihood of the substrate binding. A **noncompetitive inhibitor** binds reversibly at a site distinct from the active site. Thus, it can bind to either the enzyme or the enzyme-substrate complex. The inhibitor changes the conformation of the active site so that the enzyme cannot function properly. Noncompetitive inhibitors are important in **feedback inhibition**, in which the amount of product produced by a series of reactions is carefully controlled. The final product in a series of reactions acts as a noncompetitive inhibitor of the initial enzyme.

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

This page titled [8.3: Amino Acids, Proteins, and Enzymes \(Summary\)](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [18.S: Amino Acids, Proteins, and Enzymes \(Summary\)](#) by Anonymous is licensed [CC BY-NC-SA 3.0](#). Original source: <https://2012books.lardbucket.org/books/introduction-to-chemistry-general-organic-and-biological>.

8.4: Exercises

This page titled [8.4: Exercises](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

9: CARBOHYDRATES

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

9.1: Carbohydrates Fundamentals

9.2: Monosaccharides

9.3: Disaccharides

9.4: Oligosaccharides

9.5: Polysaccharides

9.6: Glycosides

9.7: Exercises

9.1: Carbohydrates Fundamentals

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



-  [Topic hierarchy](#) 

[Template:HideTOC](#)

Thumbnail: [Haworth formula](#) of D-glucose.

9.2: Monosaccharides

The most useful carbohydrate classification scheme divides the carbohydrates into groups according to the number of individual simple sugar units. **Monosaccharides** contain a single unit; **disaccharides** contain two sugar units; and **polysaccharides** contain many sugar units as in polymers - most contain glucose as the monosaccharide unit.



-  [Topic hierarchy](#) 

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

9.3: Disaccharides

The most useful carbohydrate classification scheme divides the carbohydrates into groups according to the number of individual simple sugar units. **Monosaccharides** contain a single unit; **disaccharides** contain two sugar units; and **polysaccharides** contain many sugar units as in polymers - most contain glucose as the monosaccharide unit.



-  [Topic hierarchy](#) 

Thumbnail Ball-and-stick model of the α -lactose molecule. Image used with permission (Public Domain; Ben Mills).

9.4: Oligosaccharides

An oligosaccharide is a carbohydrate whose molecule, upon hydrolysis, yields two to ten [Monosaccharid](#) molecules. Oligosaccharides are classified into subclasses based on the number of monosaccharide molecules that form when one molecule of the oligosaccharide is hydrolyzed. Oligosaccharides can have many functions including cell recognition and cell binding. For example, glycolipids have an important role in the immune response.



-  [Topic hierarchy](#) 

Thumbnail: Structure of galactooligosaccharide. Image used with permission (Public Domain; [Klaas1978](#))

9.5: Polysaccharides

The most useful carbohydrate classification scheme divides the carbohydrates into groups according to the number of individual simple sugar units. **Monosaccharides** contain a single unit; **disaccharides** contain two sugar units; and **polysaccharides** contain many sugar units as in polymers - most contain glucose as the monosaccharide unit.



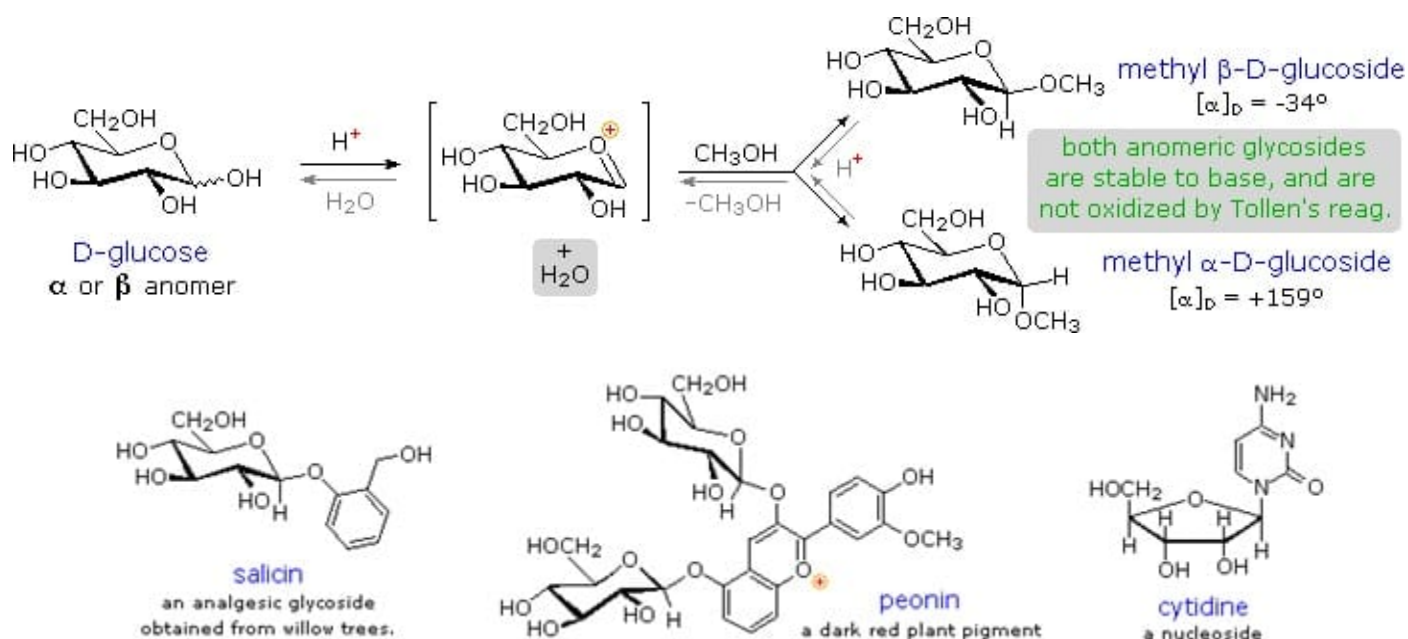
-  [Topic hierarchy](#) 

Thumbnail: Schematic two-dimensional cross-sectional view of glycogen: A core protein of glycogenin is surrounded by branches of [glucose](#) units. The entire globular granule may contain around 30,000 glucose units. By Mikael Häggström, used with permission (Public Domain).

9.6: Glycosides

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

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



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

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

Contributors

Prof. Steven Farmer (Sonoma State University)

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

9.7: Exercises

This page titled [9.7: Exercises](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

10: The Generation of Energy from Carbohydrate Metabolism

We will consider here the reverse process of photosynthesis, namely how carbohydrates, especially glucose, are converted to energy by being broken down into carbon dioxide and water.

A general summary of the several stages involved is shown in Figure 20-8. Initially, the storage fuels or foodstuffs (fats, carbohydrates, and proteins) are hydrolyzed into smaller components (fatty acids and glycerol, glucose and other simple sugars, and amino acids). In the next stage, these simple fuels are degraded further to two-carbon fragments that are delivered as the $\text{CH}_3\text{C}=\text{O}$ group (ethanoyl, or acetyl) in the form of the thioester of coenzyme A, CH_3COSCoA . The structure of this compound and the manner in which fatty acids are degraded has been considered in [Section 18-8F](#), and amino acid metabolism is discussed briefly in [Section 25-5C](#). This section is concerned mainly with the pathway by which glucose is metabolized by the process known as **glycolysis**.

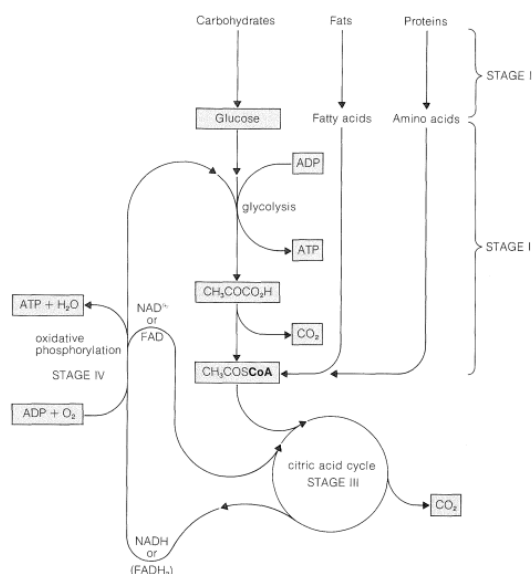
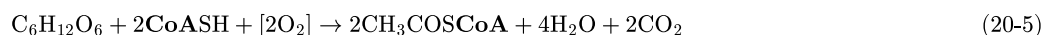


Figure 20-8: Perspective of the metabolic scheme whereby carbohydrates, fats, and proteins in foodstuffs are oxidized to CO_2 , showing the link between glycolysis, the citric acid cycle, and oxidative phosphorylation.

In the conversion of glucose to CH_3COSCoA , two carbons are oxidized to carbon dioxide with consumption of two oxygen molecules:



For further oxidation to occur, the CH_3COSCoA must enter the next stage of metabolism, whereby the $\text{CH}_3\text{C}=\text{O}$ group is converted to CO_2 and H_2O . This stage is known variously as the citric acid cycle, the tricarboxylic acid cycle, or the Krebs cycle, in honor of H. A. Krebs (Nobel Prize, 1953), who first recognized its cyclic nature in 1937. We can write an equation for the process as if it involved oxygen:



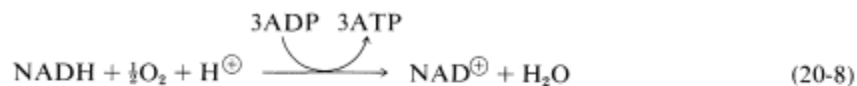
Notice that combination of the reactions of Equations 20-5 and 20-6, glycolysis plus the citric acid cycle, oxidizes glucose completely to CO_2 and H_2O :



But, as you will see, none of the steps uses molecular oxygen directly. Hence there must be a stage in metabolism whereby molecular oxygen is linked to production of oxidizing agents that are consumed in glycolysis and in the citric acid cycle.

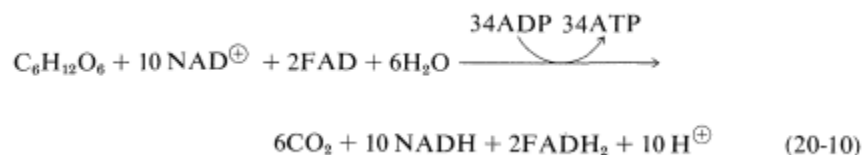
The coupling of oxygen into the metabolism of carbohydrates is an extremely complex process involving transport of the oxygen to the cells by an oxygen carrier such as hemoglobin, myoglobin, or hemocyanin. This is followed by a series of reactions, among

which NADH is converted to $\text{NAD}^{\oplus}$ with associated formation of three moles of ATP from three moles of ADP and inorganic phosphate. Another electron-carrier is flavin adenine dinucleotide (FAD; [Section 15-6C](#)), which is reduced to FADH_2 with an associated production of two moles of ATP from two moles of ADP. These processes are known as **oxidative phosphorylation** and can be expressed by the equations:



Oxidative phosphorylation resembles photophosphorylation, discussed in [Section 20-9](#), in that electron transport in photosynthesis also is coupled with ATP formation.

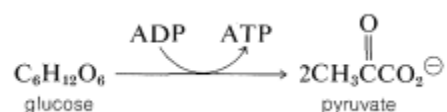
By suitably juggling Equations 20-7 through 20-9, we find that the metabolic oxidation of one mole of glucose is achieved by ten moles of $\text{NAD}^{\oplus}$ and two moles of FAD:



The overall result is production of 36 moles of ATP from ADP and phosphate per mole of glucose oxidized to CO_2 and H_2O . Of these, 34 ATPs are produced according to Equation 20-10 and, as we shall see, two more come from glycolysis.

Glycolysis

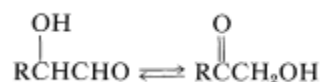
Glycolysis is the sequence of steps that converts glucose into two C_3 fragments with the production of ATP. The C_3 product of interest here is 2-oxopropanoate (pyruvate):



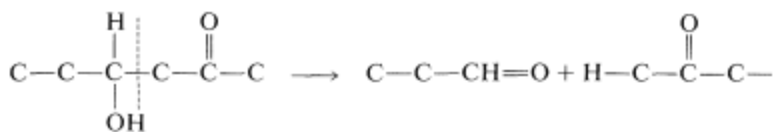
There are features in this conversion that closely resemble the dark reactions of photosynthesis, which build a C_6 chain (fructose) from C_3 chains ([Section 20-9](#)). For example, the reactants are either phosphate esters or mixed anhydrides, and the phosphorylating agent is ATP:



Furthermore, rearrangements occur that interconvert an aldose and ketose,



and the cleavage of a C_6 chain into two C_3 chains is achieved by a reverse aldol condensation:



Also, oxidation of an aldehyde to an acid is accomplished with NAD^+ . There is a related reaction in photosynthesis (Section 20-9) that accomplishes the reduction of an acid to an aldehyde and is specific for NADPH, not NADH:

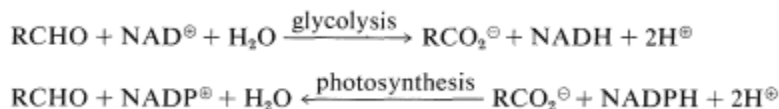


Figure 20-9 and each of the steps is identified more specifically in the ensuing discussion.

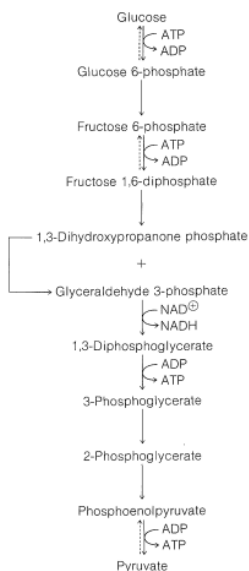
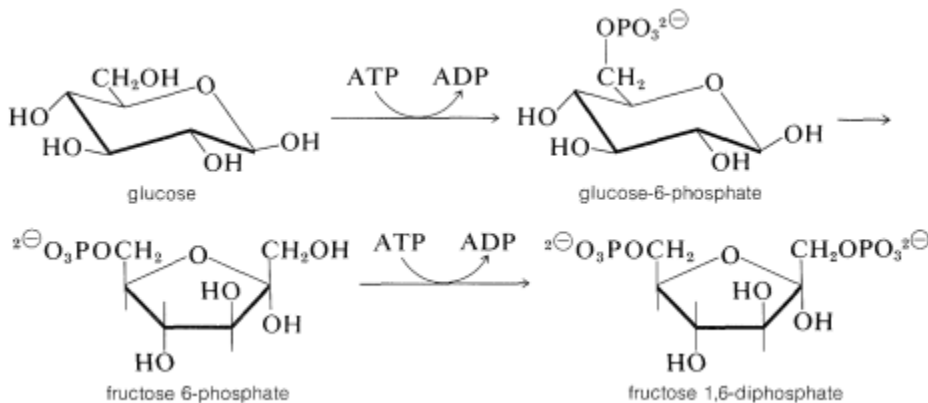
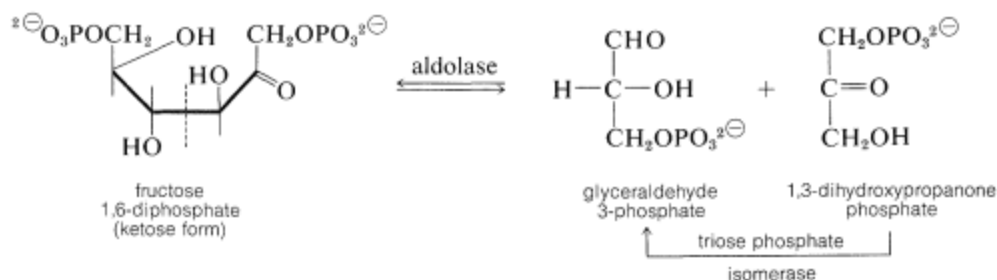


Figure 20-9: The glycolytic sequence. The dashed arrows in the reverse direction indicate the steps in the synthesis of glucose from pyruvate (gluconeogenesis) that differ from those in glycolysis.

