CHAPTER 28: INTRODUCTION TO BIOCHEMISTRY

Organic and Biochemistry Supplement to Enhanced Introductory College Chemistry

by Gregory Anderson; Jen Booth; Caryn Fahey; Adrienne Richards; Samantha Sullivan Sauer; and David Wegman

Chapter 28 Contents

Please visit the web version of Organic and Biochemistry Supplement to Enhanced Introductory College Chemistry (https://ecampusontario.pressbooks.pub/orgbiochemsupplement/) to access the complete book, interactive activities and ancillary resources.

In this chapter you will learn about:

- Classes and molecular structures of the four biomolecules: carbohydrates, lipids, proteins, and nucleic acids
- Physical and biochemical properties of these biomolecules, as well as important examples of each that affect our daily lives

To better support your learning, you should be familiar with the following concepts before starting this chapter:

- The difference between a physical and chemical property of matter
- sp^2–sp^2 hybridization, sigma and pi bonding to form double bonds in organic molecules (Chapter 21.1: Valence Bond Theory)
- Aldehydes and ketones, and the carbonyl functional group (Chapter 24: Aldehydes and Ketones)
- Carboxylic acids and esters, and the carboxyl functional group (Chapter 25: Aldehydes and Ketones)
- All chemical reactions of aldehydes, ketones, carboxylic acids, and esters

Figure 28.0a. Bread, one of the most common foods in the human diet, is composed of carbohydrates, an important biomolecule. (credit: Photo by Tom Nora on Unsplash)

Biomolecules are the fundamental building blocks of life, each playing a crucial role in various cellular processes. Carbohydrates serve as the primary energy source, providing fuel for cellular activities, and also contribute to structural support in the form of cellulose and chitin. Lipids, including fats and phospholipids, serve as energy reservoirs, insulation, and the main component of cell membranes. Proteins are the workhorses of biology, involved in countless functions, from catalyzing chemical reactions as enzymes to providing structural support as collagen and facilitating cell communication as receptors. Nucleic acids, such as DNA and RNA, store and transmit genetic information, ensuring the inheritance of traits and guiding the synthesis of proteins. These four biomolecules collectively underpin the complexity and diversity of life on Earth.

Watch Biological Polymers: Crash Course Organic Chemistry #49 – YouTube (https://youtu.be/3Pp1AY_lmR4?) (14 min)

Attribution & References

28.1 CARBOHYDRATES

Learning Objectives

By the end of this section, you will be able to:

- Recognize carbohydrates and classify them as mono-, di-, or polysaccharides
- Classify monosaccharides as aldoses or ketoses and as trioses, tetroses, pentoses, or hexoses
- Distinguish between a D sugar and an L sugar
- Identify the structures of D-glucose, D-galactose, and D-fructose and describe how they differ from each other
- Define what is meant by anomers and describe how they are formed
- Explain what is meant by mutarotation
- Identify the chemical properties of monosaccharides
- Identify the structures of sucrose, lactose, and maltose
- Identify the monosaccharides that are needed to form sucrose, lactose, and maltose
- Compare and contrast the structures and uses of starch, glycogen, and cellulose.

Carbohydrates: an Important Source of Energy

Carbohydrates, often referred to as the body's primary source of energy, have played a pivotal role in the history of nutrition and human evolution. Throughout history, carbohydrates from sources like grains, fruits, and vegetables have provided sustenance and fueled societies, enabling the growth of civilizations. In the modern era, the study of carbohydrates has revealed their crucial role in maintaining health, regulating blood sugar levels, and supporting bodily functions. From the complex polysaccharides found in whole grains to the simple sugars in fruits, carbohydrates offer a wide spectrum of nutrients essential for human well-being. Understanding the significance of carbohydrates not only contributes to a balanced diet but also promotes a healthier, more energy-efficient lifestyle, making them an indispensable component of nutrition science and public health.

All carbohydrates consist of carbon, hydrogen, and oxygen atoms and are polyhydroxy aldehydes or ketones

(as seen in Figure 28.1a.) or are compounds that can be broken down to form such compounds. Examples of carbohydrates include starch, fiber, the sweet-tasting compounds called sugars, and structural materials such as cellulose. The term *carbohydrate* had its origin in a misinterpretation of the molecular formulas of many of these substances. For example, because its formula is $C_6H_{12}O_6$, glucose was once thought to be a "carbon hydrate" with the structure $C_6 \cdot 6H_2O$.

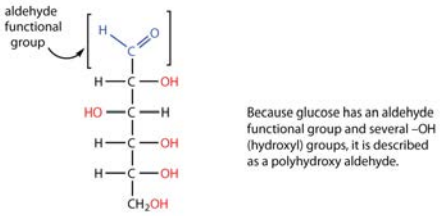


Figure 28.1a. Structure of a common carbohydrate (credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

Green plants are capable of synthesizing glucose ($C_6H_{12}O_6$) from carbon dioxide (CO_2) and water (H_2O) by using solar energy in the process known as photosynthesis:

$$\mathbf{6\,CO_2 + 6\,H_2O + 686\,kcal \rightarrow C_6H_{12}O_6 + 6\,O_2}$$

(The 686 kcal come from solar energy.) Plants can use the glucose for energy or convert it to larger carbohydrates, such as starch or cellulose. Starch provides energy for later use, perhaps as nourishment for a plant's seeds, while cellulose is the structural material of plants. We can gather and eat the parts of a plant that store energy—seeds, roots, tubers, and fruits—and use some of that energy ourselves. Carbohydrates are also needed for the synthesis of nucleic acids and many proteins and lipids.

Animals, including humans, cannot synthesize carbohydrates from carbon dioxide and water and are therefore dependent on the plant kingdom to provide these vital compounds. We use carbohydrates not only for food (about 60%–65% by mass of the average diet) but also for clothing (cotton, linen, rayon), shelter (wood), fuel (wood), and paper (wood).

Spotlight on Everyday Chemistry: Diabetes Mellitus

In the United States, 17.9 million people have been diagnosed with diabetes, and experts estimate that at least another 5.7 million people have the disease but have not been diagnosed. In 2006, diabetes was the seventh leading cause of death, listed on 72,507 death certificates. Moreover, it was a contributing factor in over 200,000 deaths in which the cause was listed as something else, such as heart or kidney disease.

Figure 28.1b. Using a Glucose Meter to Test Blood Glucose Level. (Credit: Photo by Sweet Life on Unsplash)

People with diabetes are impaired in their ability to metabolize glucose, a sugar needed by the body for energy; as a result, excessive quantities of glucose accumulate in the blood and the urine. The characteristic symptoms of diabetes are weight loss, constant hunger, extreme thirst, and frequent urination (the kidneys excrete large amounts of water in an attempt to remove the excess sugar from the blood).

An important diagnostic test for diabetes is the oral glucose tolerance test, which measures the level of glucose in blood plasma. A first measurement is made after a fast of at least 8 h, followed by another measurement 2 h after the person drinks a flavoured solution of 75 g of glucose dissolved in water. At the second measurement, the glucose plasma level should be no higher than 139 mg/dL. Individuals with a value between 140 and 199 mg/dL are diagnosed with prediabetes, while those with a value of 200 mg/dL or above are diagnosed with diabetes. Following a diagnosis of diabetes, a person will need to monitor his or her blood glucose levels daily (or more often) using a glucose meter.

The simplest carbohydrates—those that cannot be hydrolyzed to produce even smaller carbohydrates—are called monosaccharides. Two or more monosaccharides can link together to form chains that contain from two to several hundred or thousand monosaccharide units. Prefixes are used to indicate the number of such units in the chains. Disaccharide molecules have two monosaccharide units, *trisaccharide* molecules have three units, and so on. Chains with many monosaccharide units joined together are called polysaccharides. All these so-called higher saccharides can be hydrolyzed back to their constituent monosaccharides. Compounds that cannot be hydrolyzed will not react with water to form two or more smaller compounds.

Classes of Monosaccharides

The naturally occurring monosaccharides contain three to seven carbon atoms per molecule. Monosaccharides of specific sizes may be indicated by names composed of a stem denoting the number of carbon atoms and the suffix *–ose*. For example, the terms *triose*, *tetrose*, *pentose*, and *hexose* signify monosaccharides with, respectively, three, four, five, and six carbon atoms. Monosaccharides are also classified as aldoses or ketoses. Those monosaccharides that contain an aldehyde functional group are called aldoses; those containing a ketone functional group on the second carbon atom are ketoses. Combining these classification systems gives general names that indicate both the type of carbonyl group *and* the number of carbon atoms in a molecule. Thus, monosaccharides are described as aldotetroses, aldopentoses, ketopentoses, ketoheptoses, and so forth. Glucose and fructose are specific examples of an aldohexose and a ketohexose, respectively, as shown in Figure 28.1c.

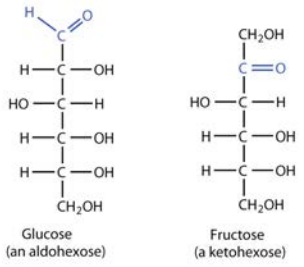


Figure 28.1c. Glucose contains an aldehyde functional group and is six carbons long; thus it is classified as an aldohexose. Fructose, on the other hand, contains a ketone functional group and is also six carbons long; thus it is classified as a ketohexose (credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

Example 28.1a

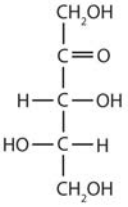
Draw an example of each type of compound.

a. a ketopentose
b. an aldotetrose

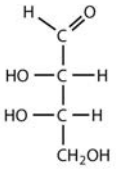
Solution

a. The structure must have five carbon atoms with the second carbon atom being a carbonyl group and the other four carbon atoms each having an OH group attached. Several structures are

possible, but one example is shown:



b. The structure must have four carbon atoms with the first carbon atom part of the aldehyde functional group. The other three carbon atoms each have an OH group attached. Several structures are possible, but one example is shown:



Example source: *Introduction to Chemistry: GOB (V. 1.0).*, CC BY-NC-SA 3.0.

Exercise 28.1a

Examining Figure 28.1c, identify the main functional groups in each molecule. List some of the physical properties of each functional group.

Check Your Answer: [1]

Source: Exercise 28.1a by Samantha Sullivan Sauer is licensed under CC BY-NC 4.0.

The simplest sugars are the trioses. The possible trioses are shown in part (a) of Figure 28.1d: glyceraldehyde is an aldotriose, while dihydroxyacetone is a ketotriose. Notice that two structures are shown for glyceraldehyde. These structures are **stereoisomers**, and hence are isomers having the same structural formula but differing in the arrangement of atoms or groups of atoms in three-dimensional space. If you make models of the two stereoisomers of glyceraldehyde, you will find that you cannot place one model on top of the other and have each functional group point in the same direction. However, if you place one of the models in front of a mirror, the image in the mirror will be identical to the second stereoisomer in part (b) of Figure 28.1d. Molecules that are nonsuperimposable (nonidentical) mirror images of each other are a type of stereoisomer

called **enantiomers** (Greek *enantios*, meaning "opposite"). It is important to note that these are another type of stereoisomer than the cis-trans (geometric) isomers previously discussed.

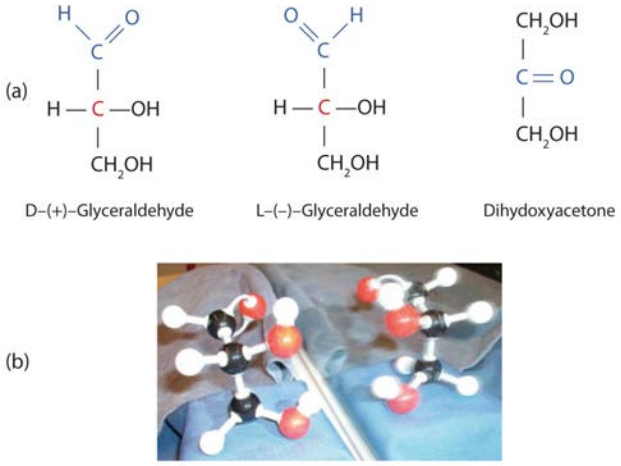


Figure 28.1d. Structures of the Trioses. (a) D- and L-glyceraldehyde are mirror images of each other and represent a pair of enantiomers. (b) A ball-and-stick model of D-glyceraldehyde is reflected in a mirror. Note that the reflection has the same structure as L-glyceraldehyde (credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

A key characteristic of enantiomers is that they have a carbon atom to which four different groups are attached. Note, for example, the four different groups attached to the central carbon atom of glyceraldehyde (part (a) of Figure 28.1d.). A carbon atom that has four different groups attached is a chiral carbon. If a molecule contains one or more chiral carbons, it is likely to exist as two or more stereoisomers. Dihydroxyacetone does not contain a chiral carbon and thus does not exist as a pair of stereoisomers. Glyceraldehyde, however, has a chiral carbon and exists as a pair of enantiomers. Except for the direction in which each enantiomer rotates plane-polarized light, these two molecules have identical physical properties. One enantiomer has a specific rotation of +8.7°, while the other has a specific rotation of −8.7°.