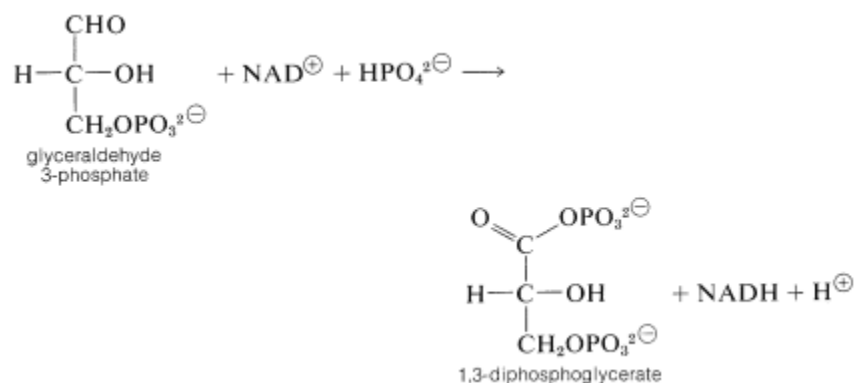
First, glucose is phosphorylated to glucose 6-phosphate with ATP. Then an aldose $\rightleftharpoons$ ketose rearrangement converts glucose 6-phosphate into fructose 6-phosphate. A second phosphorylation with ATP gives fructose 1,6-diphosphate:



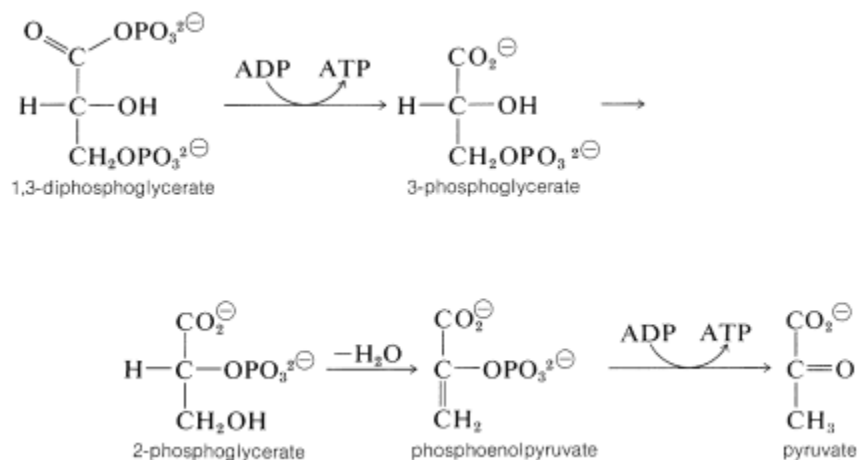
At this stage the enzyme aldolase catalyzes the aldol cleavage of fructose 1,6-diphosphate. One product is glyceraldehyde 3-phosphate and the other is 1,3-dihydroxypropanone phosphate. Another ketose $\rightleftharpoons$ aldose equilibrium converts the propanone into the glyceraldehyde derivative:



The next step oxidizes glyceraldehyde 3-phosphate with $\text{NAD}^{\oplus}$ in the presence of phosphate with the formation of 1,3-diphosphoglycerate:

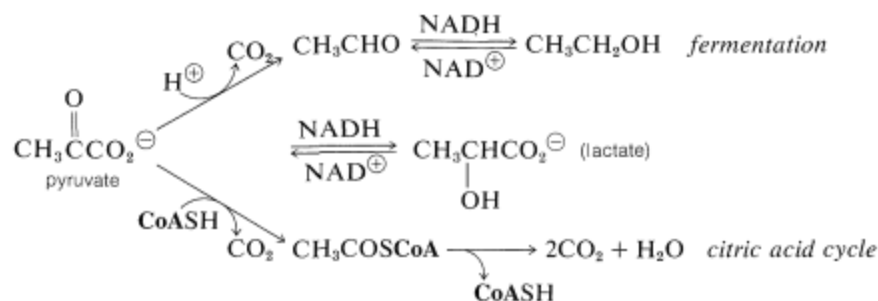


The mixed anhydride of phosphoric acid and glyceric acid then is used to convert ADP to ATP and form 3-phosphoglycerate. Thereafter the sequence differs from that in photosynthesis. The next few steps accomplish the formation of pyruvate by transfer of the phosphoryl group from C_3 to C_2 followed by dehydration to phosphoenolpyruvate. Phosphoenolpyruvate is an effective phosphorylating agent that converts ADP to ATP and forms pyruvate:



The net reaction at this point produces more ATP than is consumed in the phosphorylation of glucose and fructose.

What happens thereafter depends on the organism. With yeast and certain other microorganisms, pyruvate is decarboxylated and reduced to ethanol. The end result of glycolysis in this instance is *fermentation*. In higher organisms, pyruvate can be stored temporarily as a reduction product (lactate) or it can be oxidized further to give CH_3COSCoA and CO_2 . The CH_3COSCoA then enters the citric acid cycle to be oxidized to CO_2 and H_2O , as discussed in the next section:



The Citric Acid (Krebs) Cycle

Glycolysis to the pyruvate or lactate stage liberates heat, which can help keep the organism warm and produce ATP from ADP for future conversion into energy. However, glycolysis does not directly involve oxygen and does not liberate CO_2 , as we might expect from the overall process of the metabolic conversion of glucose to carbon dioxide and water (Equation 20-10). The liberation of CO_2 occurs subsequent to pyruvate formation in a process called variously, the citric acid cycle, the Krebs cycle, or the tricarboxylic acid (TCA) cycle.



Figure 20-10, and it will help to refer to this diagram as each of the steps in it are discussed in more detail.

The initial step, which is not really part of the cycle, is conversion of pyruvate to ethanoyl **CoA** (acetyl **CoA**):

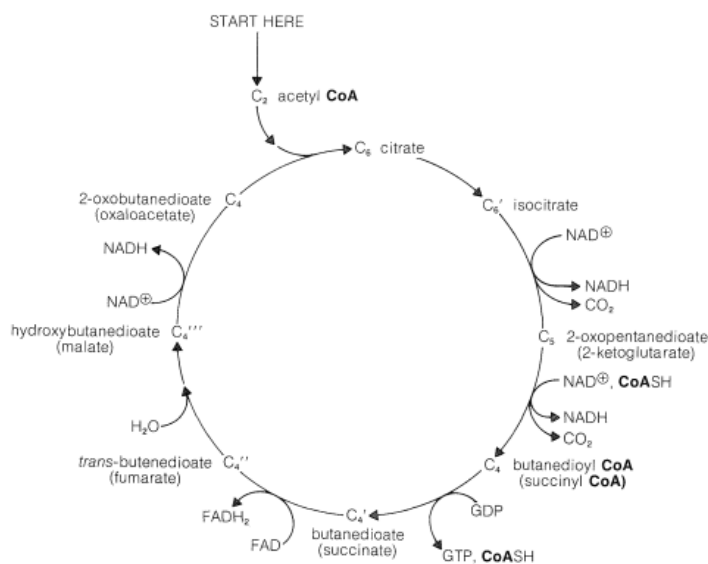
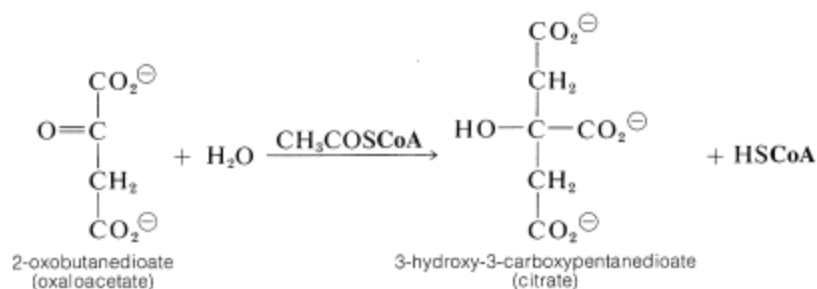


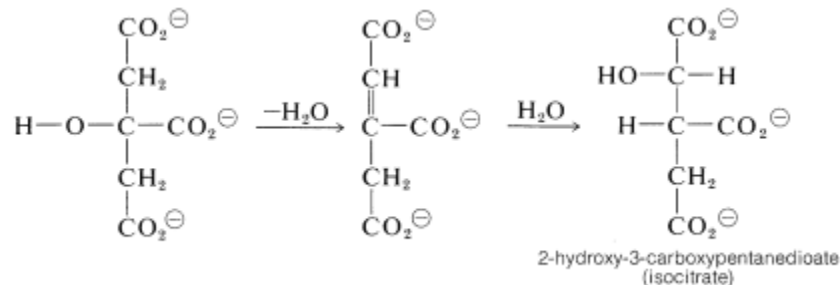
Figure 20-10: The citric acid cycle

To achieve the oxidation of acetyl **CoA** on a continuing basis, intermediates consumed in certain steps must be regenerated in others. Thus we have a situation similar to that in the Calvin cycle (Section 20-9), whereby the first stage of the cycle achieved the desired reaction (CO_2 formation) and the second stage is designed to regenerate intermediates necessary to perpetuate the cycle.

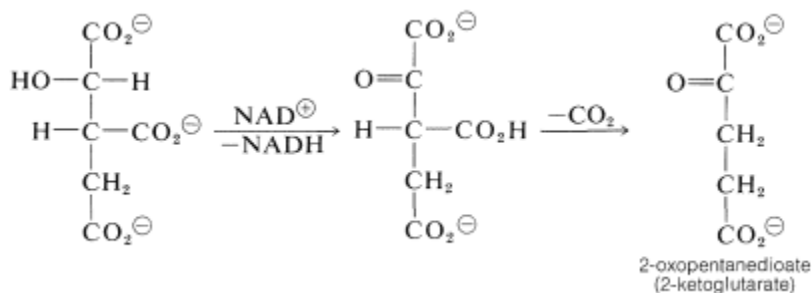
The entry point is the reaction between acetyl **CoA** and a four-carbon unit, 2-oxobutanedioic acid. An aldol-type addition of the CH_3CO group to this C_4 keto acid extends the chain to a branched C_6 acid (as citric acid):



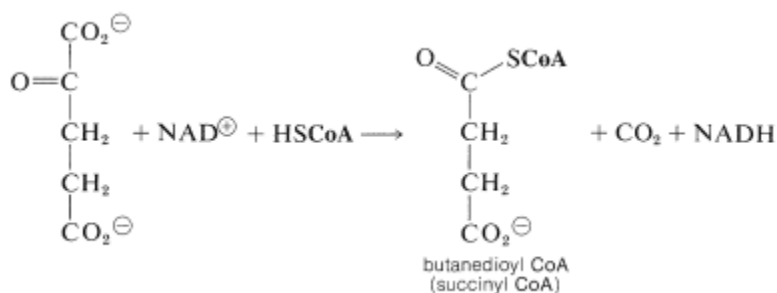
Dehydration-rehydration of citrate converts it to isocitrate:



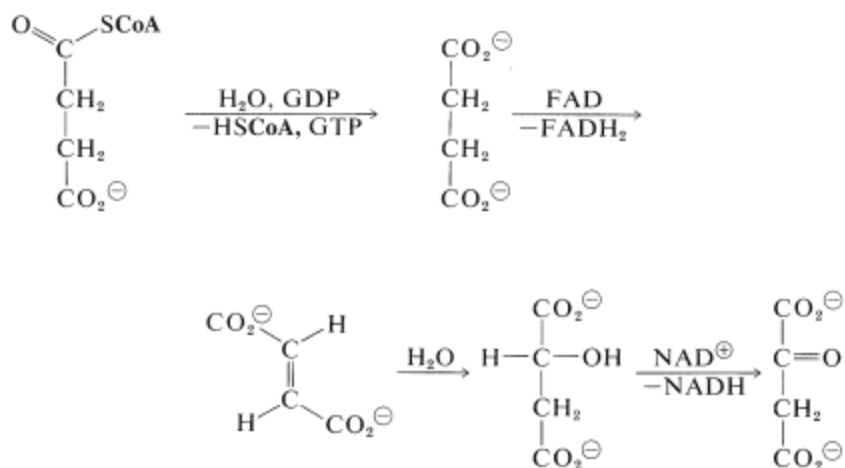
From here, oxidation of the hydroxyl function with NAD^+ gives a keto acid, which loses CO_2 readily (Section 18-4) and affords 2-oxopentanedioate:



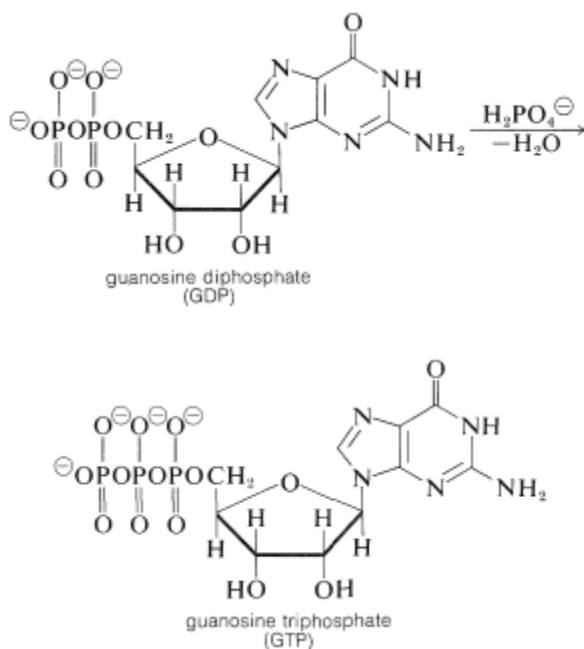
We now have a C_5 keto acid that can be oxidized in the same way as the C_3 keto acid, pyruvic acid, to give a butanedioyl **CoA**:



Two molecules of CO_2 now have been produced and the remaining part of the citric acid cycle is concerned with regeneration of the **CoA** for forming acetyl **CoA** from 2-oxopropionate, and also with regenerating the 2-oxobutanedioate, which is the precursor of citrate. The steps involved are



The hydrolysis of the acyl **CoA** in the first step is used for energy storage by conversion of guanosine diphosphate (GDP) to guanosine triphosphate (GTP):



The hydration of the *trans*-butenedioate ([Section 10-3G](#)) and the final oxidation reaction ([Section 15-6C](#)) have been discussed previously.

Alternative Routes in Carbohydrate Metabolism

There is an alternative route, called the [pentose phosphate pathway](#), by which glucose enters the glycolytic sequence to pyruvate. This route achieves the oxidative decarboxylation of glucose to give ribose, as the 5-phosphate ester. The essential steps are

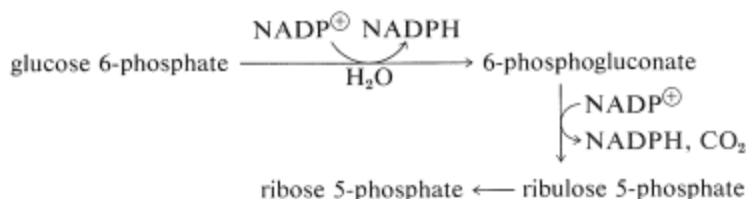
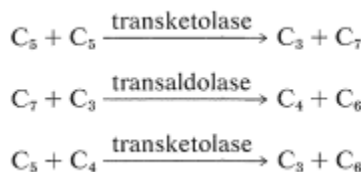


Figure 20-11). The reactions that accomplish this are very similar to those of the Calvin cycle (Section 20-9), only in reverse. They may be summarized as



The net result is that three pentoses are converted into two molecules of fructose and one of glyceraldehyde ($3\text{C}_5 \rightarrow 2\text{C}_6 + \text{C}_3$).

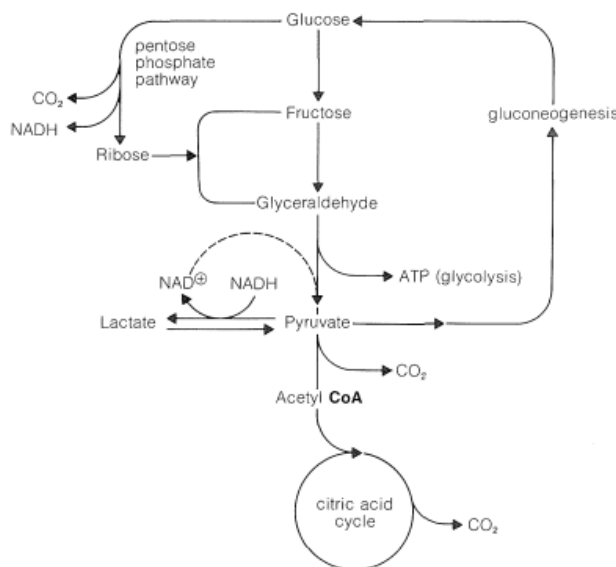


Figure 20-11: Schematic representation of metabolism of glucose by way of glycolysis and citric acid cycle, and the pentose phosphate, lactate, and gluconeogenesis links.

The relationship of the pentose-phosphate pathway to glycolysis is shown in Figure 20-11. The steps involved in the pentose shunt are readily reversible, but there are several steps in glycolysis that are not. These are the phosphorylation steps (Figure 20-9). Yet, there has to be a return route from pyruvate to glucose. This route is called **gluconeogenesis** and, in animals, takes place in the liver. We shall not discuss the steps in gluconeogenesis except to indicate again that they are not all the reverse of glycolysis. For comparison, the steps that differ are indicated in Figure 20-9 by dashed lines.

Why is lactate formed from pyruvate in the metabolism of glucose? $\text{Pyruvate} + \text{NADH} + \text{H}^+ \rightarrow \text{lactate} + \text{NAD}^+$ is a dead-end path, but it does furnish the NAD^+ needed for glycolysis in active muscle. This route for forming NAD^+ is important, because in circumstances of physical exertion, the rate of production of NAD^+ from oxidative phosphorylation may be slower than the demand for NAD^+ , in which case a temporary supply is available from the $\text{pyruvate} \rightarrow \text{lactate}$ reduction. The lactate so formed builds up in muscle tissue under conditions of physical exertion and is apt to cause muscles to "cramp". The excess lactate so formed ultimately is removed by being converted back to pyruvate by oxidation with NAD^+ .

The beauty of the metabolic cycle through pyruvate, shown in summary in Figure 20-11, is the way it can be tapped at various points according to whether the organism requires ATP (from glycolysis), NADH (from pentose shunt), or NAD^+ (from the lactate siding).

Contributors and Attributions

John D. Robert and Marjorie C. Caserio (1977) *Basic Principles of Organic Chemistry, second edition*. W. A. Benjamin, Inc. , Menlo Park, CA. ISBN 0-8053-8329-8. This content is copyrighted under the following conditions, "You are granted permission for individual, educational, research and non-commercial reproduction, distribution, display and performance of this work in any format."

This page titled [10: The Generation of Energy from Carbohydrate Metabolism](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [20.11: The Generation of Energy from Carbohydrate Metabolism](#) by John D. Roberts and Marjorie C. Caserio is licensed [CC BY-NC-SA 4.0](#).

11: OXIDATION

11.1: The Electron Transport Chain

11.2: ATP and Oxidative Phosphorylation

11.3: Exercises

Contributors

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

11.1: The Electron Transport Chain

Learning Outcomes

- Summarize the electron transport chain.
- Recognize that electron transport chain is the third and final stage of aerobic cellular respiration.
- Identify the products of the citric acid cycle.

What do trains, trucks, boats, and planes all have in common? They are ways to transport. And they all use a lot of energy. To make ATP, energy must be "transported" - first from glucose to NADH, and then somehow passed to ATP. How is this done? With an electron transport chain, the third stage of aerobic respiration. This third stage uses energy to make energy.

The Electron Transport Chain: ATP for Life in the Fast Lane

At the end of the Krebs Cycle, energy from the chemical bonds of glucose is stored in diverse energy carrier molecules: four ATPs, but also two FADH_2 and ten NADH molecules. The primary task of the last stage of cellular respiration, the **electron transport chain**, is to transfer energy from the electron carriers to even more ATP molecules, the "batteries" which power work within the cell.

Pathways for making ATP in stage 3 of aerobic respiration closely resemble the electron transport chains used in photosynthesis. In both electron transport chains, energy carrier molecules are arranged in sequence within a membrane so that energy-carrying electrons cascade from one to another, losing a little energy in each step. In both photosynthesis and aerobic respiration, the energy lost is harnessed to pump hydrogen ions into a compartment, creating an **electrochemical gradient** or **chemiosmotic gradient** across the enclosing membrane. And in both processes, the energy stored in the chemiosmotic gradient is used with **ATP synthase** to build ATP.

For aerobic respiration, the electron transport chain or "respiratory chain" is embedded in the inner membrane of the mitochondria (see figure below). The FADH_2 and NADH molecules produced in glycolysis and the Krebs Cycle, donate high-energy electrons to energy carrier molecules within the membrane. As they pass from one carrier to another, the energy they lose is used to pump hydrogen ions into the mitochondrial intermembrane space, creating an electrochemical gradient. Hydrogen ions flow "down" the gradient - from outer to inner compartment - through the ion channel/enzyme ATP synthase, which transfers their energy to ATP. Note the paradox that it requires energy to create and maintain a concentration gradient of hydrogen ions that are then used by ATP synthase to create stored energy (ATP). In broad terms, it takes energy to make energy. Coupling the electron transport chain to ATP synthesis with a hydrogen ion gradient is **chemiosmosis**, first described by Nobel laureate Peter D. Mitchell. This process, the use of energy to phosphorylate ADP and produce ATP is also known as **oxidative phosphorylation**.

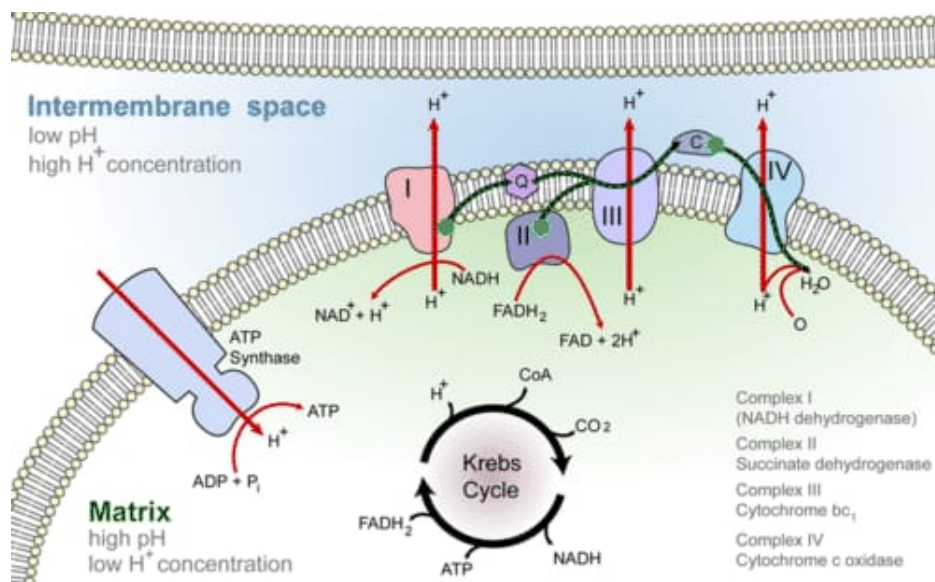
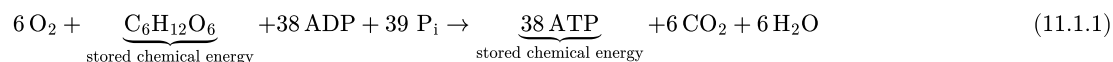


Figure 11.1.1: The third stage of cellular respiration uses the energy stored during the earlier stages in NADH and FADH₂ to make ATP. Electron transport chains embedded in the mitochondrial inner membrane capture high-energy electrons from the carrier molecules and use them to concentrate hydrogen ions in the intermembrane space. Hydrogen ions flow down their electrochemical gradient back into the matrix through ATP synthase channels which capture their energy to convert ADP to ATP. Notice that the process regenerates NAD⁺, supplying the electron acceptor molecule needed in glycolysis. (CC BY-NC 3.0; Mariana Ruiz Villarreal (LadyofHats) for the CK-12 Foundation).

After passing through the electron transport chain, low-energy electrons and low-energy hydrogen ions combine with oxygen to form water. Thus, oxygen's role is to drive the entire set of ATP-producing reactions within the mitochondrion by accepting "spent" hydrogens. Oxygen is the final electron acceptor, no part of the process - from the [Krebs Cycle](#) through the electron transport chain- can happen without oxygen.

The electron transport chain can convert the energy from one glucose molecule's worth of FADH₂ and NADH + H⁺ into as many as 34 ATP. When the four ATP produced in glycolysis and the Krebs Cycle are added, the total of 38 ATP fits the overall equation for aerobic cellular respiration:



Aerobic respiration is complete. If oxygen is available, cellular respiration transfers the energy from one molecule of glucose to 38 molecules of ATP, releasing carbon dioxide and water as waste. "Deliverable" food energy has become energy which can be used for work within the cell - transport within the cell, pumping ions and molecules across membranes, and building large organic molecules. Can you see how this could lead to "life in the fast lane" compared to anaerobic respiration (glycolysis alone)?

Contributors and Attributions

- Allison Soult, Ph.D. (Department of Chemistry, University of Kentucky)

11.2: ATP and Oxidative Phosphorylation

Learning Objectives

- explain reasons for the strongly exergonic hydrolysis of carboxylic acid anhydrides, phosphoric acid anhydrides, mixed anhydrides, and analogous structures and give approximate values for the ΔG° of hydrolysis of them;
- identify from Lewis structures molecules whose hydrolytic cleavage are strongly exergonic;
- explain how the exergonic cleavage of phosphoanhydride bonds in ATP can be coupled to the endergonic synthesis of macromolecules like proteins;
- draw mechanisms to show how oxidation and phosphorylation reactions are coupled in anaerobic metabolism through the productions of a mixed anhydride catalyzed by the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase;
- explain how arsenate can double oxidation and phosphorylation reactions in glycolysis
- explain how NAD^+ can be regenerated from NADH in anaerobic condition to allow glycolysis to continue;
- explain the general flow of electrons from NADH to dioxygen through a series of mobile and membrane protein bound electron acceptors in electron transport in the mitochondria inner member.
- explain with picture diagrams how oxidation and phosphorylation reactions (to produce ATP) are coupled in aerobic metabolism through the generation and collapse of a proton gradient in the mitochondria;
- draw pictures diagrams explaining the structure of $\text{F}_1\text{F}_0\text{ATPase}$ in the inner mitochondria member and explain using the picture how ATP synthesis is coupled to protein gradient collapse
- write an equation for the electrochemical potential and use it to calculate the available ΔG° for ATP production on proton gradient collapse, given typical values for ΔpH and ΔE across the membrane

Biological oxidation reactions serve two functions:

1. Oxidation of organic molecules can produce new molecules with different properties (e.g., an increase in solubility is observed on hydroxylation of aromatic substrates by cytochrome P450) and Likewise, amino acids can be oxidized to produce neurotransmitters.
2. Most biological oxidation reactions occur, however, to produce energy to drive thermodynamically unfavored biological processes such as protein and nucleic acid synthesis, or motility.

Chemical potential energy is not just released in biological oxidation reactions. Rather, it is transduced into a more useful form of chemical energy in the molecule ATP (adenosine triphosphate). This chapter will discuss the properties that make ATP so useful biologically, and how exergonic biological oxidation reactions are coupled to the synthesis of ATP.

[Template:HideTOC](#)

Contributors and Attributions

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

This page titled [11.2: ATP and Oxidative Phosphorylation](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Henry Jakubowski](#).

11.3: Exercises

This page titled [11.3: Exercises](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

12: The Flow of Genetic Information

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?

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, shown on the previous page, 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.

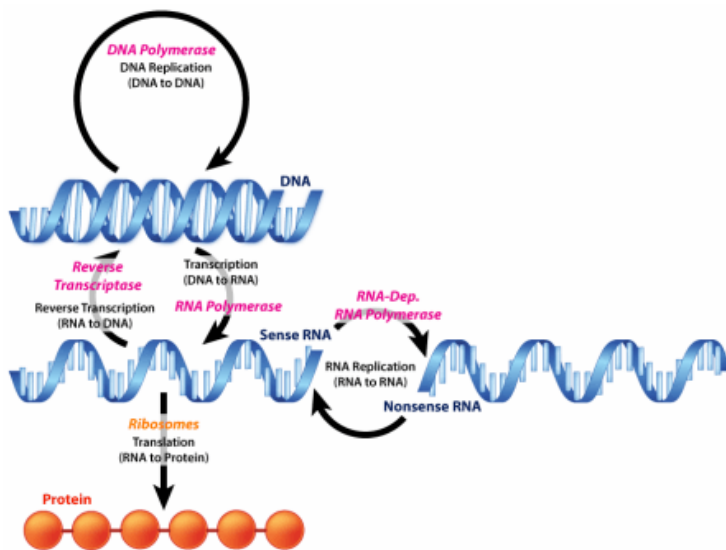


Figure 5.3.1: *Central dogma*

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.

What is 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.

General Features of Transcription

- The process of making RNA copies using a DNA as template is known as transcription
- The process of transcription produces all sorts of RNAs (mRNA, tRNA, rRNA, etc)
- One strand of the DNA serves as a template for the synthesis of RNA
- Enzymes that synthesize RNA are called RNA polymerases
- RNA polymerases synthesize RNA in the 5' to 3' direction
- RNA polymerases do not need a primer
- RNA polymerases use rNTPs (ATP, GTP, UTP and CTP) to build the new RNA strand
- RNA polymerases bind at specific DNA sequences called promoters to start transcription
- RNA polymerases stop RNA synthesis when they reach sequences called terminators

Figure 5.3.2: General Features of

Transcription

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 lifecycle. RNAs are, so to speak, temporary copies of instructions of the information in DNA and different sets of instructions are copied for use at different times.

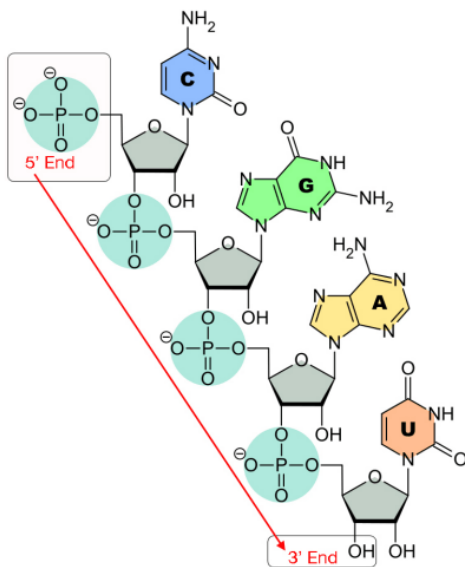


Figure 5.3.3: RNA structure

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
- Micro RNAs that regulate gene expression
- Other small RNAs that have a variety of functions.

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. What enzyme carries out transcription? Transcription is catalyzed by the enzyme RNA Polymerase. "RNA polymerase" is a general term for an enzyme that makes RNA. There are many different RNA polymerases.

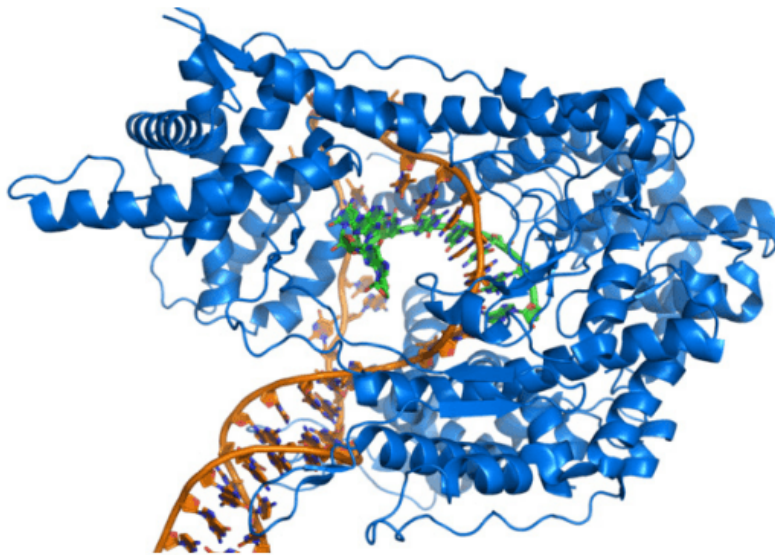


Figure 5.3.4: *T7 RNA Polymerase (blue) making RNA*

(green) using DNA template (brown)

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) rather than deoxyribonucleotides. Ribonucleotides are joined in exactly the same way as deoxyribonucleotides, which is to say that the 3'OH of the last nucleotide on the growing chain is joined to the 5' phosphate on the incoming nucleotide.

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 stringing together RNA nucleotides complementary to the DNA template.

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 genes into RNA at any given time.

Consider the challenge here: in a human cell, there are approximately 6 billion basepairs of DNA. Most of this is non- coding DNA, meaning that it won't 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? Signals in DNA indicate to RNA polymerase where it should start (and end) transcription. These signals are special sequences in DNA that 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.

A promoter is generally situated upstream of the gene that it controls. What this means is that on the DNA strand that the gene is on, the promoter sequence is "before" the gene. Remember that, by convention, DNA sequences are read from 5' to 3'. So the promoter lies 5' to the start point 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 the promoter, the gene cannot be transcribed and it will therefore, not be expressed.