H. Emil Fischer, a German chemist, developed the convention commonly used for writing two-dimensional representations of the monosaccharides, such as those in part (a) of Figure 28.1d. In these structural formulas, the aldehyde group is written at the top, and the hydrogen atoms and OH groups that are attached to each chiral carbon are written to the right or left. (If the monosaccharide is a ketose, the ketone functional group is the second carbon atom.) Vertical lines represent bonds pointing away from you, while horizontal lines represent bonds coming toward you. The formulas of chiral molecules represented in this manner are referred to as **Fischer projections**.

The two enantiomers of glyceraldehyde are especially important because monosaccharides with more than three carbon atoms can be considered as being derived from them. Thus, D- and L-glyceraldehyde provide reference points for designating and drawing all other monosaccharides. Sugars whose Fischer projections terminate in the same configuration as D-glyceraldehyde are designated as D sugars; those derived from L-

glyceraldehyde are designated as L sugars. By convention, the penultimate (next-to-last) carbon atom has been chosen as the carbon atom that determines if a sugar is D or L. It is the chiral carbon farthest from the aldehyde or ketone functional group.

Dextrorotatory or Levorotatory: Polarized Light

A beam of ordinary light can be pictured as a bundle of waves; some move up and down, some sideways, and others at all other conceivable angles. When a beam of light has been polarized, however, the waves in the bundle all vibrate in a single plane. Light altered in this way is called *plane-polarized light*. Much of what chemists know about stereoisomers comes from studying the effects they have on plane-polarized light. In Figure 28.1e., the light on the left is not polarized, while that on the right is polarized.

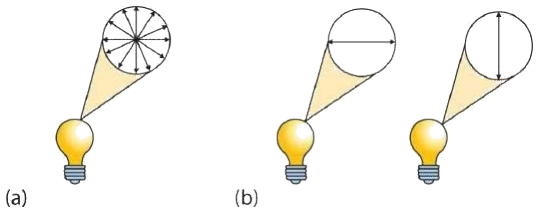


Figure 28.1e. When ordinary light is not polarized, the waves vibrate in all conceivable angles (a). However, when light is polarized, the waves will vibrate in one single plane (b). (Credit: *Introduction to Chemistry: GOB (V. 1.0).*, CC BY-NC-SA 3.0., edited by *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.)

Sunlight, in general, is not polarized; light from an ordinary light bulb or an ordinary flashlight is not polarized. One way to polarize ordinary light is to pass it through Polaroid sheets, special plastic sheets containing carefully oriented organic compounds that permit only light vibrating in a single plane to pass through. To the eye, polarized light doesn't "look" any different from nonpolarized light. We can detect polarized light, however, by using a second sheet of polarizing material, as shown in Figure 28.1f.

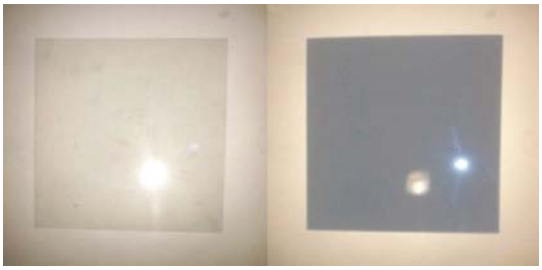


Figure 28.1f. In the photo on the left, two Polaroid sheets are aligned in the same direction; plane-polarized light from the first Polaroid sheet can pass through the second sheet. In the photo on the right, the top Polaroid sheet has been rotated 90° and now blocks the plane-polarized light that comes through the first Polaroid sheet (credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

Certain substances act on polarized light by rotating the plane of vibration. Such substances are said to be optically active. The extent of optical activity is measured by a polarimeter, an instrument that contains two polarizing lenses separated by a sample tube, as shown in Figure 28.1g. With the sample tube empty, maximum light reaches the observer's eye when the two lenses are aligned so that both pass light vibrating in the same plane. When an optically active substance is placed in the sample tube, that substance rotates the plane of polarization of the light passing through it, so that the polarized light emerging from the sample tube is vibrating in a different direction than when it entered the tube. To see the maximum amount of light when the sample is in place, the observer must rotate one lens to accommodate the change in the plane of polarization.

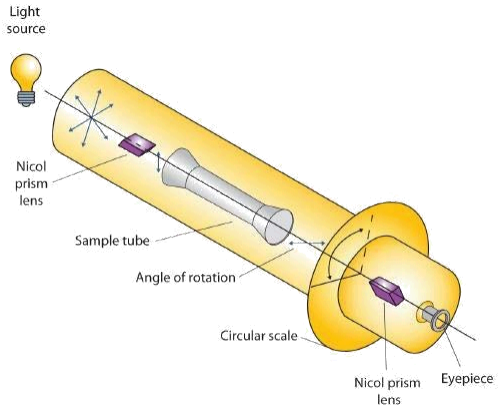


Figure 28.1g. Diagram of a Polarimeter (Credit: *Introduction to Chemistry: GOB (V. 1.0).*, CC BY-NC-SA 3.0., edited by *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.)

Some optically active substances rotate the plane of polarized light to the right (clockwise) from the observer's point of view. These compounds are said to be **dextrorotatory**; substances that rotate light to the left (counterclockwise) are **levorotatory**. To denote the direction of rotation, a positive sign (+) is given to dextrorotatory substances, and a negative sign (−) is given to levorotatory substances.

Important Hexoses

Although a variety of monosaccharides are found in living organisms, three hexoses are particularly abundant: D-glucose, D-galactose, and D-fructose (Figure 28.1h.). Glucose and galactose are both aldohexoses, while fructose is a ketohexose.

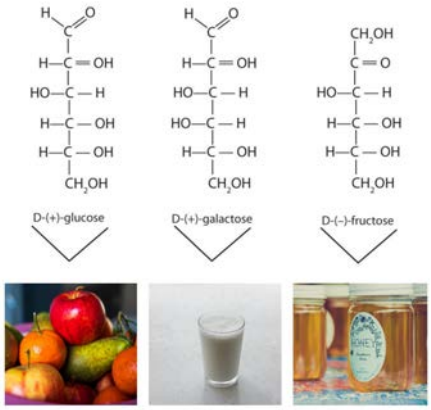


Figure 28.1h. Structures of Three Important Hexoses. Each hexose is pictured with a food source in which it is commonly found. (credit: *Introduction to Chemistry: GOB (v. 1.0)* , CC BY-NC-SA 4.0; edited by Reva to include Photo by Nick Fewings, Photo by engin akyurt and Photo by Amelia Bartlett all under Unsplash license)

Glucose

D-Glucose, generally referred to as simply glucose, is the most abundant sugar found in nature; most of the carbohydrates we eat are eventually converted to it in a series of biochemical reactions that produce energy for our cells (Figure 28.1i.). It is also known by three other names: *dextrose*, from the fact that it rotates plane-polarized light in a clockwise (dextrorotatory) direction; *corn sugar* because in the United States cornstarch is used in the commercial process that produces glucose from the hydrolysis of starch; and *blood sugar* because it is the carbohydrate found in the circulatory system of animals. Normal blood sugar values range from 70 to 105 mg glucose/dL plasma, and normal urine may contain anywhere from a trace to 20 mg glucose/dL urine.

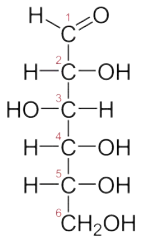


Figure 28.1i. Fischer projection of D-glucose. (Credit: Photo by Benjah-bmm27 edited by Yikrazuul , Public Domain)

Glucose is a D sugar because the OH group on the fifth carbon atom (the chiral center farthest from the carbonyl group) is on the right. In fact, all the OH groups except the one on the third carbon atom are to the right.

Indigenous Perspectives: Kivak, a Greenland Inuit Delicacy

Figure 28.1j. Little auks sitting on a rock (Image by AWeith, CC BY-SA 4.0).

A delicacy within the Inuit is Kivak. It involves the fermentation of glucose using lactic acid bacteria in a process where little auks (Figure 28.1j.) are stuffed into seal skin and stored under rocks for 18 months. This process helps with the preservation and flavours of the dish. For more information see the infographic Compound Interest: The Chemistry Advent Calendar 2023 (compoundchem.com)

Galactose

D-Galactose does not occur in nature in the uncombined state. It is released when lactose, a disaccharide found in milk, is hydrolyzed. The galactose needed by the human body for the synthesis of lactose is obtained

by the metabolic conversion of D-glucose to D-galactose. Galactose is also an important constituent of the glycolipids that occur in the brain and the myelin sheath of nerve cells. For this reason it is also known as *brain sugar*. The structure of D-galactose is shown in Figure 28.1h. Notice that the configuration differs from that of glucose only at the fourth carbon atom.

Fructose

D-Fructose, also shown in Figure 28.1h, is the most abundant ketohexose. Note that from the third through the sixth carbon atoms, its structure is the same as that of glucose. It occurs, along with glucose and sucrose, in honey (which is 40% fructose) and sweet fruits. Fructose (from the Latin *fructus*, meaning "fruit") is also referred to as *levulose* because it has a specific rotation that is strongly levorotatory (−92.4°). It is the sweetest sugar, being 1.7 times sweeter than sucrose, although many nonsugars are several hundred or several thousand times as sweet (Table 28.1a.).

Table 28.1a. The Relative Sweetness of Some Compounds

Compound	Relative Sweetness
lactose	16
maltose	32
glucose	74
sucrose	100
fructose	173
aspartame	18,000
acesulfame K	20,000
saccharin	30,000
sucralose	60,000

Spotlight on Everyday Chemistry: Artificial Sweeteners

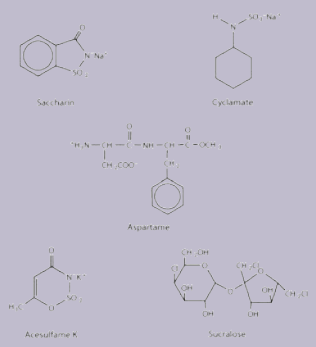


Figure 28.1k. The structures of five artificial sweeteners that are approved for consumption: saccharin, cyclamate, aspartame, acesulfame K, and sucralose (Credit: *Introduction to Chemistry: GOB (v. 1.0)* ., edited by *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.)

Although sweetness is commonly associated with mono- and disaccharides, it is not a property found only in sugars. Several other kinds of organic compounds have been synthesized that are far superior as sweetening agents. These so-called high-intensity or artificial sweeteners are useful for people with diabetes or other medical conditions that require them to control their carbohydrate intake. The synthetic compounds are noncaloric or used in such small quantities that they do not add significantly to the caloric value of food.

The first artificial sweetener—saccharin—was discovered by accident in 1879. It is 300 times sweeter than sucrose, but it passes through the body unchanged and thus adds no calories to the diet. After its discovery, saccharin was used until it was banned in the early 1900s. However, during the sugar-short years of World War I, the ban was lifted and was not reinstated at the war's end. One drawback to the use of saccharin is its bitter, metallic aftertaste. The initial solution to this problem was to combine saccharin with cyclamate, a second artificial sweetener discovered in 1937.

In the 1960s and 1970s, several clinical tests with laboratory animals implicated both cyclamate and saccharin as carcinogenic (cancer-causing) substances. The results from the cyclamate tests were completed first, and cyclamate was banned in the United States in 1969. Then a major study was released in Canada in 1977 indicating that saccharin increased the incidence of bladder cancer in rats. The US Food and Drug Administration (FDA) proposed a ban on saccharin that raised immediate public opposition because saccharin was the only artificial sweetener still available. In response, Congress passed the Saccharin Study and Labeling Act in 1977, permitting the use of saccharin as long as any product containing it was labeled with a consumer warning regarding the possible elevation of the risk of bladder cancer. Today this warning is no longer required; moreover, the FDA is currently reviewing the ban on cyclamate, as 75 additional studies and years of usage in other countries, such as Canada, have failed to show that it has any carcinogenic effect.