Gene	-35 sequence	-10 sequence	Transcription start site
<i>rmE1</i>	CAATTTTTC	TATTGAGGAATG	AGGAGAACTCCC
<i>tRNA^{tr}</i>	CAACGTAAC	ACTTTACAGCGG	CGCGTCATTG
<i>trp</i>	AAATGAGCT	GTTGACAATTA	AATCATCGAACTAG
<i>araBAD</i>	GGATCCT	ACCTGACGCTTT	TTATCGCAACTCTC
<i>araC</i>	GCCGTGAT	TATAGACACTTT	TGTTACGCGTTTT
<i>bioA</i>	TTCCAAAACG	TGTTTTTTTGTG	TTAATTCGGTG

Figure 5.3.5: Promoter Sequences

What is special about a promoter sequence? In an effort to answer this question, scientists looked at many genes and their surrounding sequences. It makes sense that because the same RNA polymerase has to bind to many different promoters, the promoters should have some similarities in their sequences. Sure enough, 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 (Figure 5.3.5):

- A transcription start site (this the base in the DNA across from which the first RNA nucleotide is paired).
- 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, which by convention, is called the +1, the sequence found at -10 in the majority of promoters studied is TATAAT).
- A -35 sequence: this is a sequence at about 35 basepairs upstream from the start of transcription. The consensus sequence at this position is TTGACA.

What is the significance of these sequences? It turns out that the sequences at -10 and -35 are recognized and bound by a subunit of prokaryotic RNA polymerase before transcription can begin.

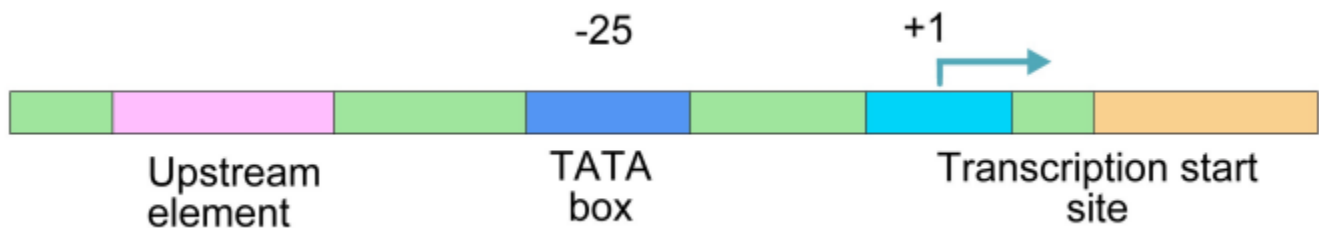


Figure 5.3.6: Promoter sequence elements

The RNA polymerase of *E. coli*, for example, has a subunit called the sigma subunit (or sigma factor) in addition to the core polymerase, which is the part of the enzyme that actually makes RNA. Together, the sigma subunit and core polymerase make up what is termed the RNA polymerase holoenzyme. The sigma subunit of the polymerase (shown in brown in Figure 5.3.7) can recognize and bind to the -10 and -35 sequences in the promoter, thus positioning the RNA polymerase (shown in green) at the right place to initiate transcription. Once transcription begins, the core polymerase and the sigma subunit separate, with the core polymerase continuing RNA synthesis and the sigma subunit wandering off to escort another core polymerase molecule to a promoter. The sigma subunit can be thought of as a sort of usher that leads the polymerase to its "seat" on the promoter.

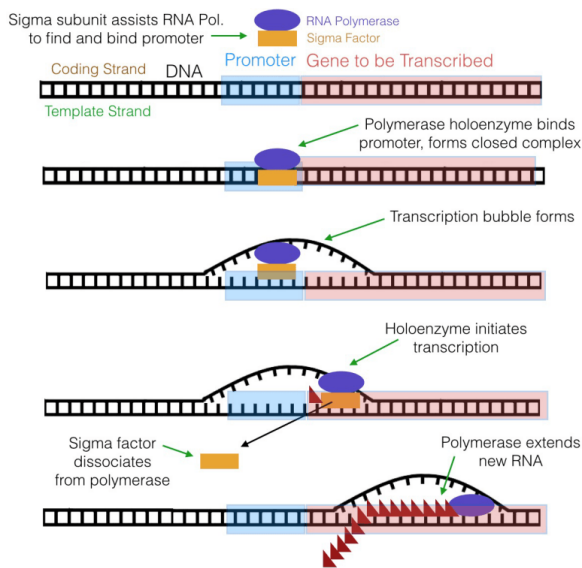


Figure 5.3.7: Transcription initiation in *E. coli*

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. How does the polymerase know where to stop? A sequence of nucleotides called the terminator is the signal to the RNA polymerase to stop transcription and dissociate from the template.

Although the process of RNA synthesis is the same in eukaryotes as in prokaryotes, there are some additional issues to keep in mind 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.

A second difference is that eukaryotes have multiple RNA polymerases, not one as in bacterial cells. The different polymerases transcribe different 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.

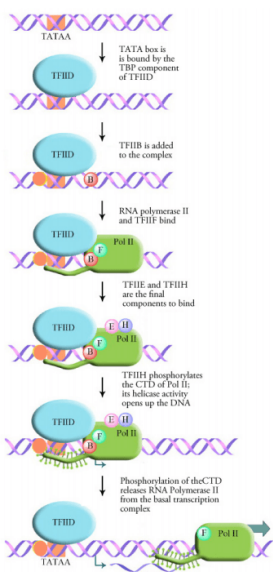


Figure 5.3.8: Assembly of basal transcription complex and initiation of transcription

All three eukaryotic RNA polymerases need additional proteins to help them get transcription started. In prokaryotes, RNA polymerase by itself can initiate transcription (remember that the sigma subunit is a subunit of the prokaryotic RNA polymerase). 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.

Finally, in eukaryotic cells, transcription is separated in space and time from translation. Transcription happens in the nucleus, and the mRNAs produced are processed further before they are sent into the cytoplasm. Protein synthesis (translation) happens in the cytoplasm. In prokaryotic cells, mRNAs can be translated as they are coming off the DNA template, and because there is no nucleus, transcription and protein synthesis occur in a single cellular compartment.

Like genes in prokaryotes, eukaryotic genes also have promoters. Eukaryotic promoters commonly have a TATA box, a sequence about 25 basepairs upstream of the start of transcription that is recognized and bound by proteins that help the RNA polymerase to position itself correctly to begin transcription. (Some eukaryotic promoters lack TATA boxes, and have, instead, other recognition sequences to help the RNA polymerase find the spot on the DNA where it binds and initiates transcription.)

We noted earlier that eukaryotic RNA polymerases need additional proteins to bind promoters and start transcription. What are these additional proteins that are needed to start transcription? General transcription factors are proteins that help eukaryotic RNA polymerases find transcription start sites and initiate RNA synthesis. We will focus on the transcription factors that assist RNA polymerase II. 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 in eukaryotes 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. 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 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. As shown in the figure at left, a number of different general transcription factors, together with RNA polymerase (Pol II) form a complex at the TATA box.

The final step in the assembly of the basal transcription complex is the binding of a general transcription factor called TFIIF. TFIIF 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 TFIIF adds a phosphate onto the C-terminal domain (CTD) of the RNA polymerase. This phosphorylation appears to be the signal that releases the RNA polymerase from the basal transcription complex and allows it to move forward and begin transcription.

12.1: The Structure of DNA

This page, looking at the structure of DNA, is the first in a sequence of pages leading on to how DNA replicates (makes copies of) itself, and then to how information stored in DNA is used to make protein molecules.

Contributors and Attributions

Jim Clark (Chemguide.co.uk)

This page titled [12.1: The Structure of DNA](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Jim Clark](#).

- [The Structure of DNA](#) by [Jim Clark](#) is licensed [CC BY-NC 4.0](#).

12.2: DNA Replication

According to the central dogma of molecular genetics, DNA is the genetically active component of the chromosomes of a cell. That is, DNA in the cell nucleus contains all the information necessary to control synthesis of the proteins, enzymes, and other molecules which are needed as that cell grows, carries on metabolism, and eventually reproduces. Thus when a cell divides, its DNA must pass on genetic information to both daughter cells. It must somehow be able to divide into duplicate copies. This process is called **replication**. Given the complementary double strands of DNA, it is relatively easy to see how DNA as a molecule is well structured for replication, as is shown in Figure 12.2.1. Each strand serves as a template for a new strand. Thus, after DNA is replicated, each new DNA double helix will have one strand from the original DNA molecule, and one newly synthesized molecule. This is referred to as semiconservative replication.

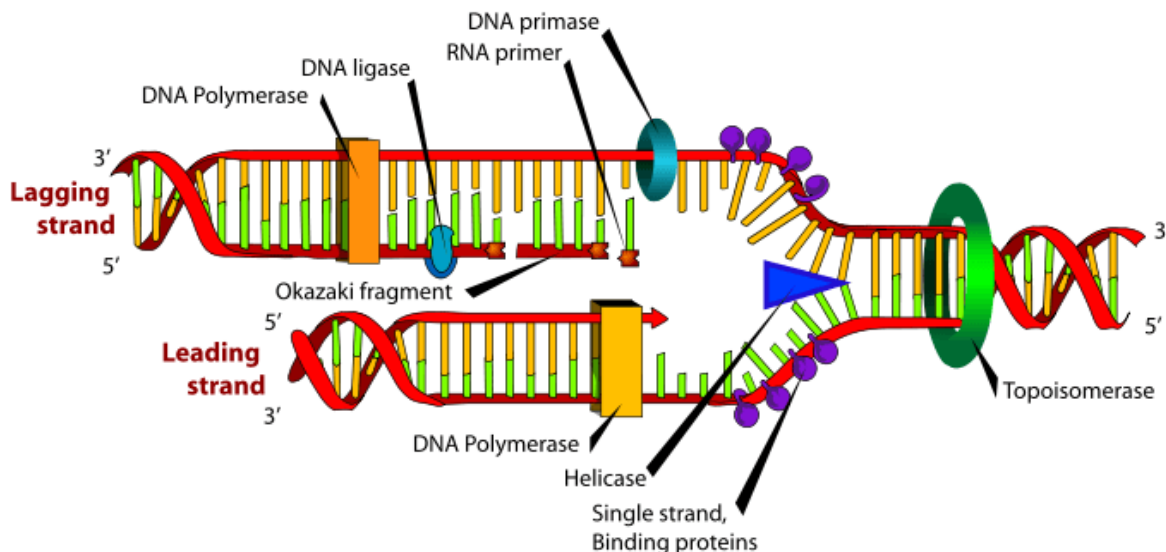


Figure 12.2.1: The replication of DNA. Replication occurs by means of partial unwinding of the two strands accompanied by synthesis of a new strand complementary to each of the originals.^[1]

A rather complex mechanism exists for DNA replication, involving many different enzymes and protein factors. Let us consider some of the more important aspects of DNA replication. First, the double strand needs to be opened up to replicate each template strand. To do this, a set of proteins and enzymes bind to and open up the double helix at an origin point in the molecule. This forms replication forks, points where double stranded DNA opens up, allowing replication to occur. A helicase enzyme binds at the replication forks, with the function of further unwinding the DNA and allowing the replication fork to move along the double strand as DNA is replicated. Another enzyme, DNA gyrase, is also required to relieve stress on the duplex caused by unwinding the double strand. Further, single strand binding proteins are needed to prevent the single strands from reforming a double strand. Another essential enzyme in this initiation phase is primase, which creates an RNA primer on each single strand of DNA to begin replication from.

All of these initial functions are necessary to prepare the DNA for the main enzyme which builds then new strands, DNA polymerase. Multiple polymerase enzymes exist, but for the moment we will DNA polymerase III, the main DNA polymerase in *E. coli*. DNA polymerase III catalyzes the reaction by which a new nucleotide is added to a growing DNA strand. That reaction is seen in Figure 12.2.2. The DNA polymerase enzymes need a free 3' OH group in order to begin synthesizing a new strand, which explains the necessity of the RNA primer, which gives a 3'OH group for DNA polymerase III to start from.^[2]

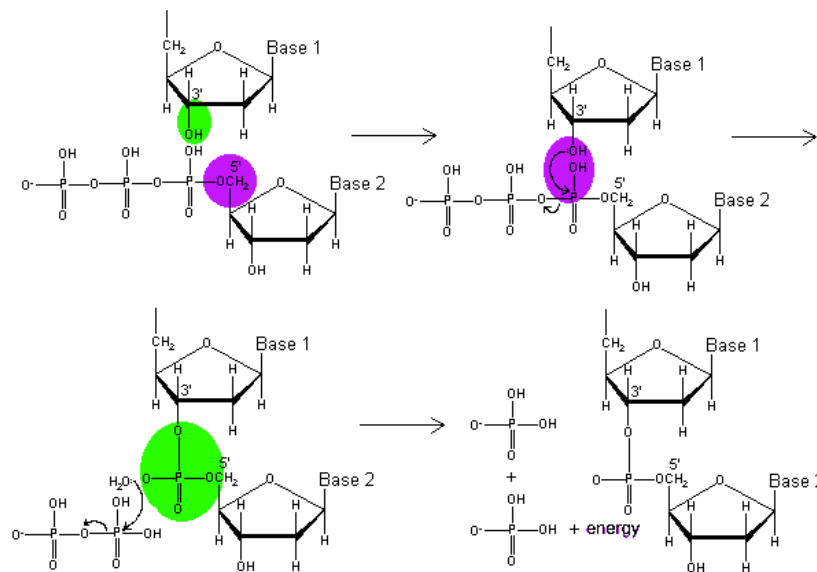


Figure 12.2.2 The polymerization of DNA nucleotides. The 3' hydroxyl group attacks the triphosphate group on the incoming nucleotide. A new phosphodiester bond is formed, and a pyrophosphate group leaves. This leaves two phosphates, the energy released from breaking down a the high energy phosphate group, and a elongated strand of DNA with a new 3' hydroxyl group to which another nucleotide may be added.

This leads to another constraint on DNA polymerase III. One strand, the **leading strand** can be polymerized continuously since the new strand being created goes 5' to 3' from the replication fork, but since the original strands are anti-parallel, the other strand, the **lagging strand** is going in the wrong direction for polymerization. In this case, the polymerization reaction starts away from the replication fork and works back toward it. This means that the lagging strand is synthesized in disconnected segments, known as Okazaki fragments, instead of continuously. Later, another DNA polymerase, in the case of *E. coli*, DNA polymerase I, removes RNA primers and fills in the missing discontinuities. Then, *another* enzyme, DNA ligase, connects breaks between 3'OH groups and 5' phosphate groups in the newly synthesized strands that exist due to these discontinuities. While the enzymes of this process differ in eukaryotes, they fulfill similar mechanisms. Even with this complexity of this process, DNA polymerase III is able to add new nucleotides at a rate of 250-1,000 nucleotides per second.^[3]

A number of advantages of the double-stranded structure held together by hydrogen bonds is evident in the process of replication. Complementary base pairing insures that the two new DNA molecules will be the same as the original. The large number of hydrogen bonds, each of which is relatively weak, makes complete separation of the two strands unlikely, but one hydrogen bond, or even a few, can be broken rather easily. The helicase portion of the replication complex can therefore separate the two strands in much the same way that a zipper operates. Like the teeth of a zipper, hydrogen bonds provide great strength when all work together, but the proper tool can separate them one at a time.

References

1. ↑ Mariana Ruiz. Wikimedia Commons. 24 January 2007. Retrieved: 11 August 2009.
2. ↑ Nelson, D.L., Cox, M.M. *Lehninger Principles of Biochemistry*(5thed). New York: W.H. Freeman and Company, 2008. pp. 985-991.
3. ↑ Nelson, D.L., Cox, M.M. *Lehninger Principles of Biochemistry*(5thed). New York: W.H. Freeman and Company, 2008. pp. 982-991.

This page titled 12.2: DNA Replication is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Ed Vitz, John W. Moore, Justin Shorb, Xavier Prat-Resina, Tim Wendorff, and Adam Hahn.

- 20.21: DNA Replication by Ed Vitz, John W. Moore, Justin Shorb, Xavier Prat-Resina, Tim Wendorff, and Adam Hahn is licensed CC BY-NC-SA 4.0.

12.2.1: Replication

The hereditary material in a cell is coded in the sequence of the heterocyclic amines of **DNA**. There are normally 46 strands of DNA called **chromosomes** in human cells. Specific regions, called **genes**, on each chromosome contain the hereditary information which distinguishes individuals from each other. The genes also contain the coded information required for the synthesis of proteins and enzymes needed for the normal functions of the cells. Bacterial cells may have 1000 genes, while the human cell contains more than a million genes. A single E. coli (bacteria) chromosome of double helical DNA consists of 3.4 million base pairs.

Introduction

Prior to cell division, the DNA material in the original cell must be duplicated so that after cell division, each new cell contains the full amount of DNA material. The process of DNA duplication is usually called **replication**. The replication is termed semiconservative since each new cell contains one strand of original DNA and one newly synthesized strand of DNA. The original polynucleotide strand of DNA serves as a template to guide the synthesis of the new complementary polynucleotide of DNA. A template is a guide that may be used for example, by a carpenter to cut intricate designs in wood.

The DNA single strand template serves to guide the synthesis of a complementary strand of DNA. DNA polymerase III is an example of this process. More explanation in the next panel.

DNA Replication Process

Several enzymes and proteins are involved with the replication of DNA. At a specific point, the double helix of DNA is caused to unwind possibly in response to an initial synthesis of a short RNA strand using the enzyme helicase. Proteins are available to hold the unwound DNA strands in position. Each strand of DNA then serves as a template to guide the synthesis of its complementary strand of DNA. DNA polymerase III is used to join the appropriate nucleotide units together. The replication process is shown in graphic on the left.

Template #1 guides the formation of a new complementary #2 strand. The DNA template guides the formation of a DNA complementary strand - not an exact copy of itself. For example looking at template # 2, this process occurs because the heterocyclic amine, adenine (A), codes or guides the incorporation of only thymine (T) to synthesize new DNA #1. The replication of DNA is guided by the base pairing principle so that no other heterocyclic amine nucleotide can hydrogen bond and fit correctly with cytosine. The next heterocyclic amine, cytosine (C), guides the incorporation of guanine (G) while similar arguments apply to the other bases. Exactly the opposite reaction occurs using template #2 (far right margin) where cytosine (C) guides the incorporation of guanine (G) to form a new complementary #2 strand.

It is so important that the cells duplicate the DNA genetic material exactly, that the sequence of newly synthesized nucleotides is checked by two different polymerase enzymes. The second enzyme can check for and actually correct any mistake of mismatched base pairs in the sequence. The mismatched nucleotides are hydrolyzed and cut out and new correct ones are inserted.

DNA Replication

Although details of DNA replication is not thoroughly understood, because so many molecules are involved in the process. This example focuses on the bacteriophage T7 DNA replication complex because it consists of relatively few proteins. The mechanism of T7 DNA replication is a good model for other DNA replication. This molecule is based on the recent work of Doublie, et al. (1998). In the graphic below, the DNA polymerase enzyme is shown with a short section of DNA. The green color represents the DNA template, while the magenta color represents the newly synthesized DNA.

In the close up, guanine triphosphate nucleotide is shown on the active site, guided by the cytosine in the template matching through hydrogen bonds. Only a few of the enzyme protein side chain interactions with the nucleotide are shown. Magnesium ions are also active in stabilizing the triphosphate through ionic interactions. Eventually the two of the phosphates are hydrolyzed and the remaining phosphate is bonded in a phosphate ester bond to the deoxyribose on the end of the newly forming DNA chain.

References

1. Doublet S., Tabor S., Long A., Richardson C., and Ellenberger T. (1998). Crystal Structure of a Bacteriophage T7 DNA Replication complex at 2.2 Å Resolution. Nature 391: 251-258.

Problems

1. In the green ribbon form of the enzyme, on the outside there are several series where the ribbons lie side by side. What protein structure is this?
2. On the inside of the "clamp", what protein structures are visible?
3. Using a single strand of DNA, write the sequence of heterocyclic amines to make its complementary strand.
Template: A C T A G G

Answers

1. These are beta pleated sheets in the tertiary structure.
2. There are numerous alpha helices that line the "hole" in the tertiary structure.
3. Complementary T G A T C C

Outside Links

- Link to: Great Animation of entire [DNA Replication](#) - John Kyrk
- More details: E. coli DNA Polymerase III Beta Subunit The Sliding DNA Clamp © David Marcey, 1997
- Reference: Kong, X-P., Onrust, R., O'Donnell, M., and J. Kuriyan (1992). Three-Dimensional Structure of the Beta Subunit of E. coli DNA Polymerase III Holoenzyme: A Sliding DNA Clamp. Cell 69: 425-437.
- More Details: The Bacteriophage T7 DNA Replication Complex, Michael E. Ward and David Marcey.

Contributors

- Charles Ophardt, Professor Emeritus, Elmhurst College; [Virtual Chembook](#)

This page titled [12.2.1: Replication](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [DNA: Replication](#) is licensed [CC BY-NC-SA 4.0](#).

12.3: RNA



-   Topic hierarchy  

12.3.1: Protein-RNA Recognition

12.3.2: RNA - Transcription

12.3.3: RNA Processing

12.3.4: Types of RNA

This page titled [12.3: RNA](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- RNA is licensed [CC BY-NC-SA 4.0](#).

12.3.1: Protein-RNA Recognition

RNA-protein interactions are behind a number of vital processes in the cell. Without the ability of particular proteins to bind RNA, the RNA would no longer be able to carry out its important functions as a component of the ribosome^{1,2} and spliceosome.³ Other examples of important RNA-protein interactions include binding of tRNA to aminoacyl-tRNA synthetases, a process vital to translation of genetic information into proteins necessary for continued biological function⁴ and regulation of post-transcriptional control of gene expression via the binding of RNA to ribonucleoproteins, or RNPs.⁵

Although not as well characterized as the binding between DNA and proteins, RNA-protein binding has been a field that has seen a great deal of growth in recent years. Although it was originally expected that RNA-protein binding motifs might fall neatly into categories the way DNA motifs did, the wide range of secondary and tertiary RNA structures that can be recognized by proteins requires more variety in binding motifs of the proteins, and the rules used to categorize them become correspondingly more complex.⁶ At this time all major families of RNA-binding proteins have been structurally characterized and these characterizations have led to a much better understanding of RNA recognition.⁷

Contributors and Attributions

- Therese Gerbich- Truman State University, Kirksville, MO
- Ashley Hoaglin- Truman State University, Kirksville, MO

This page titled [12.3.1: Protein-RNA Recognition](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

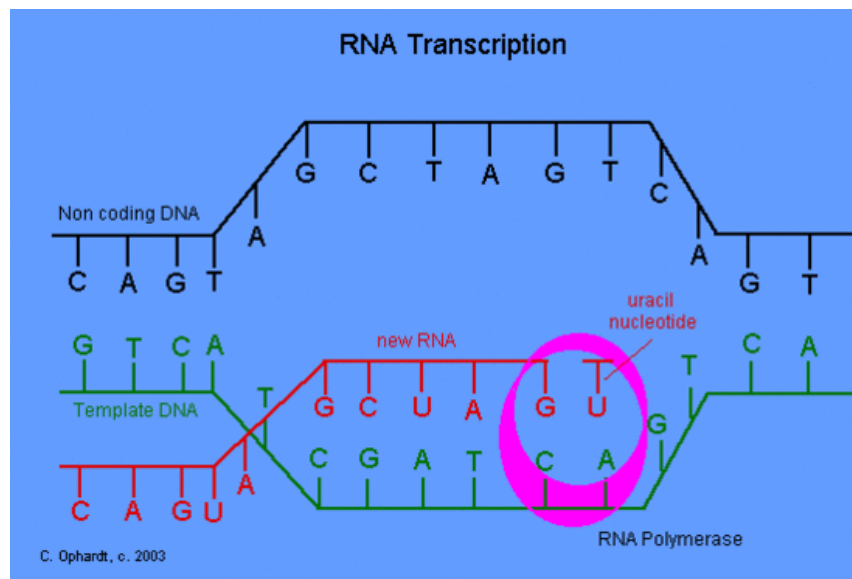
- [Protein-RNA Recognition](#) is licensed [CC BY-NC-SA 4.0](#).

12.3.2: RNA - Transcription

The biosynthesis of RNA, called **transcription**, proceeds in much the same fashion as the replication of DNA and also follows the base pairing principle. Again, a section of DNA double helix is uncoiled and only one of the DNA strands serves as a template for RNA polymerase enzyme to guide the synthesis of RNA. After the synthesis is complete, the RNA separates from the DNA and the DNA recoils into its helix.

Introduction

The differences in the composition of RNA and DNA have already been noted. In addition, RNA is not usually found as a double helix but as a single strand. However, the single polynucleotide strand may fold back on itself to form portions which have a double helix structure like the tertiary structure of proteins.



The transcription of a single RNA strand is illustrated in the graphic on the left. One major difference is that the heterocyclic amine, adenine, on DNA codes for the incorporation of uracil in RNA rather than thymine as in DNA. Remember that thymine is not found in RNA and do not confuse the replacement of uracil in RNA for thymine in DNA in the transcription process. For example, thymine in DNA still codes for adenine on RNA not uracil, while the adenine on DNA codes for uracil in RNA. Note that the new RNA (red) is identical to non coding DNA with the exception of uracil where thymine was located in DNA.

There are three major types of RNA which will be fully explained in a later section. Although RNA is synthesized in the nucleus, it migrates out of the nucleus into the cytoplasm where it is used in the synthesis of proteins.

RNA Transcription Process

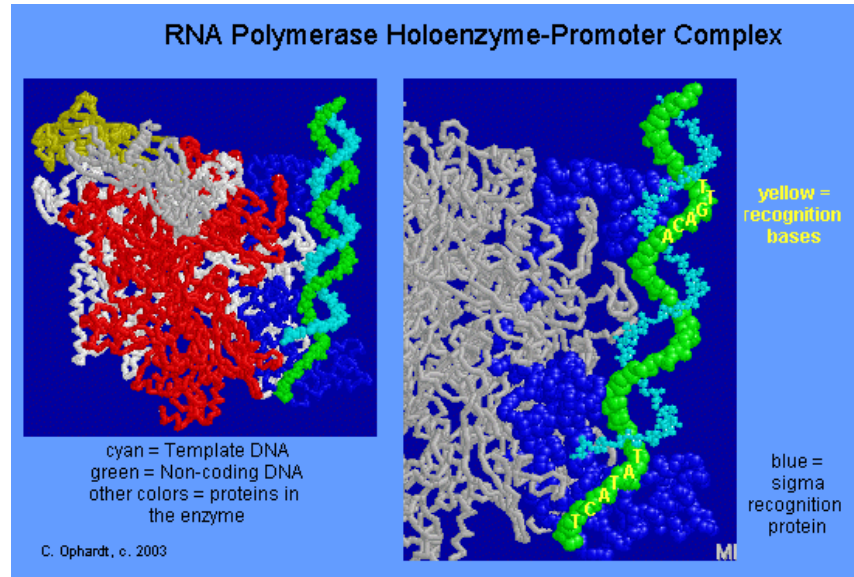
The RNA transcription process occurs in three stages: initiation, chain elongation, and termination.

The first stage occurs when the **RNA Polymerase-Promoter Complex** binds to the promoter gene in the DNA. This also allows for the finding of the start sequence for the RNA polymerase. The promoter enzyme will not work unless the sigma protein is present (shown in blue in graphic). Specific sequences on the non coding strand of DNA are recognized as the signal to start the unwinding process.

The recognition sequences are as follows:

Non-coding DNA -5' recognition sections in bold
 GGCCGCT**TTGAC**AAAAGTGTTAAATTGTGCT**TATACT**

Once the process has been initiated, then the RNA polymerase elongation enzyme takes over and is described in the next panel.



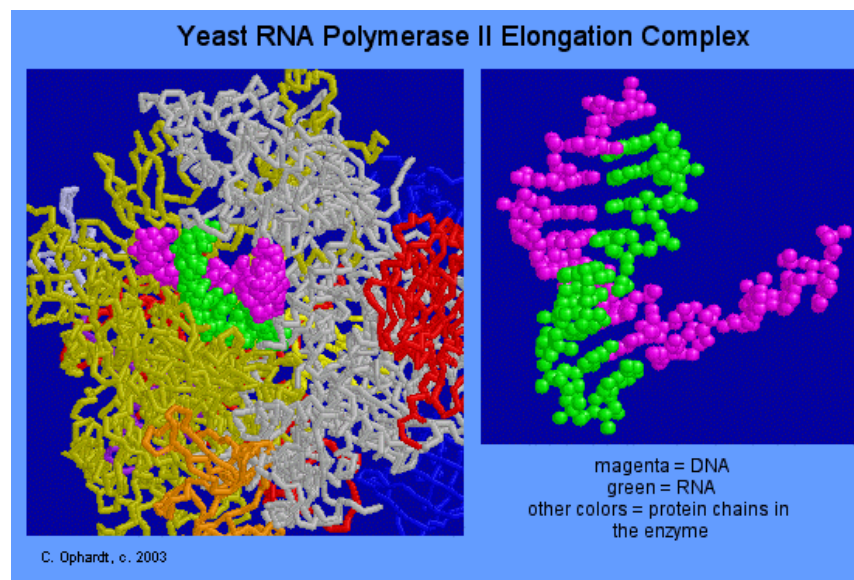
RNA Polymerase - Elongation

The elongation begins when the RNA polymerase "reads" the template DNA. Only one strand of the DNA is read for the base sequence. The RNA which is synthesized is the complementary strand of the DNA. The RNA (top strand) and DNA (bottom strand) sequences in the model are:

5' -GACCAGGCA-3'

3'-TCTGGTCCGTA-5'

In the graphic, the magenta color is the template DNA, while the green is the RNA strand. In the next reaction step, uracil triphosphate (UTP) is the next to be added to the RNA by bind and pairing with the adenine (A) nucleotide on the template DNA strand.



A phosphodiester bond is formed; the RNA chain is then elongated to 10 nucleotides; and diphosphate left over would dissociate. Note: The coordinates used in this display have only the alpha carbons of the proteins RNA Polymerase.

Outside Links

- Link with more details - RNA Polymerase - Promoter Complex
- Credit: Murakami, K. S., Masuda, S., Campbell, E. A., Muzzin, O., Darst, S. A. (2002) "Structural Basis of Transcription Initiation: An RNA Polymerase Holoenzyme-DNA Complex" Science 296:1285.
- More details: Yeast RNA Polymerase II Elongation Complex
- Credit: Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., Kornberg, R. D. (2001) "Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 Å Resolution" Science 292:1876.

Contributors

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

This page titled [12.3.2: RNA - Transcription](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [RNA - Transcription](#) is licensed [CC BY-NC-SA 4.0](#).

12.3.3: RNA Processing

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 (the product of transcription is known as a transcript) is further processed before it is functional. Both prokaryotes and eukaryotes process their ribosomal and transfer RNAs.

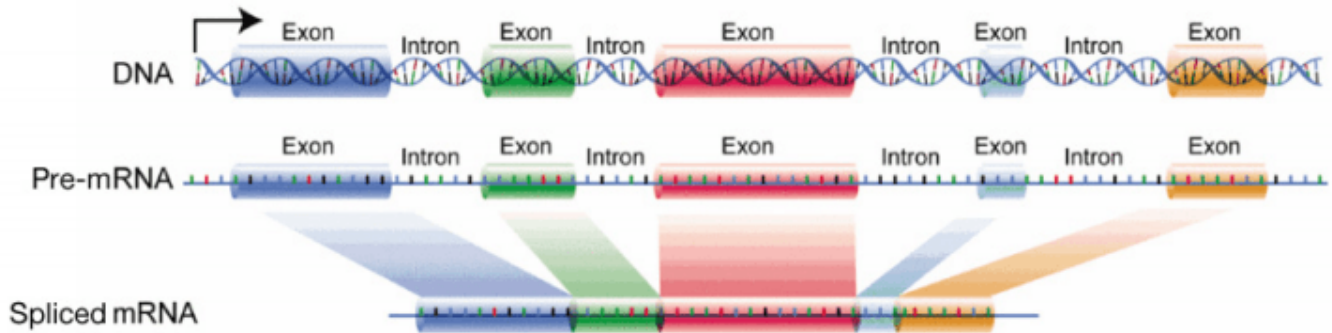


Figure 5.5.1: *RNA splicing*

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 discussion. 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. In addition, eukaryotic genes have introns, noncoding regions that interrupt the gene's coding sequence. The mRNA copied from genes containing introns will also therefore have regions that interrupt the information in the gene. These regions must be removed before the mRNA is sent out of the nucleus to be used to direct protein synthesis. The process of removing the introns and rejoining the coding sections or exons, of the mRNA, is called splicing. Once the mRNA has been capped, spliced and had a polyA tail added, it is sent from the nucleus into the cytoplasm for translation.