A third artificial sweetener, aspartame, was discovered in 1965. This white crystalline compound is about 180 times sweeter than sucrose and has no aftertaste. It was approved for use in 1981 and is

used to sweeten a wide variety of foods because it blends well with other food flavours. Aspartame is not used in baked goods, however, because it is not heat stable.

In the body (or when heated), aspartame is initially hydrolyzed to three molecules: the amino acids aspartic acid and phenylalanine and an alcohol methanol. Repeated controversy regarding the safety of aspartame arises partly from the fact that the body metabolizes the released methanol to formaldehyde. It should be noted, though, that a glass of tomato juice has six times as much methanol as a similar amount of a diet soda containing aspartame. The only documented risk connected to aspartame use is for individuals with the genetic disease *phenylketonuria* (PKU); these individuals lack the enzyme needed to metabolize the phenylalanine released when aspartame is broken down by the body. Because of the danger to people with PKU, all products containing aspartame must carry a warning label.

Acesulfame K, discovered just two years after aspartame (1967), was approved for use in the United States in 1988. It is 200 times sweeter than sugar and, unlike aspartame, is heat stable. It has no lingering aftertaste.

One of the newest artificial sweeteners to gain FDA approval (April 1998) for use in the United States is sucralose, a white crystalline solid approximately 600 times sweeter than sucrose. Sucralose is synthesized from sucrose and has three chlorine atoms substituted for three OH groups. It is noncaloric because it passes through the body unchanged. It can be used in baking because it is heat stable.

The structures of these artificial sweeteners can be seen in Figure 28.1k. All of the extensive clinical studies completed to date have indicated that these artificial sweeteners approved for use in the United States are safe for consumption by healthy individuals in moderate amounts.

Cyclic Structures of Monosaccharides

So far, we have represented monosaccharides as linear molecules, but many of them also adopt cyclic structures. This conversion occurs because of the ability of aldehydes and ketones to react with alcohols, as shown in Figure 28.1l.

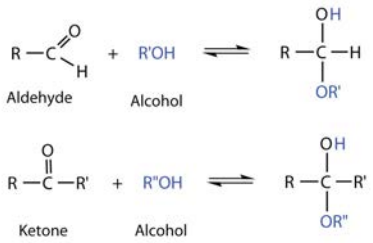


Figure 28.1l. Reaction schemes for an aldehyde reacting with an alcohol, and a ketone reacting with an alcohol. R stands in for any alkyl group (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Take a look at the reaction mechanism for how the linear monosaccharide D-glucose adopts a ring structure in Figure 28.1m. You might wonder why the aldehyde reacts with the OH group on the fifth carbon atom rather than the OH group on the second carbon atom next to it. Recall that cyclic alkanes containing five or six carbon atoms in the ring are the most stable. The same is true for monosaccharides that form cyclic structures: rings consisting of five or six carbon atoms are the most stable.

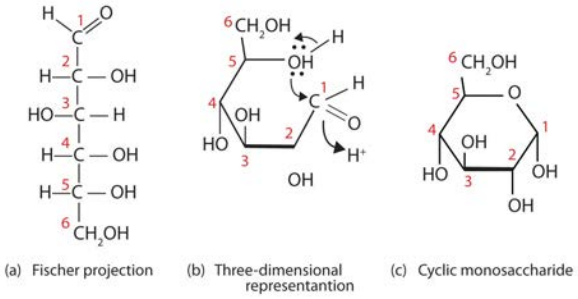


Figure 28.1m. Cyclization of D-Glucose. D-Glucose can be represented with a Fischer projection (a) or three dimensionally (b). By reacting the OH group on the fifth carbon atom with the aldehyde group, the cyclic monosaccharide (c) is produced (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

When a straight-chain monosaccharide, such as any of the structures shown in figure 28.1h, forms a cyclic structure, the carbonyl oxygen atom may be pushed either up or down, giving rise to two stereoisomers, as shown in Figure 28.1n. The structure shown on the left side of Figure 28.1n., with the OH group on the first carbon atom projected downward, represent what is called the alpha (α) form. The structures on the right side, with the OH group on the first carbon atom pointed upward, is the beta (β) form. These two stereoisomers of a cyclic monosaccharide are known as anomers; they differ in structure around the anomeric carbon—that is, the carbon atom that was the carbonyl carbon atom in the straight-chain form.

It is possible to obtain a sample of crystalline glucose in which all the molecules have the α structure or all have the β structure. The α form melts at 146°C and has a specific rotation of +112°, while the β form melts at 150°C and has a specific rotation of +18.7°. When the sample is dissolved in water, however, a mixture is soon

produced containing both anomers as well as the straight-chain form, in dynamic equilibrium (part (a) of Figure 28.1n.). You can start with a pure crystalline sample of glucose consisting entirely of either anomer, but as soon as the molecules dissolve in water, they open to form the carbonyl group and then reclose to form either the α or the β anomer. The opening and closing repeats continuously in an ongoing interconversion between anomeric forms and is referred to as mutarotation (Latin *mutare*, meaning "to change"). At equilibrium, the mixture consists of about 36% α-D-glucose, 64% β-D-glucose, and less than 0.02% of the open-chain aldehyde form. The observed rotation of this solution is +52.7°.

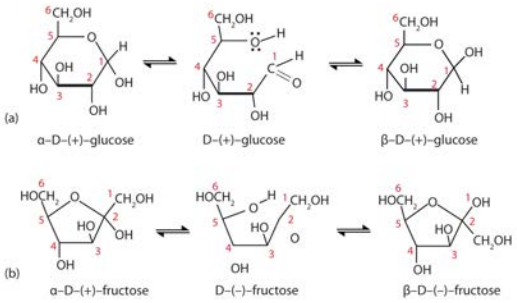


Figure 28.1n. Monosaccharides. In an aqueous solution, monosaccharides exist as an equilibrium mixture of three forms. The interconversion between the forms is known as mutarotation, which is shown for D-glucose (a) and D-fructose (b). (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Even though only a small percentage of the molecules are in the open-chain aldehyde form at any time, the solution will nevertheless exhibit the characteristic reactions of an aldehyde. As the small amount of free aldehyde is used up in a reaction, there is a shift in the equilibrium to yield more aldehyde. Thus, all the molecules may eventually react, even though very little free aldehyde is present at a time.

Commonly, the cyclic forms of sugars are depicted using a convention first suggested by Walter N. Haworth, an English chemist. The molecules are drawn as planar hexagons with a darkened edge representing the side facing toward the viewer. The structure is simplified to show only the functional groups attached to the carbon atoms. Any group written to the right in a Fischer projection appears below the plane of the ring in a Haworth projection, and any group written to the left in a Fischer projection appears above the plane in a Haworth projection.

The difference between the α and the β forms of sugars may seem trivial, but such structural differences are often crucial in biochemical reactions. This explains why we can get energy from the starch in potatoes and other plants but not from cellulose, even though both starch and cellulose are polysaccharides composed of glucose molecules linked together.

Chemical Properties of Monosaccharides

Monosaccharides such as glucose and fructose are crystalline solids at room temperature, but they are quite soluble in water, each molecule having several OH groups that readily engage in hydrogen bonding. The chemical behavior of these monosaccharides is likewise determined by their functional groups.

An important reaction of monosaccharides is the oxidation of the aldehyde group, one of the most easily oxidized organic functional groups. Aldehyde oxidation can be accomplished with any mild oxidizing agent, such as Tollens' reagent or Benedict's reagent. With the latter, complexed copper(II) ions are reduced to copper(I) ions that form a brick-red precipitate [copper(I) oxide; Figure 28.1o.].

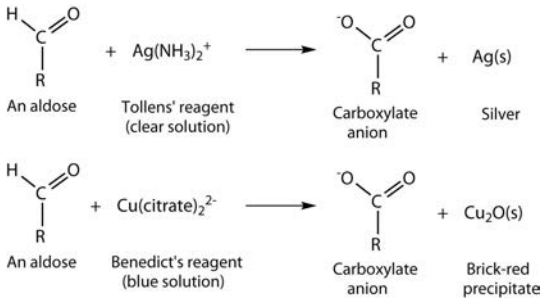


Figure 28.1o. Oxidation reactions of monosaccharides using either Tollens' reagent or Benedict's reagent (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Any carbohydrate capable of reducing either Tollens' or Benedict's reagents without first undergoing hydrolysis is said to be a reducing sugar (Figure 28.1p.). Because both the Tollens' and Benedict's reagents are basic solutions, ketoses (such as fructose) also give positive tests due to an equilibrium that exists between ketoses and aldoses in a reaction known as *tautomerism*.

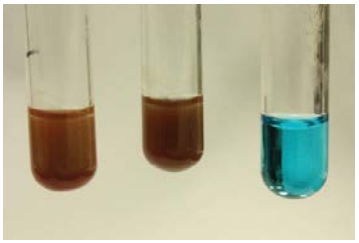


Figure 28.1p. Benedict's Test. Benedict's test was performed on three carbohydrates, depicted from left to right: fructose, glucose, and sucrose. The solution containing sucrose remains blue because sucrose is a nonreducing sugar (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

These reactions have been used as simple and rapid diagnostic tests for the presence of glucose in blood or

urine. For example, Clinitest tablets, which are used to test for sugar in the urine, contain copper(II) ions and are based on Benedict's test. A green colour indicates very little sugar, whereas a brick-red colour indicates sugar in excess of 2 g/100 mL of urine.

Disaccharides

Previously, you learned that monosaccharides can form cyclic structures by the reaction of the carbonyl group with an OH group. These cyclic molecules can in turn react with another alcohol. Disaccharides ($C_{12}H_{22}O_{11}$) are sugars composed of two monosaccharide units that are joined by a carbon–oxygen-carbon linkage known as a glycosidic linkage. This linkage is formed from the reaction of the anomeric carbon of one cyclic monosaccharide with the OH group of a second monosaccharide, as shown in Figure 28.1q.

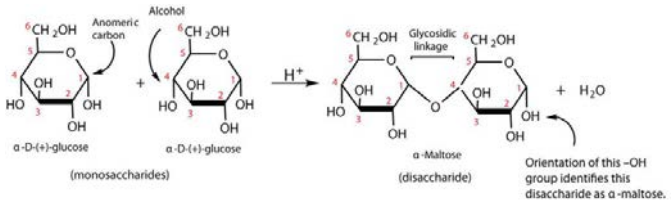


Figure 28.1q. Reaction mechanism depicting how two monosaccharides can link to form a disaccharide. (Credit: *Introduction to Chemistry: GOB(v. 1.0)*, edited by *(Ball et al.)* , CC BY-NC-SA 4.0)

The disaccharides differ from one another in their monosaccharide constituents and in the specific type of glycosidic linkage connecting them. There are three common disaccharides: maltose, lactose, and sucrose. All three are white crystalline solids at room temperature and are soluble in water. We'll consider each sugar in more detail.

Maltose

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

In the body, such hydrolysis reactions are catalyzed by enzymes such as *maltase*. The same reactions can be carried out in the laboratory with dilute acid as a catalyst, although in that case the rate is much slower, and high temperatures are required. Whether it occurs in the body or a glass beaker, the hydrolysis of maltose produces two molecules of D-glucose:

$$\text{maltose} \xrightarrow{H^+ \text{ or maltase}} 2\ \text{D-glucose}$$

Maltose is a reducing sugar. Thus, its two glucose molecules must be linked in such a way as to leave one anomeric carbon that can open to form an aldehyde group. The glucose units in maltose are joined in a *head-to-tail* fashion through an α-linkage from the first carbon atom of one glucose molecule to the fourth carbon atom of the second glucose molecule (that is, an α-1,4-glycosidic linkage; see Figure 28.1q.). The bond from the anomeric carbon of the first monosaccharide unit is directed downward, which is why this is known as an α-glycosidic linkage. The OH group on the anomeric carbon of the second glucose can be in either the α or the β position, as shown in Figure 28.1r.