The initial product of transcription of a protein coding gene is called the pre-mRNA (or primary transcript). After it has been processed and is ready to be exported from the nucleus, it is called the mature mRNA or processed mRNA.

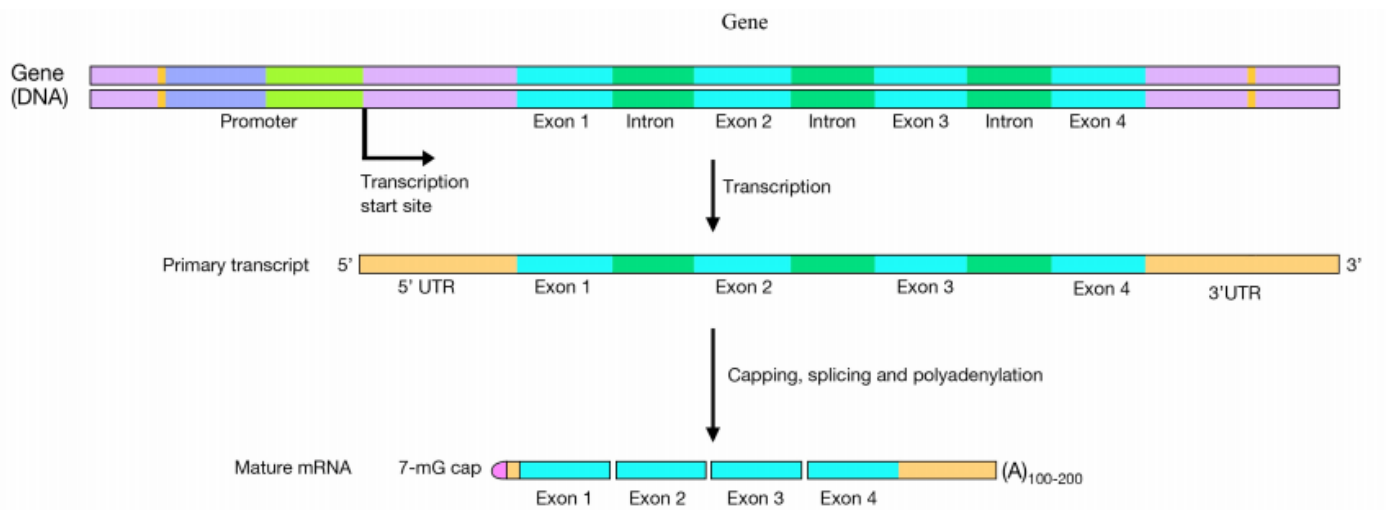


Figure 5.5.2: *Steps in processing eukaryotic messenger RNAs*

What are the processing steps for messenger RNAs?

In eukaryotic cells, pre-mRNAs undergo three main processing steps:

- Capping at the 5' end

- Addition of a polyA tail at the 3' end. and
- Splicing to remove introns

In the capping step of mRNA processing, a 7-methyl guanosine (shown at left) is added at the 5' end of the mRNA. 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.

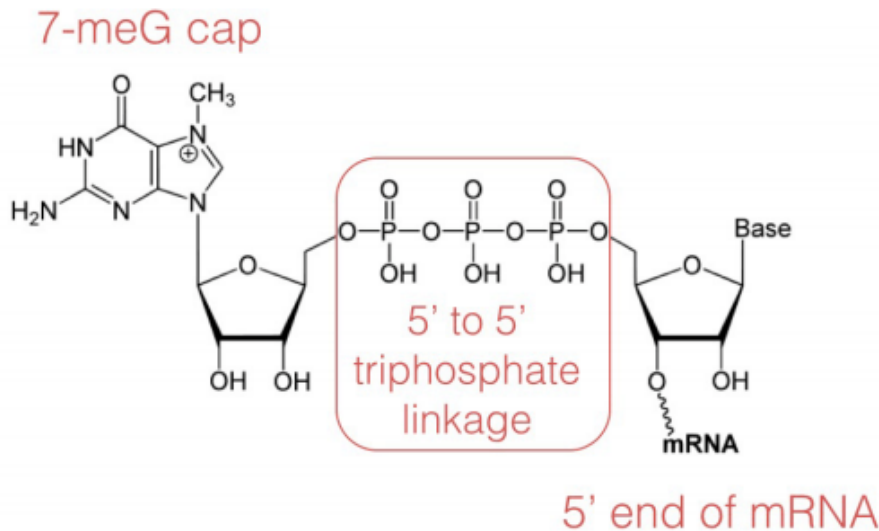


Figure 5.5.3: mRNA capping structure

The 3' end of a eukaryotic mRNA is first trimmed, then an enzyme called PolyA Polymerase adds a "tail" of about 200 'A' nucleotides to the 3' end. 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. The cap and the polyA tail on an mRNA are also indications that the mRNA is complete (i.e., not defective). 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). The splicing machinery must be able to recognize splice junctions (i.e., the end of each exon and the start of the next) in order to correctly cut out the introns and join the exons to make the mature, spliced mRNA.

What signals indicate where an intron starts and ends? The base sequence at the start (5' or left end, also called the donor site) of an intron is GU while the sequence at the 3' or right end (a.k.a. acceptor site) is AG. There is also a third important sequence within the intron, called a branch point, that is important for splicing.

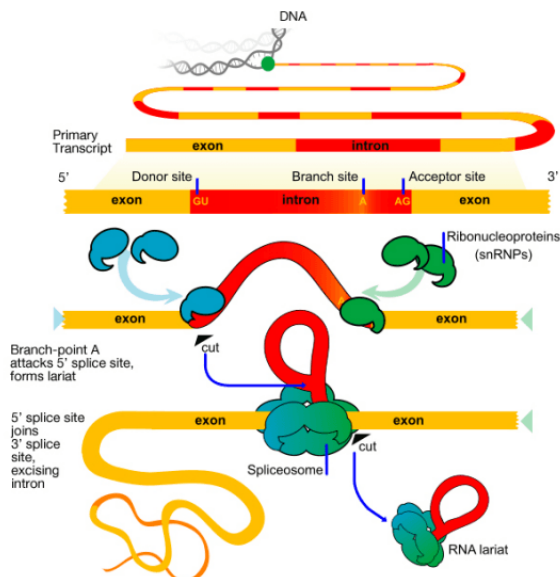


Figure 5.5.4: Splicing

There are two main steps in splicing:

- In the first step, the pre-mRNA is cut at the 5' splice site (the junction of the 5' exon and the intron). The 5' end of the intron then is joined to the branch point within the intron. This generates the lariat-shaped molecule characteristic of the splicing process
- In the second step, the 3' splice site is cut, and the two exons are joined together, and the intron is released.

Many pre-mRNAs have a large number of exons that 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 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. 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.

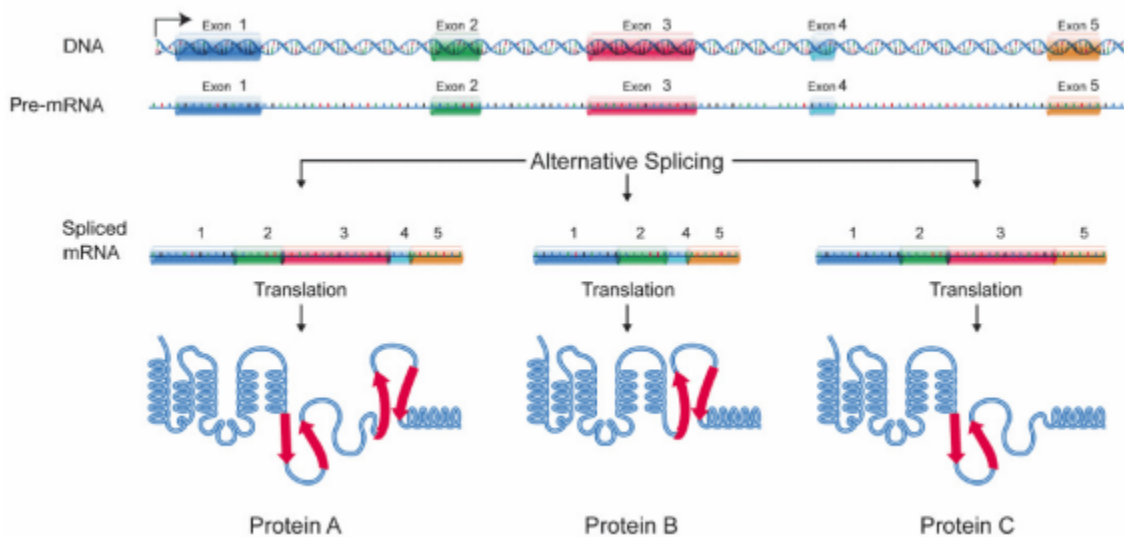


Figure 5.5.5:

Splicing and protein diversity

This page titled [12.3.3: RNA Processing](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Kevin Ahern](#) and [Indira Rajagopal](#).

- [Uncertainty Principle](#) is licensed [CC BY-NC-SA 4.0](#).

12.3.4: Types of RNA

Three general types of RNA exist: messenger, ribosomal, and transfer.

Messenger RNA

Messenger RNA (mRNA) is synthesized from a gene segment of DNA which ultimately contains the information on the primary sequence of amino acids in a protein to be synthesized. The genetic code as translated is for m-RNA not DNA. The **messenger RNA carries the code** into the cytoplasm where protein synthesis occurs.

Genetic Code

Each gene (or distinct segment) on DNA contains instructions for making one specific protein with order of amino acids coded by the precise sequence of heterocyclic amines on the nucleotides. Since proteins have a variety of functions including those of enzymes mistakes in the primary sequence of amino acids in proteins may have lethal effects.

How can a polymeric nucleotide with only four different heterocyclic amines specify the sequence of 20 or more different amino acids? If each nucleotide coded for a single amino acid, then obviously only 4 of the 20 amino acids could be accommodated. If the nucleotides were used in groups of two, there are 16 different combinations possible which is still inadequate.

It has been determined that the genetic code is actually based upon **triplets** of nucleotides which provide 64 different codes using the 4 nucleotides. During the 1960's, a tremendous effort was devoted to proving that the code was read as triplets, and also to solving the genetic code. The genetic code was originally translated for the bacteria E. Coli, but its universality has since been established. The genetic code is "read" from a type of RNA called **messenger RNA (mRNA)**. **Each nucleotide triplet, called a codon**, can be "read" and translated into an amino acid to be incorporated into a protein being synthesized. The genetic code is shown in Figure 7.

Several distinctive features of the genetic code are clearly evident. First, all of the 64 codons or triplets have a known function, with 61 coding for amino acids and the other 3 serving as a stop or termination signal for protein synthesis. Secondly, the code is degenerate, meaning that there are usually several codons for each amino acid. Only methionine and tryptophan have a single codon. More specifics on the importance of the degeneracy of the genetic code will be discussed in a later section.

Ribosomal RNA

In the cytoplasm, **ribosomal RNA (rRNA)** and protein combine to form a nucleoprotein called a ribosome. **The ribosome serves as the site and carries the enzymes necessary for protein synthesis.** In the graphic on the left, the ribosome is shown as made from two sub units, 50S and 30S. There are about equal parts rRNA and protein. The far left graphic shows the complete ribosome with three tRNA attached.

The ribosome attaches itself to m-RNA and provides the stabilizing structure to hold all substances in position as the protein is synthesized. Several ribosomes may be attached to a single RNA at any time. In upper right corner is the 30S sub unit with mRNA and tRNA attached.

Note: The coordinates used in this display have only the alpha carbons of the proteins (CA) and the DNA backbone atoms.

Transfer RNA

Transfer RNA (tRNA) contains about 75 nucleotides, three of which are called anticodons, and one amino acid. **The tRNA reads the code and carries the amino acid to be incorporated into the developing protein.**

There are at least 20 different tRNA's - one for each amino acid. The basic structure of a tRNA is shown in the left graphic. Part of the tRNA doubles back upon itself to form several double helical sections. On one end, the amino acid, phenylalanine, is attached. On the opposite end, a specific base triplet, called the **anticodon**, is used to actually "read" the codons on the mRNA.

The tRNA "reads" the mRNA codon by using its own anticodon. The actual "reading" is done by matching the base pairs through hydrogen bonding following the base pairing principle. Each codon is "read" by various tRNA's until the appropriate match of the anticodon with the codon occurs. In this example, the tRNA anticodon (AAG) reads the codon (UUC) on the mRNA. The UUC codon codes for phenylalanine which is attached to the tRNA. Remember that the codons read from the mRNA make up the genetic code as read by humans.

Problems

Quiz: Read and translate the codons on mRNA into the appropriate amino acids.
G U A C G A A A

Answer ▼

Quiz: Read and translate the codon on mRNA into the appropriate amino acid.
AGA. What is the anticodon on the appropriate tRNA?

Answer ▼

Outside Links

- Charles Ophardt (Professor Emeritus, Elmhurst College); [Virtual Chembook](#)

This page titled [12.3.4: Types of RNA](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.

- [Types of RNA](#) is licensed [CC BY-NC-SA 4.0](#).

12.4: Regulation of Transcription

The processes described above are required whenever any gene is transcribed. But what determines which genes are transcribed at a given time. What are the molecular switches that turn transcription on or off? Although there are entire books written on this one topic, the basic mechanism by which transcription is regulated depends on highly specific interactions between transcription regulating proteins and regulatory sequences on DNA.

We know that promoters indicate where transcription begins, but what determines that a given gene will be transcribed? In addition to the promoter sequences required for transcription initiation, genes have additional regulatory sequences (sequences of DNA on the same DNA molecule as the gene) that control when a gene is transcribed. Regulatory sequences are bound tightly and specifically by transcriptional regulators, proteins that can recognize DNA sequences and bind to them. The binding of such proteins to the DNA can regulate transcription by preventing or increasing transcription from a particular promoter.

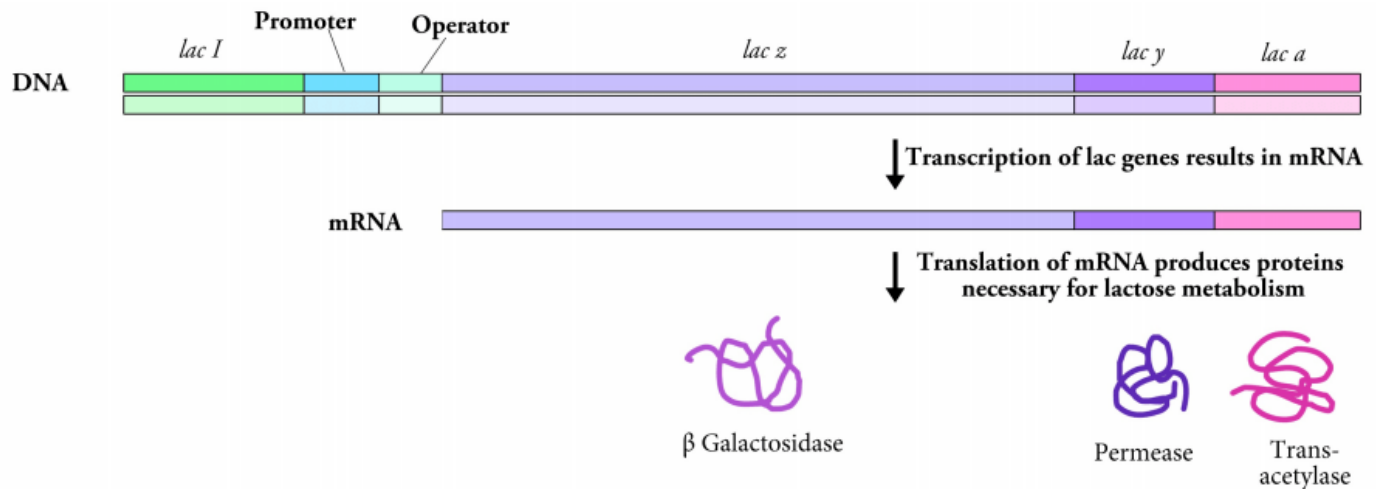


Figure 5.4.1: The genes *lac z*, *lac y*, and *lac a* are all under the control of a single promoter in the *lac* operon

Regulation in Prokaryotes

Let us first consider an example 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. The *lac* operon, shown in Figure 5.4.2, is one such group of genes that encode proteins needed for the uptake and breakdown of the sugar lactose. The three genes of the *lac* operon, *lac z*, *lac y* and *lac a* are controlled by a single promoter.

Bacterial cells generally prefer to use glucose for their energy needs, but if glucose is unavailable, and lactose is present, the bacteria will take up lactose and break it down for energy. Since the proteins for taking up and breaking down lactose are only needed when glucose is absent and lactose is available, the bacterial cells need a way to express the genes of the *lac* operon only under those conditions. At times when lactose is absent, the cells do not need to express these genes.