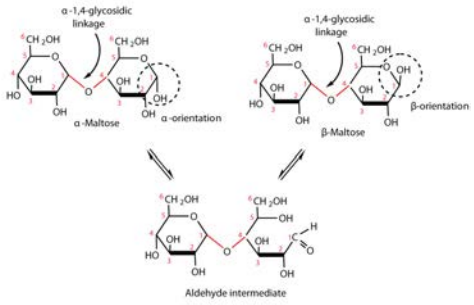


Figure 28.1r. An Equilibrium Mixture of Maltose Isomers (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Lactose

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

Lactose is a reducing sugar composed of one molecule of D-galactose and one molecule of D-glucose joined by a β-1,4-glycosidic bond (the bond from the anomeric carbon of the first monosaccharide unit being directed upward). The two monosaccharides are obtained from lactose by acid hydrolysis or the catalytic action of the enzyme *lactase,* as shown in Figure 28.1s.

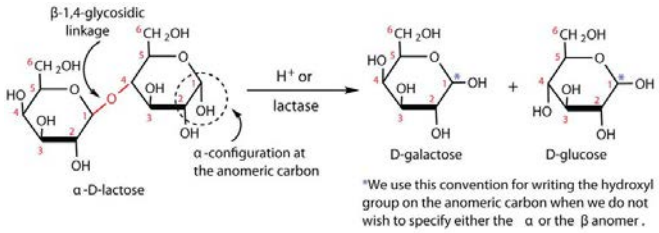
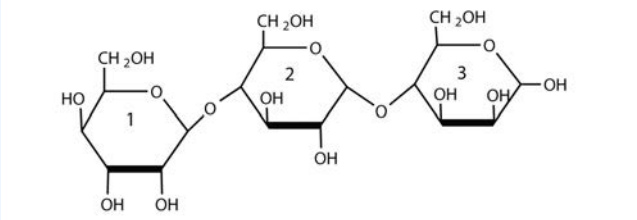


Figure 28.1s. Reaction mechanism of the acid hydrolysis of lactose into its constituent monosaccharides by the enzyme lactase. (Credit: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)* , CC BY-NC-SA 4.0)

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

Example 28.1b

For this trisaccharide, indicate whether each glycosidic linkage is α or β.



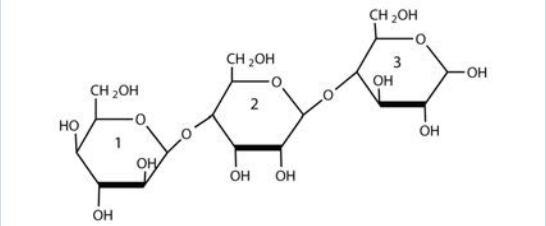
(credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Solution

The glycosidic linkage between sugars 1 and 2 is β because the bond is directed up from the anomeric carbon. The glycosidic linkage between sugars 2 and 3 is α because the bond is directed down from the anomeric carbon.

Exercise 28.1b

For this trisaccharide, indicate whether each glycosidic linkage is α or β.



(credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Check Your Answer[2]

Spotlight on Everyday Chemistry: Lactose Intolerance and Galactosemia

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

In people with lactose intolerance, some of the unhydrolyzed lactose passes into the colon, where it tends to draw water from the interstitial fluid into the intestinal lumen by osmosis. At the same time, intestinal bacteria may act on the lactose to produce organic acids and gases. The buildup of water and bacterial decay products leads to abdominal distention, cramps, and diarrhea, which are symptoms of the condition.

The symptoms disappear if milk or other sources of lactose are excluded from the diet or consumed only sparingly. Alternatively, many food stores now carry special brands of milk that have been

pretreated with lactase to hydrolyze the lactose. Cooking or fermenting milk causes at least partial hydrolysis of the lactose, so some people with lactose intolerance are still able to enjoy cheese, yogurt, or cooked foods containing milk. The most common treatment for lactose intolerance, however, is the use of lactase preparations (e.g., Lactaid), which are available in liquid and tablet form at drugstores and grocery stores. These are taken orally with dairy foods—or may be added to them directly—to assist in their digestion.

Galactosemia is a condition in which one of the enzymes needed to convert galactose to glucose is missing. Consequently, the blood galactose level is markedly elevated, and galactose is found in the urine. An infant with galactosemia experiences a lack of appetite, weight loss, diarrhea, and jaundice. The disease may result in impaired liver function, cataracts, mental retardation, and even death. If galactosemia is recognized in early infancy, its effects can be prevented by the exclusion of milk and all other sources of galactose from the diet. As a child with galactosemia grows older, he or she usually develops an alternate pathway for metabolizing galactose, so the need to restrict milk is not permanent. The incidence of galactosemia in the United States is 1 in every 65,000 newborn babies.

Sucrose

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

The sucrose molecule is unique among the common disaccharides in having an α-1,β-2-glycosidic (head-to-head) linkage. Because this glycosidic linkage is formed by the OH group on the anomeric carbon of α-D-glucose and the OH group on the anomeric carbon of β-D-fructose, it ties up the anomeric carbons of both glucose and fructose, as shown in Figure 28.1t.

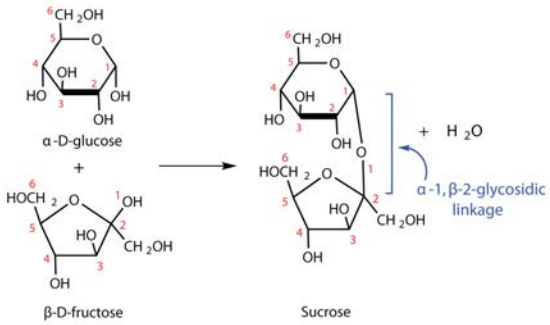


Figure 28.1t. Reaction mechanism of the formation of sucrose from the starting monosaccharides glucose and fructose (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

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

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

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

Polysaccharides

The polysaccharides are the most abundant carbohydrates in nature and serve a variety of functions, such as energy storage or as components of plant cell walls. Polysaccharides are very large polymers composed of tens to thousands of monosaccharides joined together by glycosidic linkages. The three most abundant polysaccharides are starch, glycogen, and cellulose. These three are referred to as *homopolymers* because each yields only one type of monosaccharide (glucose) after complete hydrolysis. *Heteropolymers* may contain sugar

acids, amino sugars, or noncarbohydrate substances in addition to monosaccharides. Heteropolymers are common in nature (gums, pectins, and other substances) but will not be discussed further in this textbook. The polysaccharides are nonreducing carbohydrates, are not sweet tasting, and do not undergo mutarotation.

Starch

Starch is the most important source of carbohydrates in the human diet and accounts for more than 50% of our carbohydrate intake. It occurs in plants in the form of granules, and these are particularly abundant in seeds (especially the cereal grains) and tubers, where they serve as a storage form of carbohydrates. The breakdown of starch to glucose nourishes the plant during periods of reduced photosynthetic activity. We often think of potatoes as a "starchy" food, yet other plants contain a much greater percentage of starch (potatoes 15%, wheat 55%, corn 65%, and rice 75%). Commercial starch is a white powder.

Starch is a mixture of two polymers: amylose and amylopectin. Natural starches consist of about 10%–30% amylose and 70%–90% amylopectin. Amylose is a linear polysaccharide composed entirely of D-glucose units joined by the α-1,4-glycosidic linkages we saw in maltose (part (a) of Figure 28.1u.). Experimental evidence indicates that amylose is not a straight chain of glucose units but instead is coiled like a spring, with six glucose monomers per turn (part (b) of Figure 28.1u.). When coiled in this fashion, amylose has just enough room in its core to accommodate an iodine molecule. The characteristic blue-violet colour that appears when starch is treated with iodine is due to the formation of the amylose-iodine complex. This colour test is sensitive enough to detect even minute amounts of starch in solution.

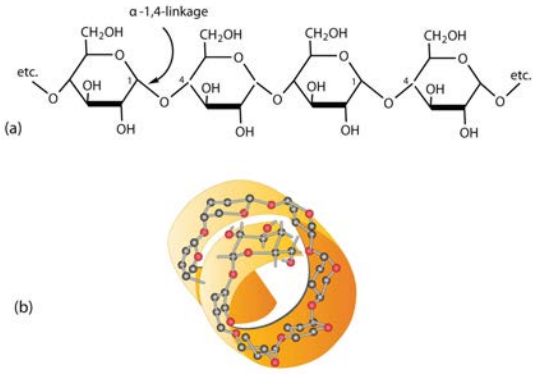


Figure 28.1u. Amylose. (a) Amylose is a linear chain of α-D-glucose units joined together by α-1,4-glycosidic bonds. (b) Because of hydrogen bonding, amylose acquires a spiral structure that contains six glucose units per turn (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Amylopectin is a branched-chain polysaccharide composed of glucose units linked primarily by α-1,4-glycosidic bonds but with occasional α-1,6-glycosidic bonds, which are responsible for the branching. A molecule of amylopectin may contain many thousands of glucose units with branch points occurring about every 25–30 units (Figure 28.1v.). The helical structure of amylopectin is disrupted by the branching of the

chain, so instead of the deep blue-violet colour amylose gives with iodine, amylopectin produces a less intense reddish brown.

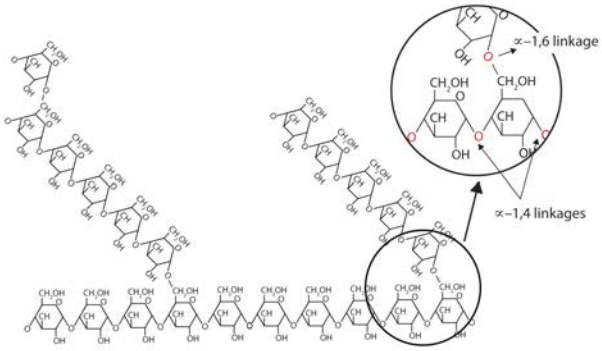


Figure 28.1v. Representation of the Branching in Amylopectin and Glycogen. Both amylopectin and glycogen contain branch points that are linked through **α**-1,6-linkages. These branch points occur more often in glycogen (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Dextrins are glucose polysaccharides of intermediate size. The shine and stiffness imparted to clothing by starch are due to the presence of dextrins formed when clothing is ironed. Because of their characteristic stickiness with wetting, dextrins are used as adhesives on stamps, envelopes, and labels; as binders to hold pills and tablets together; and as pastes. Dextrins are more easily digested than starch and are therefore used extensively in the commercial preparation of infant foods.

The complete hydrolysis of starch yields, in successive stages, glucose:

$$\text{starch} \rightarrow \text{dextrins} \rightarrow \text{maltose} \rightarrow \text{glucose}$$

In the human body, several enzymes known collectively as amylases degrade starch sequentially into usable glucose units.

Glycogen

Glycogen is the energy reserve carbohydrate of animals. Practically all mammalian cells contain some stored carbohydrates in the form of glycogen, but it is especially abundant in the liver (4%–8% by weight of tissue) and in skeletal muscle cells (0.5%–1.0%). Like starch in plants, glycogen is found as granules in liver and muscle cells. When fasting, animals draw on these glycogen reserves during the first day without food to obtain the glucose needed to maintain metabolic balance.

Glycogen is structurally quite similar to amylopectin, although glycogen is more highly branched (8–12 glucose units between branches) and the branches are shorter. When treated with iodine, glycogen gives a reddish brown colour. Glycogen can be broken down into its D-glucose subunits by acid hydrolysis or by the same enzymes that catalyze the breakdown of starch. In animals, the enzyme phosphorylase catalyzes the breakdown of glycogen to phosphate esters of glucose. About 70% of the total glycogen in the body is stored

in muscle cells. Although the percentage of glycogen (by weight) is higher in the liver, the much greater mass of skeletal muscle stores a greater total amount of glycogen.

Cellulose

Cellulose, a fibrous carbohydrate found in all plants, is the structural component of plant cell walls. Because the earth is covered with vegetation, cellulose is the most abundant of all carbohydrates, accounting for over 50% of all the carbon found in the vegetable kingdom. Cotton fibrils and filter paper are almost entirely cellulose (about 95%), wood is about 50% cellulose, and the dry weight of leaves is about 10%–20% cellulose. The largest use of cellulose is in the manufacture of paper and paper products. Although the use of noncellulose synthetic fibers is increasing, rayon (made from cellulose) and cotton still account for over 70% of textile production.

Like amylose, cellulose is a linear polymer of glucose. It differs, however, in that the glucose units are joined by β-1,4-glycosidic linkages, producing a more extended structure than amylose (part (a) of Figure 28.1w.). This extreme linearity allows a great deal of hydrogen bonding between OH groups on adjacent chains, causing them to pack closely into fibers (part (b) of Figure 28.1w.). As a result, cellulose exhibits little interaction with water or any other solvent. Cotton and wood, for example, are completely insoluble in water and have considerable mechanical strength. Because cellulose does not have a helical structure, it does not bind to iodine to form a coloured product.

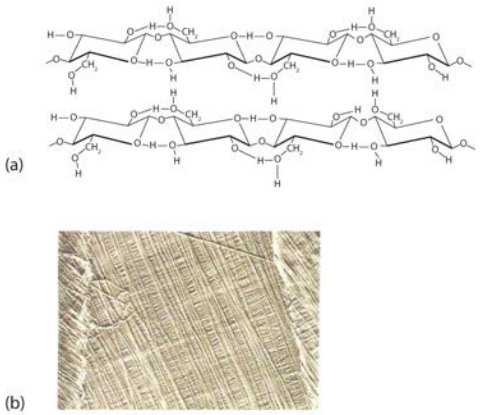


Figure 28.1w. Cellulose. (a) There is extensive hydrogen bonding in the structure of cellulose. (b) In this electron micrograph of the cell wall of an alga, the wall consists of successive layers of cellulose fibers in parallel arrangement (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Cellulose yields D-glucose after complete acid hydrolysis, yet humans are unable to metabolize cellulose as a source of glucose. Our digestive juices lack enzymes that can hydrolyze the β-glycosidic linkages found in cellulose, so although we can eat potatoes, we cannot eat grass. However, certain microorganisms can digest cellulose because they make the enzyme cellulase, which catalyzes the hydrolysis of cellulose. The presence of

these microorganisms in the digestive tracts of herbivorous animals (such as cows, horses, and sheep) allows these animals to degrade the cellulose from plant material into glucose for energy. Termites also contain cellulase-secreting microorganisms and thus can subsist on a wood diet. This example once again demonstrates the extreme stereospecificity of biochemical processes.

Attribution & References

Except where otherwise noted, this page is adapted by Gregory A. Anderson and Samantha Sullivan Sauer from "16: Carbohydrates" Basics of General, Organic, and Biological Chemistry (Ball et al.) by David W. Ball, John W. Hill, and Rhonda J. Scott via LibreTexts, CC BY-NC-SA 4.0./ A LibreTexts version of *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0. Content has been condensed into one page, edited for student understanding and enhanced to match this OER.

Notes

1. Glucose has an aldehyde functional group and multiple alcohol functional groups making it polar and soluble in water. Fructose has a ketone functional group and multiple alcohol functional groups making it polar and soluble in water.

2. The glycosidic linkage between sugars 1 and 2 and the linkage between sugars 2 and 3 are both β because the bonds are directed up from the anomeric carbon in each case.

28.2 LIPIDS

Learning Objectives

By the end of this section, you will be able to:

- Describe the classification of lipids
- Recognize the structures of common fatty acids and classify them as saturated, monounsaturated, or polyunsaturated
- Explain why fats and oils are referred to as triglycerides
- Explain how the fatty acid composition of the triglycerides determines whether a substance is a fat or oil
- Describe the importance of key reactions of triglycerides, such as hydrolysis, hydrogenation, and oxidation
- Identify the distinguishing characteristics of membrane lipids
- Describe membrane components and how they are arranged
- Identify the functions of steroids produced in mammals

Nature's Storehouse of Energy and the Backbone of the Cell

Lipids, a diverse group of organic molecules that includes fats, oils, phospholipids, and steroids, have played a pivotal role in the history and importance of biochemistry. Their significance lies in their multifaceted functions, such as energy storage, cellular membrane structure, and as crucial components of hormones and signaling molecules. The study of lipids has deep roots in biochemistry, dating back to the early 19th century when scientists began unraveling their structure and metabolic roles. Since then, research in lipid biochemistry has greatly expanded, shedding light on their role in health, disease, and cellular regulation, making them a cornerstone of our understanding of biological systems. The exploration of lipids has paved the way for breakthroughs in areas like nutrition, metabolism, and the development of life-saving drugs, underscoring their enduring importance in the field of biochemistry.

Spotlight on Everyday Chemistry: Diet and Metabolism

On July 11, 2003, the Food and Drug Administration amended its food labeling regulations to require that manufacturers list the amount of *trans* fatty acids on Nutrition Facts labels of foods and dietary supplements, effective January 1, 2006. This amendment was a response to published studies demonstrating a link between the consumption of *trans* fatty acids and an increased risk of heart disease. *Trans* fatty acids are produced in the conversion of liquid oils to solid fats, as in the creation of many commercial margarines and shortenings. They have been shown to increase the levels of low-density lipoproteins (LDLs)—complexes that are often referred to as bad cholesterol—in the blood. In this chapter, you will learn about fatty acids and what is meant by a *trans* fatty acid, as well as the difference between fats and oils. You will also learn what cholesterol is and why it is an important molecule in the human body.

Fats and oils, found in many of the foods we eat, belong to a class of biomolecules known as lipids. Gram for gram, they pack more than twice the caloric content of carbohydrates: the oxidation of fats and oils supplies about 9 kcal of energy for every gram oxidized, whereas the oxidation of carbohydrates supplies only 4 kcal/g. Although the high caloric content of fats may be bad news for the dieter, it says something about the efficiency of nature's designs. Our bodies use carbohydrates, primarily in the form of glucose, for our *immediate* energy needs. Our capacity for storing carbohydrates for later use is limited to tucking away a bit of glycogen in the liver or in muscle tissue. We store our *reserve* energy in lipid form, which requires far less space than the same amount of energy stored in carbohydrate form. Lipids have other biological functions besides energy storage. They are a major component of the membranes of the 10 trillion cells in our bodies. They serve as protective padding and insulation for vital organs. Furthermore, without lipids in our diets, we would be deficient in the fat-soluble vitamins A, D, E, and K.

Classification of Lipids

Lipids are not defined by the presence of specific functional groups, as carbohydrates are, but by a physical property—solubility. Compounds isolated from body tissues are classified as lipids if they are more soluble in organic solvents, such as dichloromethane, than in water. By this criterion, the lipid category includes not only fats and oils, which are esters of the trihydroxy alcohol glycerol and fatty acids, but also compounds that incorporate functional groups derived from phosphoric acid, carbohydrates, or amino alcohols, as well as

steroid compounds such as cholesterol (Figure 28.2a. presents one scheme for classifying the various kinds of lipids). We will discuss the various kinds of lipids by considering one subclass at a time and pointing out structural similarities and differences as we go.

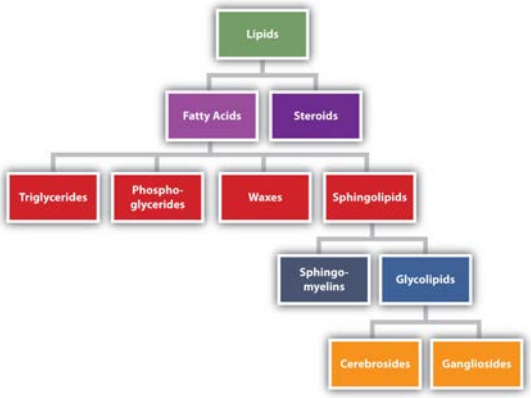


Figure 28.2a. Lipid Organization Based on Structural Relationships. Lipid categorized into fatty acids and steroids. Fatty acids are further separated into triglycerides, phospho-glycerides, waxes, and sphingolipids. Sphingolipids are separated into sphingo-myelins and glycolipids. Glycolipids are separated into cerebrosides and gangliosides (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Fatty Acids

Fatty acids are carboxylic acids that are structural components of fats, oils, and all other categories of lipids, except steroids. More than 70 have been identified in nature. They usually contain an even number of carbon atoms (typically 12–20), are generally unbranched, and can be classified by the presence and number of carbon-to-carbon double bonds. Thus, saturated fatty acids contain no carbon-to-carbon double bonds, monounsaturated fatty acids contain one carbon-to-carbon double bond, and polyunsaturated fatty acids contain two or more carbon-to-carbon double bonds. The presence of carbon-to-carbon double bonds allows for both *cis* and *trans* isomers of fatty acids to exist, as shown in figure 28.2b.

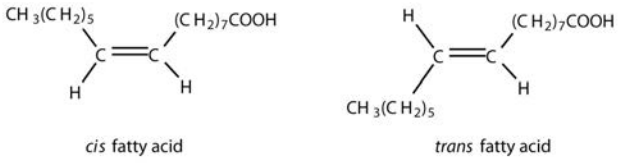


Figure 28.2b. *Cis* fatty acids have two H atoms on the same side of the plane of the fatty acid while *trans* fatty acids have the two H atoms on opposite sides (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Table 28.2a. lists some common fatty acids and one important source for each. The atoms or groups around

the double bonds in unsaturated fatty acids can be arranged in either the *cis* or *trans* isomeric form. Naturally occurring fatty acids are generally in the *cis* configuration.

Table 28.2a. Some Common Fatty Acids Found in Natural Fats

Name	Abbreviated Structural Formula	Condensed Structural Formula	Melting Point (°C)	Source
lauric acid	$C_{11}H_{23}COOH$	$CH_3(CH_2)_{10}COOH$	44	palm kernel oil
myristic acid	$C_{13}H_{27}COOH$	$CH_3(CH_2)_{12}COOH$	58	oil of nutmeg
palmitic acid	$C_{15}H_{31}COOH$	$CH_3(CH_2)_{14}COOH$	63	palm oil
palmitoleic acid	$C_{15}H_{29}COOH$	$CH_3(CH_2)_5CH{=}CH(CH_2)_7COOH$	0.5	macadamia oil
stearic acid	$C_{17}H_{35}COOH$	$CH_3(CH_2)_{16}COOH$	70	cocoa butter
oleic acid	$C_{17}H_{33}COOH$	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$	16	olive oil
linoleic acid	$C_{17}H_{31}COOH$	$CH_3(CH_2)_3(CH_2CH{=}CH)_2(CH_2)_7COOH$	−5	canola oil
α-linolenic acid	$C_{17}H_{29}COOH$	$CH_3(CH_2CH{=}CH)_3(CH_2)_7COOH$	−11	flaxseed
arachidonic acid	$C_{19}H_{31}COOH$	$CH_3(CH_2)_4(CH_2CH{=}CH)_4(CH_2)_2COOH$	−50	liver

Source: "17.1: Fatty Acids" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

Link to Enhanced Learning

Review Compound Interest: A Guide to the Different Types of Fat [New tab] (https://www.compoundchem.com/2015/08/25/fat/).

Two polyunsaturated fatty acids—linoleic and α-linolenic acids—are termed essential fatty acids because humans must obtain them from their diets. Both substances are required for normal growth and development, but the human body does not synthesize them. The body uses linoleic acid to synthesize many of the other unsaturated fatty acids, such as arachidonic acid, a precursor for the synthesis of prostaglandins. In addition, the essential fatty acids are necessary for the efficient transport and metabolism of cholesterol. The average daily diet should contain about 4–6 g of the essential fatty acids.

Prostaglandins

Prostaglandins are chemical messengers synthesized in the cells in which their physiological activity is expressed. They are unsaturated fatty acids containing 20 carbon atoms and are synthesized from arachidonic acid—a polyunsaturated fatty acid—when needed by a particular cell. They are called *prostaglandins* because they were originally isolated from semen found in the prostate gland. It is now known that they are synthesized in nearly all mammalian tissues and affect almost all organs in the body. The five major classes of prostaglandins are designated as PGA, PGB, PGE, PGF, and PGI. Subscripts are attached at the end of these abbreviations to denote the number of double bonds outside the five-carbon ring in a given prostaglandin.

The prostaglandins are among the most potent biological substances known. Slight structural differences give them highly distinct biological effects; however, all prostaglandins exhibit some ability to induce smooth muscle contraction, lower blood pressure, and contribute to the inflammatory response. Aspirin and other nonsteroidal anti-inflammatory agents, such as ibuprofen, obstruct the synthesis of prostaglandins by inhibiting cyclooxygenase, the enzyme needed for the initial step in the conversion of arachidonic acid to prostaglandins.

Their wide range of physiological activity has led to the synthesis of hundreds of prostaglandins and their analogs. Derivatives of PGE_2 are now used in the United States to induce labor (Figure 28.2c.). Other prostaglandins have been employed clinically to lower or increase blood pressure, inhibit stomach secretions, relieve nasal congestion, relieve asthma, and prevent the formation of blood clots, which are associated with heart attacks and strokes.