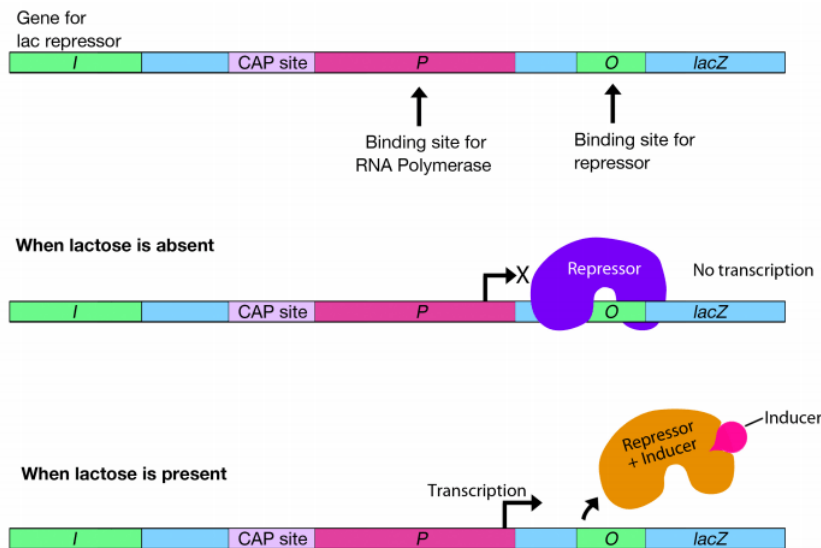


Figure 5.4.2: *Lac operon regulation*

How do bacteria achieve this? Transcription of the lac cluster of genes is primarily controlled by a repressor protein that binds to a region of the DNA just downstream of the -10 sequence of the lac promoter. Recall that the promoter is where the RNA polymerase must bind to begin transcription. The place where the repressor is bound is called the operator (labeled O in the figure). When the repressor is bound at this position, it physically blocks the RNA polymerase from transcribing the genes, 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. How is the repressor removed?

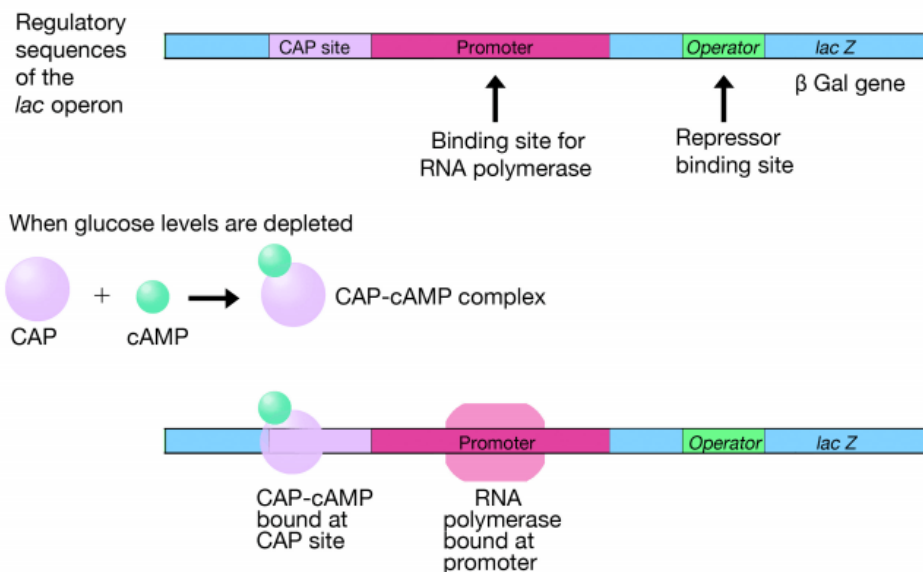


Figure 5.4.3: *Lac operon activation*

When the sugar lactose is present, it binds to the repressor, changing its conformation so that it no longer binds to the operator. When the repressor is no longer bound at the operator, the "road-block" in front of the RNA polymerase is removed, permitting the transcription of the genes of the lac operon.

Because the binding of the lactose induces the expression of the genes in the lac operon, lactose is called an inducer. (Technically, the inducer is allolactose, a molecule made from lactose by the cell, but the principle is the same.)

What makes this an especially effective control system is that the genes of the lac operon encode proteins that break down lactose. Turning on these genes requires lactose to be present. Once the lactose is broken down, the repressor binds to the operator once more and the lac genes are no longer expressed. This allows the genes to be expressed only when they are needed.

But how do glucose levels affect the expression of the lac genes? We noted earlier that if glucose was present, lactose would not be used. A second level of control is exerted by a protein called CAP that binds to a site adjacent to the promoter and recruits RNA polymerase to bind the lac promoter. When glucose is depleted, there is an increase in levels of cAMP which binds to CAP. The CAP cAMP complex then binds the CAP site, as shown in Figure 5.4.3. The combination of CAP binding and the lac repressor dissociating from the operator when lactose levels are high ensures high levels of transcription of the lac operon just when it is most needed. The CAP protein binding may be thought of as a green light for the RNA polymerase, while the removal of repressor is like the lifting of a barricade in front of it. When both conditions are met, the RNA polymerase transcribes the downstream genes.

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. An example of this is the trp operon in bacterial cells, which encodes enzymes necessary for the synthesis of the amino acid tryptophan. These genes are expressed at all times, except when tryptophan is available from the cell's surroundings. This means that these genes must be prevented from being expressed in the presence of tryptophan. This is achieved by having a repressor protein that will bind to the operator only in the presence of tryptophan.

Regulation in Eukaryotes

Transcription in eukaryotes is also regulated by the binding of proteins to specific DNA sequences, but with some differences from the simple schemes outlined above. For most eukaryotic genes, general transcription factors and RNA polymerase (i.e., the basal transcription complex) are necessary, but not sufficient, for high levels of transcription.

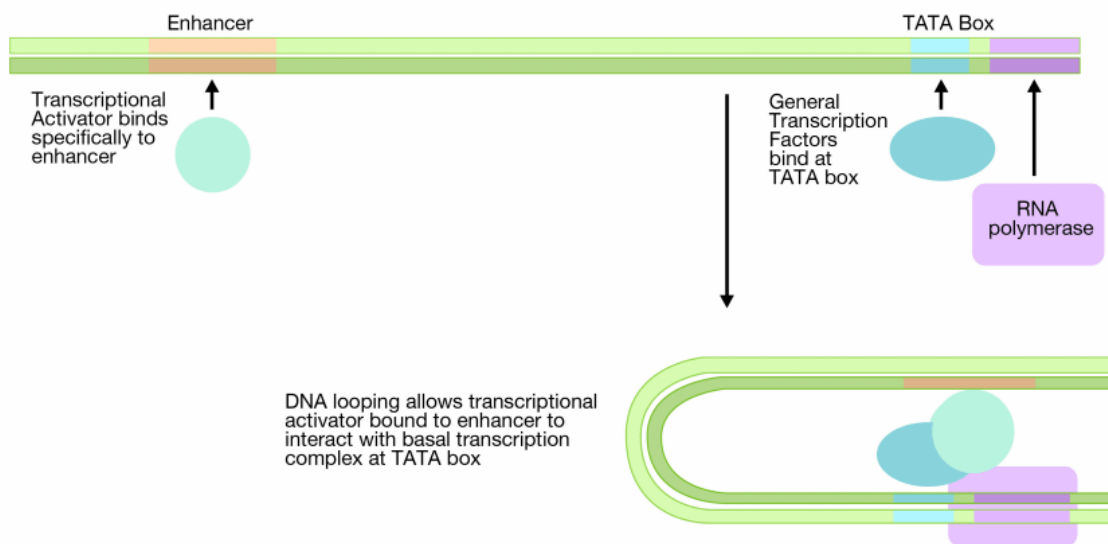


Figure 5.4.4:

Enhancer Mechanism

In eukaryotes, additional regulatory sequences called enhancers and the proteins that bind to the enhancers are needed to achieve high levels of transcription. Enhancers are DNA sequences that regulate the transcription of genes. Unlike prokaryotic regulatory sequences, enhancers don't need to be next to the gene they control. Often they are many kilobases away on the DNA. As the name suggests, enhancers can enhance (increase) transcription of a particular gene.

How can a DNA sequence far from the gene being transcribed affect the level of its transcription?

Enhancers work by binding proteins (transcriptional activators) that can, in turn, interact with the proteins bound at the promoter. The enhancer region of the DNA, with its associated transcriptional activator(s) can come in contact with the basal transcription complex that is bound at a distant TATA box by looping of the DNA (previous page). This allows the protein bound at the enhancer to make contact with the proteins in the basal transcription complex.

One way that the transcriptional activator bound to the enhancer increases the transcription from a distant promoter is that it increases the frequency and efficiency with which the basal transcription complex is formed at the promoter.

Another mechanism by which proteins bound at the enhancer can affect transcription is by recruiting to the promoter other proteins that can modify the structure of the chromatin in that region. As we noted earlier, in eukaryotes, DNA is packaged with proteins to form chromatin. When the DNA is tightly associated with these proteins, it is difficult to access for transcription. So proteins that can make the DNA more accessible to the transcription machinery can also play a role in the extent to which transcription occurs.

In addition to enhancers, there are also negative regulatory sequences called silencers. Such regulatory sequences bind to transcriptional repressor proteins. Transcriptional activators and repressors are modular proteins- they have a part that binds DNA and a part that activates or represses transcription by interacting with the basal transcription complex.

This page titled [12.4: Regulation of Transcription](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by Kevin Ahern and Indira Rajagopal.

- [Quantum Theory](#) is licensed [CC BY-NC-SA 4.0](#).

12.5: The Genetic Code

This page looks at how the base sequences in DNA and RNA are used to code for particular amino acids when it comes to building protein chains.

Contributors and Attributions

Jim Clark (Chemguide.co.uk)

This page titled [12.5: The Genetic Code](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Jim Clark](#).

- [The Genetic Code](#) by [Jim Clark](#) is licensed [CC BY-NC 4.0](#).

12.6: Translation

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 RNAs (rRNAs) and proteins. The proteins and rRNAs are organized into two subunits, a large and a small. The large subunit has an enzymatic activity, known as a peptidyl transferase, that makes the peptide bonds that join amino acids to make a polypeptide. The small and large subunits assemble on the mRNA at its 5' end to initiate translation. 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.

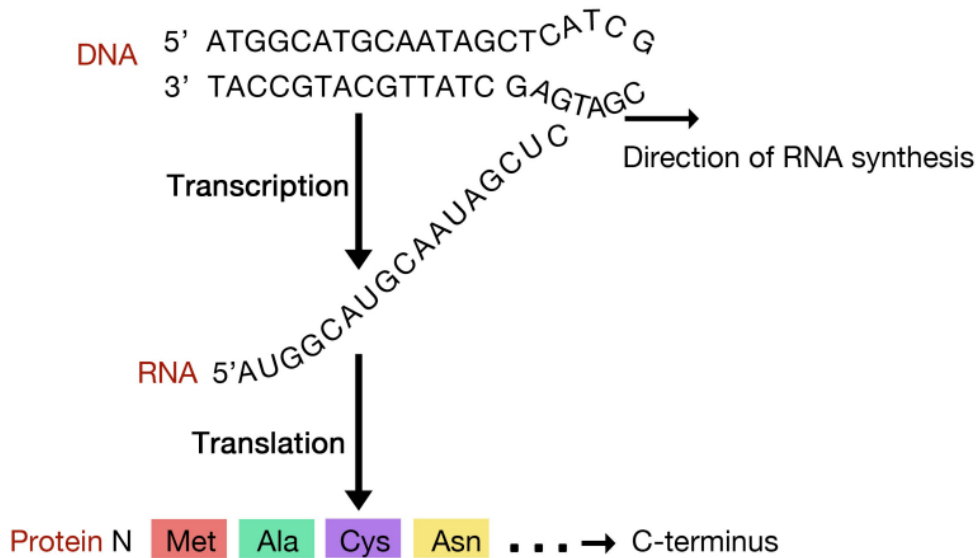
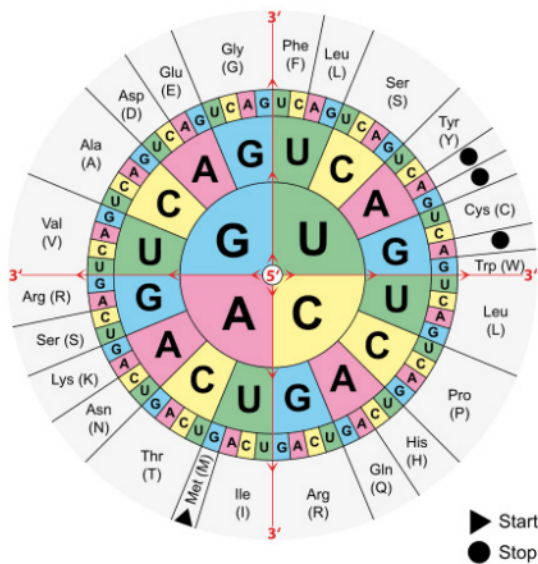


Figure 5.6.1: *Coupled transcription*

and translation in prokaryotes

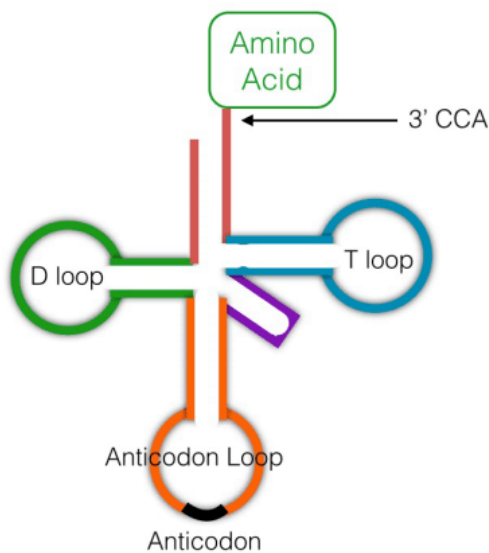
The sequence of an mRNA directly specifies the sequence of amino acids in the protein it encodes. Each amino acid in the protein is specified by a sequence of 3 bases called a codon in the mRNA. For example, the amino acid tryptophan is encoded by the sequence 5'UGG3' on an mRNA. Given that there are 4 bases in RNA, the number of different 3-base combinations that are possible is 4^3 , or 64. There are, however, only 20 amino acids that are used in building proteins. 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. The figure above shows the codons that are used for each of the twenty amino acids.



Three of the 64 codons 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 start, or initiation, codon.

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 111 for a red bead, followed by 222 for a green bead, 333 for yellow, and so on, till you came to 000, indicating that you should stop stringing beads.

While the ribosomes are literally the protein 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. Transfer RNAs (see figure, left) are small RNA molecules, about 75-80 nucleotides long, that function to 'interpret' the instructions in the mRNA during protein synthesis. In terms of the bead analogy above, someone, or something, has to be able to bring a red bead in when the instructions indicate 111, and a green bead when the instructions say 222. Unlike a human, who can choose a red bead when 111 is present in the instructions, neither ribosomes nor tRNAs can think. The system, therefore, relies, like so many processes in cells, solely on molecular recognition.



A given transfer RNA is specific for a particular amino acid. It is linked covalently to this amino acid at its 3' end by an enzyme called aminoacyl tRNA synthetase. 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. Another region of the tRNA has 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, and the amino acid attached to it will be brought in to the ribosome to be added on to the growing protein chain.

With an idea of the various components necessary for translation 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.



Figure 5.6.4: mRNA alignment by Shine-Dalgarno sequence

In bacterial cells, translation is coupled with transcription and begins even before the mRNA has been completely synthesized. How does the ribosome recognize and bind to the mRNA. Many bacterial mRNAs carry a short purine-rich sequence known as the Shine-Dalgarno site upstream of the AUG start codon, as shown in the figure below. This sequence is recognized and bound by a complementary sequence in the 16S rRNA that is part of the small ribosomal subunit as shown above. Because the Shine-Dalgarno site serves to recruit and bind the ribosome, it is also referred to as the Ribosome Binding Site or RBS.