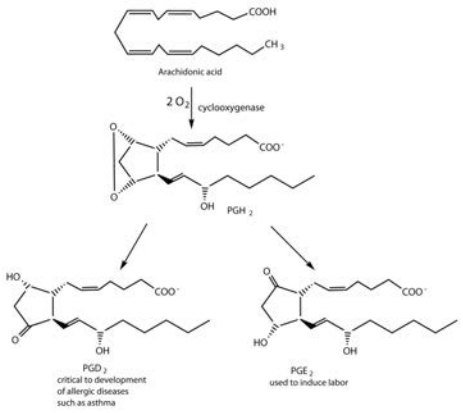


Figure 28.2c. Arachidonic acid can be converted into a prostaglandin by the enzyme cyclooxygenase, which is the enzyme that several drugs including Ibuprofen inhibit. Prostaglandins can then be used as starting materials to to produce other products, such as PGD_2 and PGE_2 (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Although we often draw the carbon atoms in a straight line, they actually have more of a zigzag configuration (part a of Figure 28.2d.). Viewed as a whole, however, the saturated fatty acid molecule is relatively straight

(part b of Figure 28.2d.). Such molecules pack closely together into a crystal lattice, maximizing the strength of dispersion forces and causing fatty acids and the fats derived from them to have relatively high melting points. In contrast, each *cis* carbon-to-carbon double bond in an unsaturated fatty acid produces a pronounced bend in the molecule, so that these molecules do not stack neatly. As a result, the intermolecular attractions of unsaturated fatty acids (and unsaturated fats) are weaker, causing these substances to have lower melting points. Most are liquids at room temperature.

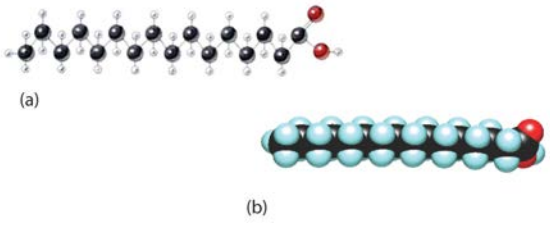


Figure 28.2d. The Structure of Saturated Fatty Acids. (a) There is a zigzag pattern formed by the carbon-to-carbon single bonds in the ball-and-stick model of a palmitic acid molecule. (b) A space-filling model of palmitic acid shows the overall straightness of a saturated fatty acid molecule (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Waxes

Waxes are esters formed from long-chain fatty acids and long-chain alcohols. Most natural waxes are mixtures of such esters. Plant waxes on the surfaces of leaves, stems, flowers, and fruits protect the plant from dehydration and invasion by harmful microorganisms. Carnauba wax, used extensively in floor waxes, automobile waxes, and furniture polish, is largely myricyl cerotate, obtained from the leaves of certain Brazilian palm trees (Figure 28.2e.). Animals also produce waxes that serve as protective coatings, keeping the surfaces of feathers, skin, and hair pliable and water repellent. In fact, if the waxy coating on the feathers of a water bird is dissolved as a result of the bird swimming in an oil slick, the feathers become wet and heavy, and the bird, unable to maintain its buoyancy, drowns.

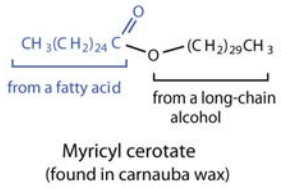


Figure 28.2e. Myricyl cerotate, the main component of carnauba wax, is an ester derived from a fatty acid and a long-chain alcohol. This is an important wax that is found in furniture, automobile, and floor waxes (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Fats and Oils

Fats and oils are the most abundant lipids in nature. They provide energy for living organisms, insulate body organs, and transport fat-soluble vitamins through the blood. Fats and oils are called triglycerides (or triacylcylgerols) because they are esters composed of three fatty acid units joined to glycerol, a trihydroxy alcohol, as seen in Figure 28.2f.

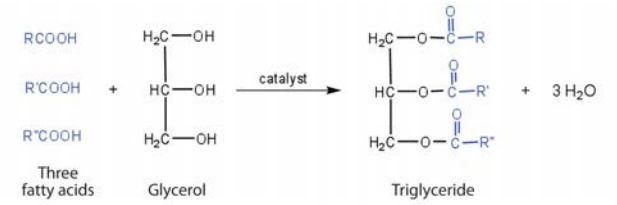


Figure 28.2f. Triglycerides are formed from the addition of fatty acid molecules with the glycerol backbone (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

If all three OH groups on the glycerol molecule are esterified with the same fatty acid, the resulting ester is called a simple triglyceride. Although simple triglycerides have been synthesized in the laboratory, they rarely occur in nature. Instead, a typical triglyceride obtained from naturally occurring fats and oils contains two or three different fatty acid components and is thus termed a mixed triglyceride. Examples of both can be seen in Figure 28.2g.

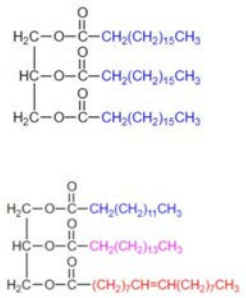


Figure 28.2g. Tristearin, a simple triglyceride is shown at the top. Most naturally occurring triglycerides are mixed triglycerides, such as the one shown on the bottom (Credit: *Basics of GOB (Ball et al.)*, CC BY-NC-SA 4.0)

A triglyceride is called a fat if it is a solid at 25°C; it is called an oil if it is a liquid at that temperature. These differences in melting points reflect differences in the degree of unsaturation and number of carbon atoms in the constituent fatty acids. Triglycerides obtained from animal sources are usually solids, while those of plant origin are generally oils. Therefore, we commonly speak of animal fats and vegetable oils.

No single formula can be written to represent the naturally occurring fats and oils because they are highly

complex mixtures of triglycerides in which many different fatty acids are represented. Table 28.2b. shows the fatty acid compositions of some common fats and oils. The composition of any given fat or oil can vary depending on the plant or animal species it comes from as well as on dietetic and climatic factors. To cite just one example, lard from corn-fed hogs is more highly saturated than lard from peanut-fed hogs. Palmitic acid is the most abundant of the saturated fatty acids, while oleic acid is the most abundant unsaturated fatty acid.

Table 28.2b. Average Fatty Acid Composition of Some Common Fats and Oils (%)*

Fat/Oil	Lauric	Myristic	Palmitic	Stearic	Oleic	Linoleic	Linolenic
Fat – butter (cow)	3	11	27	12	29	2	1
Fat – tallow		3	24	19	43	3	1
Fat – lard		2	26	14	44	10	
canola oil			4	2	62	22	10
coconut oil†	47	18	9	3	6	2	
corn oil			11	2	28	58	1
olive oil			13	3	71	10	1
peanut oil			11	2	48	32	
soybean oil			11	4	24	54	7

***Totals less than 100% indicate the presence of fatty acids with fewer than 12 carbon atoms or more than 18 carbon atoms.**

†**Coconut oil is highly saturated. It contains an unusually high percentage of the low-melting C_8, C_{10}, and C_{12} saturated fatty acids.**

Source: "17.2: Fats and Oils" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

Terms such as saturated fat or unsaturated oil are often used to describe the fats or oils obtained from foods. Saturated fats contain a high proportion of saturated fatty acids, while unsaturated oils contain a high proportion of unsaturated fatty acids. The high consumption of saturated fats is a factor, along with the high consumption of cholesterol, in increased risks of heart disease.

Physical Properties of Fats and Oils

Contrary to what you might expect, pure fats and oils are colourless, odourless, and tasteless. The characteristic colours, odours, and flavours that we associate with some of them are imparted by foreign substances that are lipid soluble and have been absorbed by these lipids. For example, the yellow color of butter is due to the presence of the pigment carotene; the taste of butter comes from two compounds—diacetyl and 3-hydroxy-2-butanone—produced by bacteria in the ripening cream from which the butter is made (Figure 28.2h.).

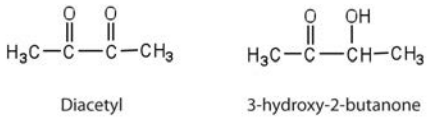


Figure 28.2h. Bacteria produce the molecules diacetyl and 3-hydroxy-2-butanone during the ripening process of cream used to make butter. These molecules, and not the fat, are what give butter its characteristic flavour (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Fats and oils are lighter than water, having densities of about 0.8 g/cm^3. They are poor conductors of heat and electricity and therefore serve as excellent insulators for the body, slowing the loss of heat through the skin.

Chemical Reactions of Fats and Oils

Fats and oils can participate in a variety of chemical reactions—for example, because triglycerides are esters, they can be hydrolyzed in the presence of an acid, a base, or specific enzymes known as lipases. The hydrolysis of fats and oils in the presence of a base is used to make soap and is called saponification. Today most soaps are prepared through the hydrolysis of triglycerides (often from tallow, coconut oil, or both) using water under high pressure and temperature [700 lb/in^2 (~50 atm or 5,000 kPa) and 200°C]. Sodium carbonate or sodium hydroxide is then used to convert the fatty acids to their sodium salts (soap molecules), as seen in Figure 28.2i.

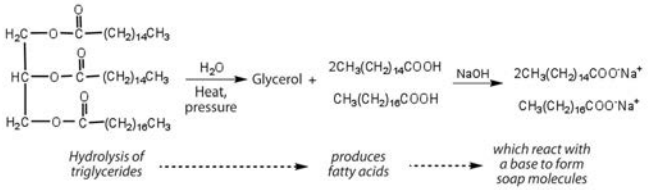


Figure 28.2i. The formation of soaps, a process called saponification, starts with the hydrolysis of triglyceride molecules (which are esters) to produce fatty acids. These fatty acids are then reacted with a base, such as sodium hydroxide, to form soap molecules (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Spotlight on Everyday Chemistry: Soaps

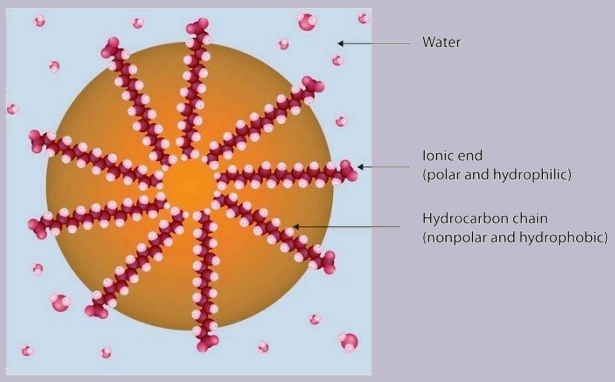


Figure 28.2j. Typical structure of a soap molecule (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Ordinary soap is a mixture of the sodium salts of various fatty acids, produced in one of the oldest organic syntheses practiced by humans (second only to the fermentation of sugars to produce ethyl alcohol). Both the Phoenicians (600 BCE) and the Romans made soap from animal fat and wood ash. Even so, the widespread production of soap did not begin until the 1700s. Soap was traditionally made by treating molten lard or tallow with a slight excess of alkali in large open vats. The mixture was heated, and steam was bubbled through it. After saponification was completed, the soap was precipitated from the mixture by the addition of sodium chloride (NaCl), removed by filtration, and washed several times with water. It was then dissolved in water and reprecipitated by the addition of more NaCl. The glycerol produced in the reaction was also recovered from the aqueous wash solutions.

Pumice or sand is added to produce scouring soap, while ingredients such as perfumes or dyes are added to produce fragrant, coloured soaps. Blowing air through molten soap produces a floating soap. Soft soaps, made with potassium salts, are more expensive but produce a finer lather and are more soluble. They are used in liquid soaps, shampoos, and shaving creams.

Dirt and grime usually adhere to skin, clothing, and other surfaces by combining with body oils, cooking fats, lubricating greases, and similar substances that act like glues. Because these substances are not miscible in water, washing with water alone does little to remove them. Soap removes them, however, because soap molecules have a dual nature. One end, called the *head*, carries an ionic charge (a carboxylate anion) and therefore dissolves in water; the other end, the *tail*, has a hydrocarbon structure and dissolves in oils (Figure 28.2j.). The hydrocarbon tails dissolve in the soil; the ionic heads remain in the aqueous phase, and the soap breaks the oil into tiny soap-

enclosed droplets called *micelles*, which disperse throughout the solution. The droplets repel each other because of their charged surfaces and do not coalesce. With the oil no longer "gluing" the dirt to the soiled surface (skin, cloth, dish), the soap-enclosed dirt can easily be rinsed away.

The double bonds in fats and oils can undergo hydrogenation and also oxidation. The hydrogenation of vegetable oils to produce semisolid fats is an important process in the food industry. Chemically, it is essentially identical to the catalytic hydrogenation reaction described for alkenes, as seen in Figure 28.2k.

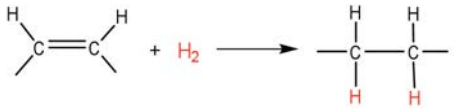


Figure 28.2k. Hydrogenation of an alkene produce an alkane. This same reaction occurs in vegetable oils to produce fats (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

In commercial processes, the number of double bonds that are hydrogenated is carefully controlled to produce fats with the desired consistency (soft and pliable). Inexpensive and abundant vegetable oils (canola, corn, soybean) are thus transformed into margarine and cooking fats. In the preparation of margarine, for example, partially hydrogenated oils are mixed with water, salt, and nonfat dry milk, along with flavouring agents, colouring agents, and vitamins A and D, which are added to approximate the look, taste, and nutrition of butter. (Preservatives and antioxidants are also added.) In most commercial peanut butter, the peanut oil has been partially hydrogenated to prevent it from separating out. Consumers could decrease the amount of saturated fat in their diet by using the original unprocessed oils on their foods, but most people would rather spread margarine on their toast than pour oil on it.

Many people have switched from butter to margarine or vegetable shortening because of concerns that saturated animal fats can raise blood cholesterol levels and result in clogged arteries. However, during the hydrogenation of vegetable oils, an isomerization reaction occurs that produces the *trans* fatty acids. However, studies have shown that *trans* fatty acids also raise cholesterol levels and increase the incidence of heart disease. *Trans* fatty acids do not have the bend in their structures, which occurs in *cis* fatty acids and thus pack closely together in the same way that the saturated fatty acids do. Consumers are now being advised to use polyunsaturated oils and soft or liquid margarine and reduce their total fat consumption to less than 30% of their total calorie intake each day.

Fats and oils that are in contact with moist air at room temperature eventually undergo oxidation and hydrolysis reactions that cause them to turn rancid, acquiring a characteristic disagreeable odour. One cause of the odour is the release of volatile fatty acids by hydrolysis of the ester bonds. Butter, for example, releases

foul-smelling butyric, caprylic, and capric acids. Microorganisms present in the air furnish lipases that catalyze this process. Hydrolytic rancidity can easily be prevented by covering the fat or oil and keeping it in a refrigerator.

Rancidity is a major concern of the food industry, which is why food chemists are always seeking new and better antioxidants, substances added in very small amounts (0.001%–0.01%) to prevent oxidation and thus suppress rancidity. Antioxidants are compounds whose affinity for oxygen is greater than that of the lipids in the food; thus they function by preferentially depleting the supply of oxygen absorbed into the product. Because vitamin E has antioxidant properties, it helps reduce damage to lipids in the body, particularly to unsaturated fatty acids found in cell membrane lipids.

Membranes and Membrane Lipids

All living cells are surrounded by a cell membrane. Plant cells (Figure 28.2l., top) and animal cells (Figure 28.2l., bottom) contain a cell nucleus that is also surrounded by a membrane and holds the genetic information for the cell. Everything between the cell membrane and the nuclear membrane—including intracellular fluids and various subcellular components such as the mitochondria and ribosomes—is called the cytoplasm. The membranes of all cells have a fundamentally similar structure, but membrane function varies tremendously from one organism to another and even from one cell to another within a single organism. This diversity arises mainly from the presence of different proteins and lipids in the membrane.

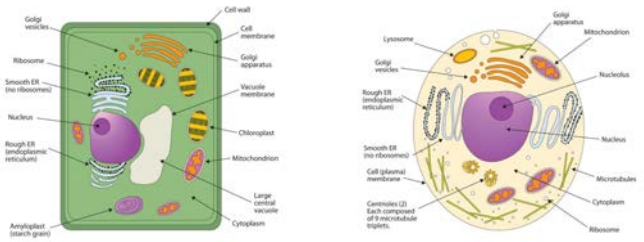


Figure 28.2l. (A) An Idealized Plant Cell. Not all the structures shown here occur in every type of plant cell. (B) An Idealized Animal Cell. The structures shown here will seldom all be found in a single animal cell (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The lipids in cell membranes are highly polar but have dual characteristics: part of the lipid is ionic and therefore dissolves in water, whereas the rest has a hydrocarbon structure and therefore dissolves in nonpolar substances. Often, the ionic part is referred to as hydrophilic, meaning "water loving," and the nonpolar part as hydrophobic, meaning "water fearing" (repelled by water). When allowed to float freely in water, polar lipids spontaneously cluster together in any one of three arrangements: micelles, monolayers, and bilayers, as seen in Figure 28.2m.

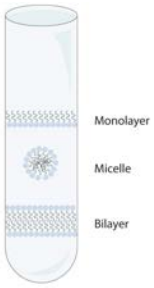


Figure 28.2m. Spontaneously Formed Polar Lipid Structures in Water: Monolayer, Micelle, and Bilayer. (Credit: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Micelles are aggregations in which the lipids' hydrocarbon tails—being hydrophobic—are directed toward the center of the assemblage and away from the surrounding water while the hydrophilic heads are directed outward, in contact with the water. Each micelle may contain thousands of lipid molecules. Polar lipids may also form a monolayer, a layer one molecule thick on the surface of the water. The polar heads face into water, and the nonpolar tails stick up into the air. Bilayers are double layers of lipids arranged so that the hydrophobic tails are sandwiched between an inner surface and an outer surface consisting of hydrophilic heads. The hydrophilic heads are in contact with water on either side of the bilayer, whereas the tails, sequestered inside the bilayer, are prevented from having contact with the water. Bilayers like this make up every cell membrane (Figure 28.2n.).

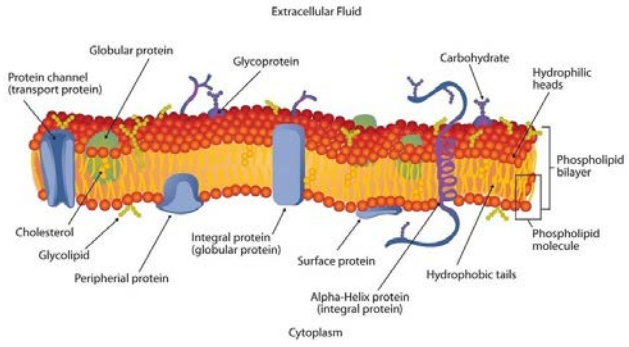


Figure 28.2n. Schematic Diagram of a Cell Membrane. The membrane enclosing a typical animal cell is a phospholipid bilayer with embedded cholesterol and protein molecules. Short oligosaccharide chains are attached to the outer surface. (Credit: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

In the bilayer interior, the hydrophobic tails (that is, the fatty acid portions of lipid molecules) interact by means of dispersion forces. The interactions are weakened by the presence of unsaturated fatty acids. As a result, the membrane components are free to mill about to some extent, and the membrane is described as fluid.

Exercise 28.2a

Using your knowledge of functional groups and their physical properties, explain why membrane lipids have water loving and water fearing regions.

Check Your Answer:[1]

The lipids found in cell membranes can be categorized in various ways. Phospholipids are lipids containing phosphorus. Glycolipids are sugar-containing lipids. The latter are found exclusively on the outer surface of the cell membrane, acting as distinguishing surface markers for the cell and thus serving in cellular recognition and cell-to-cell communication. Sphingolipids are phospholipids or glycolipids that contain the unsaturated amino alcohol sphingosine rather than glycerol. Diagrammatic structures of representative membrane lipids are presented in Figure 28.2o.

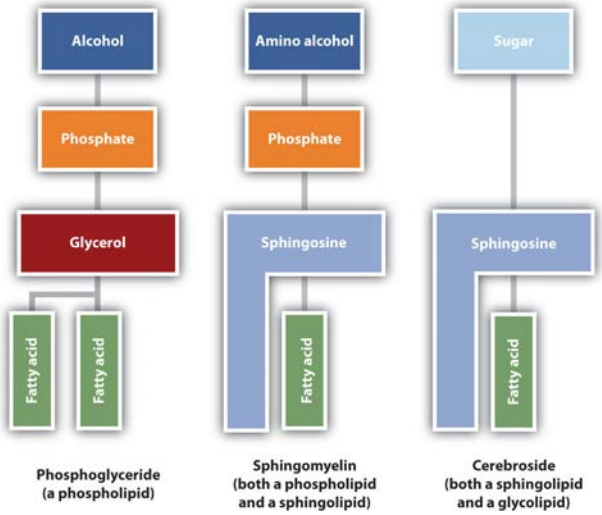


Figure 28.2o. Component Structures of Some Important Membrane Lipids. Phosphoglyceride are composed of alcohol, phosphate, glycerol and 2 fatty acids. A sphingomyelin contains phosphate, sphingosine, and a fatty acid. A cerbroside contains sugar, sphingosine and a fatty acid (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Phosphoglycerides (also known as glycerophospholipids) are the most abundant phospholipids in cell membranes. They consist of a glycerol unit with fatty acids attached to the first two carbon atoms, while a phosphoric acid unit, esterified with an alcohol molecule (usually an amino alcohol, as in part (a) of Figure 28.2p.) is attached to the third carbon atom of glycerol (part (b) of Figure 28.12e.). Notice that the phosphoglyceride molecule is identical to a triglyceride up to the phosphoric acid unit (part (b) of Figure 28.2p.).

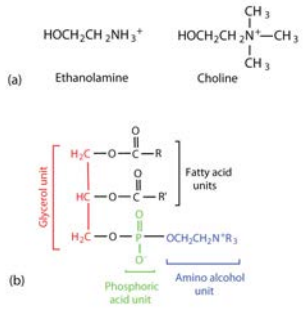


Figure 28.2p. Phosphoglycerides. (a) Amino alcohols are commonly found in phosphoglycerides, which are evident in its structural formula (b). Structural formula of ethanolamine and choline are shown. The structural formula of a phosphoglyceride is shown with the glycerol unit, phosphoric unit, and amino alcohol unit highlighted in different colours (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

There are two common types of phosphoglycerides. Phosphoglycerides containing ethanolamine as the amino

alcohol are called *phosphatidylethanolamines* or *cephalins*. Cephalins are found in brain tissue and nerves and also have a role in blood clotting (Figure 28.2q.). Phosphoglycerides containing choline as the amino alcohol unit are called *phosphatidylcholines* or *lecithins*. Lecithins occur in all living organisms (Figure 28.2q.). Like cephalins, they are important constituents of nerve and brain tissue. Egg yolks are especially rich in lecithins. Commercial-grade lecithins isolated from soybeans are widely used in foods as emulsifying agents. An emulsifying agent is used to stabilize an emulsion—a dispersion of two liquids that do not normally mix, such as oil and water. Many foods are emulsions. Milk is an emulsion of butterfat in water. The emulsifying agent in milk is a protein called *casein*. Mayonnaise is an emulsion of salad oil in water, stabilized by lecithins present in egg yolk.

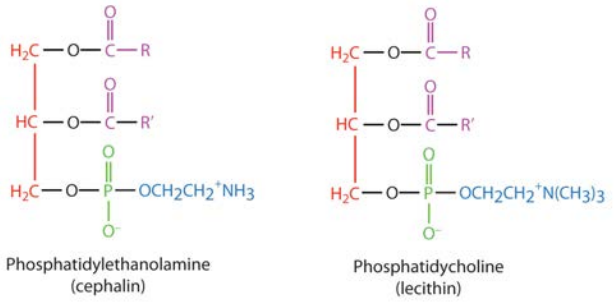


Figure 28.2q. Structures of cephalin and lecithin, two common phosphoglycerides (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Sphingomyelins, the simplest sphingolipids, each contain a fatty acid, a phosphoric acid, sphingosine, and choline (Figure 28.2r.). Because they contain phosphoric acid, they are also classified as phospholipids. Sphingomyelins are important constituents of the myelin sheath surrounding the axon of a nerve cell. Multiple sclerosis is one of several diseases resulting from damage to the myelin sheath.