<i>araB</i>	U U U G G A U	G G A G	U G A A A C G	A U G	G C G
<i>galE</i>	A G C C U A A U	G G A G	C G A A U U	A U G	A G A
<i>lacI</i>	C A A U U C A	G G G U G G U	G A U U	G U G	A A A
<i>lacZ</i>	U U C A C A C	A G G A	A A C A G C U	A U G	A C C
<i>trpE</i>	C A A A A U U A	G A G	A A U A A C A	A U G	C A A
<i>trpL</i> leader	G U A A A	A A G G G	U A U C G A C A	A U G	A A A

Figure

5.6.5: Shine-Dalgarno sequences

A variation of this process of ribosome assembly operates in eukaryotic cells. We already know that in eukaryotic cells, processed mRNAs are sent from the nucleus to the cytoplasm.

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.

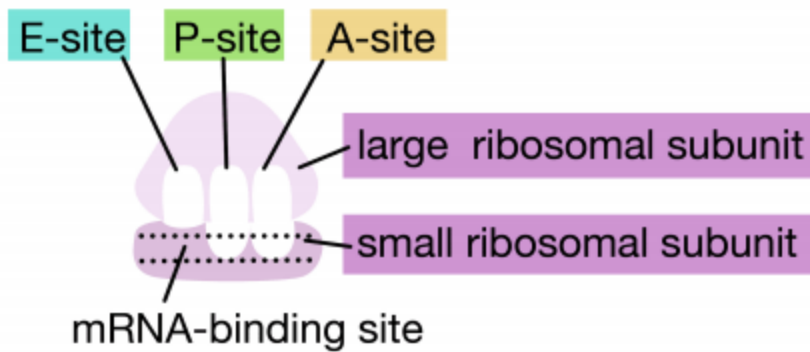


Figure 5.6.6: A, P, and E sites of the ribosome

Protein synthesis in eukaryotes starts with the binding of the small subunit of the ribosome to the 5' end of the mRNA. The assembly of the translation machinery begins with the binding of the small ribosomal subunit to the 7-methyl guanosine cap on the 5' end of an mRNA. Meanwhile, the initiator tRNA pairs with the start codon. (Recall that the start codon is AUG, and codes for methionine. The initiator tRNA carries the amino acid methionine). The large subunit of the ribosome then joins the complex, which is now ready to start protein synthesis.

Ribosomes have two sites for binding charged tRNAs, each of which is positioned to make two adjacent codons on the mRNA available for binding by tRNAs. The initiation codon occupies the first of these ribosomal sites, the P-site. The anticodon complementary to this is on the initiator tRNA, which brings in the first amino acid of the protein. This initial phase of translation is called initiation and requires the help of protein factors called initiation factors.

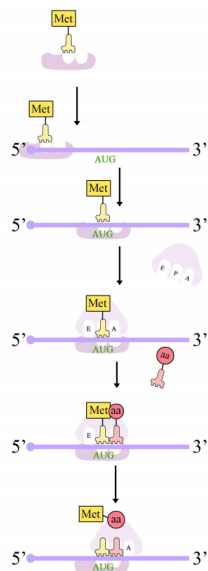


Figure 5.6.7: Initiation of translation

The second codon of the mRNA is positioned adjacent to the second site on the ribosome, the A site. This is where the tRNA carrying the amino acid specified by the second codon binds. The binding of aminoacyl tRNA to the A-site is mediated by proteins called elongation factors and requires the input of energy. Once the appropriate charged tRNAs have "docked" on the codons by base-pairing between the anticodon on the tRNA and the codon on the mRNA, the ribosome joins the amino acids carried by the two tRNAs by making a peptide bond (see figure at right).

Interestingly, the formation of the peptide bond is catalyzed by a catalytic RNA (the 23S rRNA in prokaryotes) rather than by a protein enzyme.

This and subsequent steps in the synthesis of the polypeptide are called the elongation phase of translation. Once the first two amino acids are linked, the first tRNA dissociates, and moves out of the P-site and into the E, or Exit site. The second tRNA then moves into the P-site, vacating the A-site for the tRNA corresponding to the next codon.

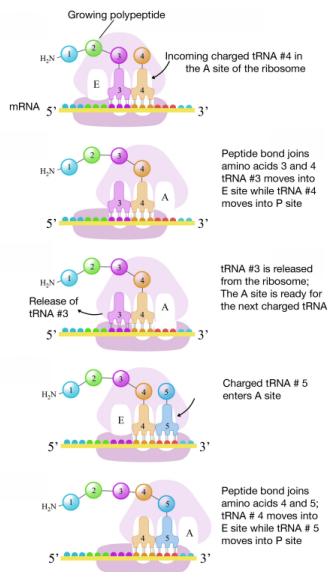


Figure 5.6.8: *Elongation of translation*

The process repeats till the stop codon is in the A-site. At this point, a release factor binds at the A-site, adds a water molecule to the polypeptide at the P-site, and releases the completed polypeptide from the ribosome, which itself, then dissociates into subunits.

As described in Chapter 3, polypeptides made in this way are then folded into their three dimensional shapes, post-translationally modified and delivered to the appropriate cellular compartments to carry out their functions.

This page titled [12.6: Translation](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Kevin Ahern and Indira Rajagopal](#).

- [Wave-Particle Duality](#) is licensed [CC BY-NC-SA 4.0](#).

INDEX

A

affinity chromatography

3.4.1: Affinity Chromatography

allosteric competitive inhibition

8.2: Control of Metabolism Through Enzyme Regulation

Alzheimer's disease

2.2: Protein Folding

amino acids

1.1: Nomenclature of Amino acids

B

Blaber

1.3.2: Stereochemistry of Amino Acids

C

Cell Disruption

3.2: Cell Disruption

Chymotrypsin

6.2: Chymotrypsin

citric acid cycle

10: The Generation of Energy from Carbohydrate Metabolism

cofactor

8.1: Regulation of Enzyme Activity

competitive inhibition

8.1: Regulation of Enzyme Activity
8.2: Control of Metabolism Through Enzyme Regulation

D

denaturation

2.3: Denaturation of proteins

E

electron transport chain

11.1: The Electron Transport Chain

enhancers

12.4: Regulation of Transcription

enzyme regulation

8.2: Control of Metabolism Through Enzyme Regulation

ethidium bromide

3.5.2: DNA Analysis - Gel Electrophoresis

F

Fractionation

3.3: Fractionation

G

gel electrophoresis

3.5: Electrophoresis
3.5.2: DNA Analysis - Gel Electrophoresis

Gel Exclusion Chromatography

3.4.2: Gel Exclusion Chromatography

gel filtration chromatography

3.4.2: Gel Exclusion Chromatography

gluconeogenesis

10: The Generation of Energy from Carbohydrate Metabolism

glycolysis

10: The Generation of Energy from Carbohydrate Metabolism

I

Ion Exchange Chromatography

3.4.3: Ion Exchange Chromatography

Isoelectric focusing

3.5: Electrophoresis

isoelectric point

1.4.2: Acid-Base Reactions of Amino Acids

K

KNF model

8.1: Regulation of Enzyme Activity

Krebs cycle

10: The Generation of Energy from Carbohydrate Metabolism

L

lac operon

12.4: Regulation of Transcription

lagging strand

12.2: DNA Replication

leading strand

12.2: DNA Replication

Lysis

3.2: Cell Disruption

M

mixed inhibition

8.1: Regulation of Enzyme Activity

MWC model

8.1: Regulation of Enzyme Activity

N

noncompetitive inhibition

8.1: Regulation of Enzyme Activity
8.2: Control of Metabolism Through Enzyme Regulation

northern blot

3.5.1: Blotting and Hybridization

O

oxidative phosphorylation

10: The Generation of Energy from Carbohydrate Metabolism
11.1: The Electron Transport Chain
11.2: ATP and Oxidative Phosphorylation

P

pentose phosphate pathway

10: The Generation of Energy from Carbohydrate Metabolism

peptide bond

1.2: Structure of Amino Acids
1.5.1: Polypeptides and Proteins

phosphorylation

11: Oxidation

polypeptides

1.5.1: Polypeptides and Proteins

polysaccharide

9.5: Polysaccharides

primary protein structure

1.5.1: Polypeptides and Proteins

primary transcript

12.3.3: RNA Processing

Promoters

12.4: Regulation of Transcription

protein folding

2.2: Protein Folding

proteins

1.5.1: Polypeptides and Proteins

Q

quaternary protein structure

1.5.1: Polypeptides and Proteins

R

Renaturation

2.3: Denaturation of proteins

Replication

12.2: DNA Replication

repressible system

8: Enzyme Regulation

Repressors

8: Enzyme Regulation

RNA

12.3: RNA

RNA splicing

12.3.3: RNA Processing

S

secondary protein structure

1.5.1: Polypeptides and Proteins

Sonication

3.2: Cell Disruption

Southern blot

3.5.1: Blotting and Hybridization

stop codons

12.6: Translation

T

tertiary protein structure

1.5.1: Polypeptides and Proteins

U

uncompetitive inhibition

8.1: Regulation of Enzyme Activity

W

Western blot

3.5.1: Blotting and Hybridization

GLOSSARY

S

Sample Word 1 | Sample Definition 1

DETAILED LICENSING

Overview

Title: Chem 4320/5320: Biochemistry 1

Webpages: 89

Applicable Restrictions: Noncommercial

All licenses found:

- [CC BY-NC-SA 4.0](#): 79.8% (71 pages)
- [Undeclared](#): 10.1% (9 pages)
- [CC BY-NC 4.0](#): 4.5% (4 pages)
- [CC BY 4.0](#): 3.4% (3 pages)
- [CC BY-SA 4.0](#): 2.2% (2 pages)

By Page

- Chem 4320/5320: Biochemistry 1 - [CC BY-NC-SA 4.0](#)
 - Front Matter - [Undeclared](#)
 - TitlePage - [Undeclared](#)
 - InfoPage - [Undeclared](#)
 - Table of Contents - [Undeclared](#)
 - Licensing - [Undeclared](#)
 - 1: Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.1: Nomenclature of Amino acids - [CC BY-NC-SA 4.0](#)
 - 1.2: Structure of Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.3: Properties of Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.3.1: Charged Nature of Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.3.2: Stereochemistry of Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.4: Reactions of Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.4.1: Acid-base Chemistry of Amino Acids - [CC BY-NC-SA 4.0](#)
 - 1.4.2: Acid-Base Reactions of Amino Acids - [CC BY-NC 4.0](#)
 - 1.4.3: Introduction to Amino Acid Reactivity - [CC BY-NC-SA 4.0](#)
 - 1.5: Amino Acid Analysis and Chemical Sequencing - [CC BY-NC-SA 4.0](#)
 - 1.5.1: Polypeptides and Proteins - [CC BY 4.0](#)
 - 1.5.2: Sequence Determination Using Mass Spectrometry - [CC BY-NC-SA 4.0](#)
 - 1.5.3: The Edman Degradation - [CC BY-NC-SA 4.0](#)
 - 2: Protein Structure - [CC BY-NC-SA 4.0](#)
 - 2.1: The Structure of Proteins - [CC BY-NC 4.0](#)
 - 2.2: Protein Folding - [CC BY-NC-SA 4.0](#)
 - 2.3: Denaturation of proteins - [CC BY-NC-SA 4.0](#)
 - 2.4: Amino Acids and Proteins (Exercises) - [CC BY-NC-SA 4.0](#)
 - 3: Methods of Protein Purification and Characterization - [CC BY-NC-SA 4.0](#)
 - 3.1: Protein Purification - [CC BY-NC-SA 4.0](#)
 - 3.2: Cell Disruption - [CC BY-NC-SA 4.0](#)
 - 3.3: Fractionation - [CC BY-NC-SA 4.0](#)
 - 3.4: Chromatography - [CC BY-NC-SA 4.0](#)
 - 3.4.1: Affinity Chromatography - [CC BY-NC-SA 4.0](#)
 - 3.4.2: Gel Exclusion Chromatography - [CC BY-NC-SA 4.0](#)
 - 3.4.3: Ion Exchange Chromatography - [CC BY-NC-SA 4.0](#)
 - 3.5: Electrophoresis - [CC BY-NC-SA 4.0](#)
 - 3.5.1: Blotting and Hybridization - [CC BY-SA 4.0](#)
 - 3.5.2: DNA Analysis - Gel Electrophoresis - [CC BY-SA 4.0](#)
 - 3.5.3: Exercises - [CC BY-NC-SA 4.0](#)
 - 4: Overview of Hemoglobin and Myoglobin - [CC BY-NC-SA 4.0](#)
 - 4.1: Myoglobin, Hemoglobin, and their Ligands - [CC BY-NC-SA 4.0](#)
 - 4.2: Oxygen Transport by the Proteins Myoglobin and Hemoglobin - [CC BY-NC-SA 4.0](#)
 - 4.3: Exercises - [CC BY-NC-SA 4.0](#)
 - 5: Michaelis-Menten Enzyme Kinetics - [CC BY-NC-SA 4.0](#)
 - 5.1: Catalytic Efficiency of Enzymes - [CC BY-NC-SA 4.0](#)
 - 5.2: Michaelis-Menten Kinetics - [CC BY-NC-SA 4.0](#)

- 4.0
 - 5.3: Enzyme Parameters - [CC BY-NC-SA 4.0](#)
 - 5.4: Enzyme Inhibition - [CC BY-NC-SA 4.0](#)
 - 5.4.1: Enzyme Inhibition - [CC BY-NC-SA 4.0](#)
 - 5.5: Temperature, pH, and enzyme concentration on the rate of a reaction - [CC BY-NC-SA 4.0](#)
 - 5.6: Multi-Substrate Sequential Mechanisms - [CC BY-NC-SA 4.0](#)
 - 5.7: Double displacement reaction - [CC BY 4.0](#)
- 6: Catalytic Strategies of Enzymes - [CC BY-NC-SA 4.0](#)
 - 6.1: Serine Proteases - [CC BY-NC-SA 4.0](#)
 - 6.2: Chymotrypsin - [CC BY-NC-SA 4.0](#)
 - 6.3: Trypsin - [CC BY-NC-SA 4.0](#)
- 7: Enzyme Active Site and Substrate Specificity - [CC BY-NC-SA 4.0](#)
- 8: Enzyme Regulation - [CC BY 4.0](#)
 - 8.1: Regulation of Enzyme Activity - [CC BY-NC-SA 4.0](#)
 - 8.2: Control of Metabolism Through Enzyme Regulation - [CC BY-NC-SA 4.0](#)
 - 8.3: Amino Acids, Proteins, and Enzymes (Summary) - [CC BY-NC-SA 4.0](#)
 - 8.4: Exercises - [CC BY-NC-SA 4.0](#)
- 9: Carbohydrates - [CC BY-NC-SA 4.0](#)
 - 9.1: Carbohydrates Fundamentals - [CC BY-NC-SA 4.0](#)
 - 9.2: Monosaccharides - [CC BY-NC-SA 4.0](#)
 - 9.3: Disaccharides - [CC BY-NC-SA 4.0](#)
 - 9.4: Oligosaccharides - [CC BY-NC-SA 4.0](#)
 - 9.5: Polysaccharides - [CC BY-NC-SA 4.0](#)
 - 9.6: Glycosides - [CC BY-NC-SA 4.0](#)
 - 9.7: Exercises - [CC BY-NC-SA 4.0](#)
- 10: The Generation of Energy from Carbohydrate Metabolism - [CC BY-NC-SA 4.0](#)
- 11: Oxidation - [CC BY-NC-SA 4.0](#)
 - 11.1: The Electron Transport Chain - [CC BY-NC-SA 4.0](#)
 - 11.2: ATP and Oxidative Phosphorylation - [CC BY-NC-SA 4.0](#)
 - 11.3: Exercises - [CC BY-NC-SA 4.0](#)
- 12: The Flow of Genetic Information - [CC BY-NC-SA 4.0](#)
 - 12.1: The Structure of DNA - [CC BY-NC 4.0](#)
 - 12.2: DNA Replication - [CC BY-NC-SA 4.0](#)
 - 12.2.1: Replication - [CC BY-NC-SA 4.0](#)
 - 12.3: RNA - [CC BY-NC-SA 4.0](#)
 - 12.3.1: Protein-RNA Recognition - [CC BY-NC-SA 4.0](#)
 - 12.3.2: RNA - Transcription - [CC BY-NC-SA 4.0](#)
 - 12.3.3: RNA Processing - [CC BY-NC-SA 4.0](#)
 - 12.3.4: Types of RNA - [CC BY-NC-SA 4.0](#)
 - 12.4: Regulation of Transcription - [CC BY-NC-SA 4.0](#)
 - 12.5: The Genetic Code - [CC BY-NC 4.0](#)
 - 12.6: Translation - [CC BY-NC-SA 4.0](#)
- Back Matter - *Undeclared*
 - Index - *Undeclared*
 - Glossary - *Undeclared*
 - Detailed Licensing - *Undeclared*