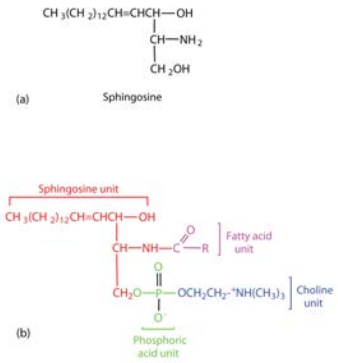


Figure 28.2r. Sphingolipids. (a) Sphingosine, an amino alcohol, is found in all sphingolipids. (b) A sphingomyelin is also known as a phospholipid, as evidenced by the phosphoric acid unit in its structure. Structural formula of sphingosine is shown. The general structure of a sphingolipid is shown with the sphingosine unit, fatty acid unit, phosphoric acid unit, and choline unit highlighted in different colours (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Most animal cells contain sphingolipids called cerebrosides (Figure 28.2s.). Cerebrosides are composed of sphingosine, a fatty acid, and galactose or glucose. They therefore resemble sphingomyelins but have a sugar unit in place of the choline phosphate group. Cerebrosides are important constituents of the membranes of nerve and brain cells.

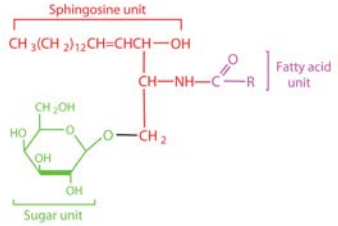


Figure 28.2s. Cerebrosides. Cerebrosides are sphingolipids that contain a sugar unit. General structure of a cerebroside with its sugar unit, sphingosine unit, and fatty acid unit highlighted in different colours (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The sphingolipids called gangliosides are more complex, usually containing a branched chain of three to eight monosaccharides and/or substituted sugars. Because of considerable variation in their sugar components, about 130 varieties of gangliosides have been identified. Most cell-to-cell recognition and communication processes (e.g., blood group antigens) depend on differences in the sequences of sugars in these compounds. Gangliosides are most prevalent in the outer membranes of nerve cells, although they also occur in smaller quantities in the outer membranes of most other cells. Because cerebrosides and gangliosides contain sugar groups, they are also classified as glycolipids.

Membrane Proteins

If membranes were composed only of lipids, very few ions or polar molecules could pass through their hydrophobic "sandwich filling" to enter or leave any cell. However, certain charged and polar species do cross the membrane, aided by proteins that move about in the lipid bilayer. The two major classes of proteins in the cell membrane are integral proteins, which span the hydrophobic interior of the bilayer, and peripheral proteins, which are more loosely associated with the surface of the lipid bilayer (Figure 28.2c.). Peripheral proteins may be attached to integral proteins, to the polar head groups of phospholipids, or to both by hydrogen bonding and electrostatic forces.

Small ions and molecules soluble in water enter and leave the cell by way of channels through the integral proteins. Some proteins, called *carrier proteins*, facilitate the passage of certain molecules, such as hormones and neurotransmitters, by specific interactions between the protein and the molecule being transported.

Steroids

All the lipids discussed so far are saponifiable, reacting with aqueous alkali to yield simpler components, such as glycerol, fatty acids, amino alcohols, and sugars. Lipid samples extracted from cellular material, however, also contain a small but important fraction that does not react with alkali. The most important nonsaponifiable lipids are the steroids. These compounds include the bile salts, cholesterol and related compounds, and certain hormones (such as cortisone and the sex hormones).

Steroids occur in plants, animals, yeasts, and molds but not in bacteria. They may exist in free form or combined with fatty acids or carbohydrates. All steroids have a characteristic structural component consisting of four fused rings. Chemists identify the rings by capital letters and number the carbon atoms as shown in Figure 28.2t. Slight variations in this structure or in the atoms or groups attached to it produce profound differences in biological activity.

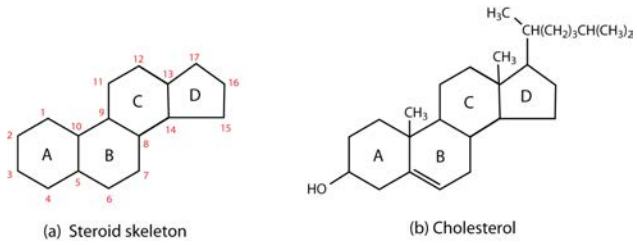


Figure 28.2t. Steroids. (a) The four-fused-ring steroid skeleton uses letter designations for each ring and the numbering of the carbon atoms. (b) The cholesterol molecule follows this pattern (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Cholesterol

Cholesterol (Figure 28.2t., part b) does not occur in plants, but it is the most abundant steroid in the human body (240 g is a typical amount). Excess cholesterol is believed to be a primary factor in the development of atherosclerosis and heart disease, which are major health problems in the United States and Canada today. About half of the body's cholesterol is interspersed in the lipid bilayer of cell membranes. Much of the rest is converted to cholic acid, which is used in the formation of bile salts. Cholesterol is also a precursor in the synthesis of sex hormones, adrenal hormones, and vitamin D.

Excess cholesterol not metabolized by the body is released from the liver and transported by the blood to the gallbladder. Normally, it stays in solution there until being secreted into the intestine (as a component of bile) to be eliminated. Sometimes, however, cholesterol in the gallbladder precipitates in the form of gallstones (Figure 28.2u). Indeed, the name cholesterol is derived from the Greek *chole*, meaning "bile," and *stereos*, meaning "solid."

Figure 28.2u. Numerous small gallstones made up largely of cholesterol, all removed from one patient. Grid scale = 1 mm. (Credit: Photo by George Chernilevsky, CC BY-SA 4.0)

Exercise 28.2b

Is cholesterol water soluble? Explain.

Check Your Answer:[2]

Source: Exercise 28.2b by Samantha Sullivan Sauer, licensed under CC BY-NC 4.0

Cholesterol and Heart Disease

Heart disease is the leading cause of death in the United States and Canada for both men and women. The Centers for Disease Control and Prevention reported that heart disease claimed 631,636 lives in the United States (26% of all reported deaths) in 2006.

Scientists agree that elevated cholesterol levels in the blood, as well as high blood pressure, obesity, diabetes, and cigarette smoking, are associated with an increased risk of heart disease. A long-term investigation by the National Institutes of Health showed that among men ages 30 to 49, the incidence of heart disease was five times greater for those whose cholesterol levels were above 260 mg/100 mL of serum than for those with cholesterol levels of 200 mg/100 mL or less. The cholesterol content of blood varies considerably with age, diet, and sex. Young adults average about 170 mg of cholesterol per 100 mL of blood, whereas males at age 55 may have cholesterol levels at 250 mg/100 mL or higher because the rate of cholesterol breakdown decreases with age. Females tend to have lower blood cholesterol levels than males.

To understand the link between heart disease and cholesterol levels, it is important to understand how cholesterol and other lipids are transported in the body. Lipids, such as cholesterol, are not soluble in water and therefore cannot be transported in the blood (an aqueous medium) unless they are complexed with proteins that are soluble in water, forming assemblages called lipoproteins (Figure 28.2v.). Lipoproteins are classified according to their density, which is dependent on the relative amounts of protein and lipid they contain. Lipids are less dense than proteins, so lipoproteins containing a greater proportion of lipid are less dense than those containing a greater proportion of protein.

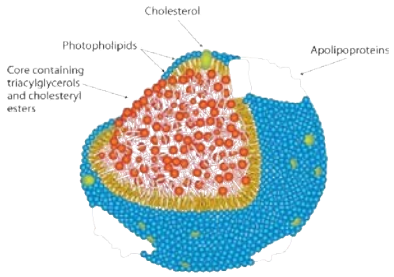


Figure 28.2v. Artist rendition of a typical lipoprotein molecule. (Credit: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Research on cholesterol and its role in heart disease has focused on serum levels of low-density lipoproteins (LDLs) and high-density lipoproteins (HDLs). One of the most fascinating discoveries is that high levels of HDLs reduce a person's risk of developing heart disease, whereas high levels of LDLs increase that risk. Thus the serum LDL:HDL ratio is a better predictor of heart disease risk than the overall level of serum cholesterol. Persons who, because of hereditary or dietary factors, have high LDL:HDL ratios in their blood have a higher incidence of heart disease.

How do HDLs reduce the risk of developing heart disease? No one knows for sure, but one role of HDLs

appears to be the transport of excess cholesterol to the liver, where it can be metabolized. Therefore, HDLs aid in removing cholesterol from blood and from the smooth muscle cells of the arterial wall.

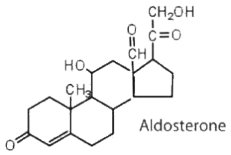
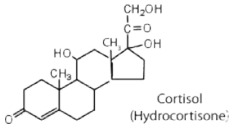
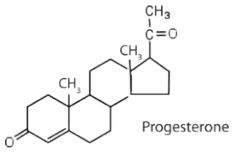
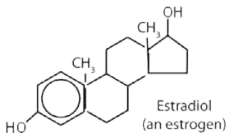
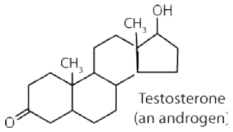
Dietary modifications and increased physical activity can help lower total cholesterol and improve the LDL:HDL ratio. The average American consumes about 600 mg of cholesterol from animal products each day and also synthesizes approximately 1 g of cholesterol each day, mostly in the liver. The amount of cholesterol synthesized is controlled by the cholesterol level in the blood; when the blood cholesterol level exceeds 150 mg/100 mL, the rate of cholesterol biosynthesis is halved. Hence, if cholesterol is present in the diet, a feedback mechanism suppresses its synthesis in the liver. However, the ratio of suppression is not a 1:1 ratio; the reduction in biosynthesis does not equal the amount of cholesterol ingested. Thus, dietary substitutions of unsaturated fat for saturated fat, as well as a reduction in consumption of *trans* fatty acids, is recommended to help lower serum cholesterol and the risk of heart disease.

Steroid Hormones

Hormones are chemical messengers that are released in one tissue and transported through the circulatory system to one or more other tissues. One group of hormones is known as steroid hormones because these hormones are synthesized from cholesterol, which is also a steroid. There are two main groups of steroid hormones: adrenocortical hormones and sex hormones.

The adrenocortical hormones, such as aldosterone and cortisol (Table 28.2c.), are produced by the adrenal gland, which is located adjacent to each kidney. Aldosterone acts on most cells in the body, but it is particularly effective at enhancing the rate of reabsorption of sodium ions in the kidney tubules and increasing the secretion of potassium ions and/or hydrogen ions by the tubules. Because the concentration of sodium ions is the major factor influencing water retention in tissues, aldosterone promotes water retention and reduces urine output. Cortisol regulates several key metabolic reactions (for example, increasing glucose production and mobilizing fatty acids and amino acids). It also inhibits the inflammatory response of tissue to injury or stress. Cortisol and its analogs are therefore used pharmacologically as immunosuppressants after transplant operations and in the treatment of severe skin allergies and autoimmune diseases, such as rheumatoid arthritis.

Table 28.2c. Representative Steroid Hormones and Their Physiological Effects (Image Credits: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Hormone	Effect
 Aldosterone	regulates salt metabolism; stimulates kidneys to retain sodium and excrete potassium
 Cortisol (Hydrocortisone)	stimulates the conversion of proteins to carbohydrates
 Progesterone	regulates the menstrual cycle; maintains pregnancy
 Estradiol (an estrogen)	stimulates female sex characteristics; regulates changes during the menstrual cycle
 Testosterone (an androgen)	stimulates and maintains male sex characteristics

The sex hormones are a class of steroid hormones secreted by the gonads (ovaries or testes), the placenta, and the adrenal glands. Testosterone and androstenedione are the primary male sex hormones, or androgens, controlling the primary sexual characteristics of males, or the development of the male genital organs and the continuous production of sperm. Androgens are also responsible for the development of secondary male characteristics, such as facial hair, deep voice, and muscle strength. Two kinds of sex hormones are of particular importance in females: progesterone, which prepares the uterus for pregnancy and prevents the further release of eggs from the ovaries during pregnancy, and the estrogens, which are mainly responsible for the development of female secondary sexual characteristics, such as breast development and increased deposition of fat tissue in the breasts, the buttocks, and the thighs. Both males and females produce androgens and estrogens, differing in the amounts of secreted hormones rather than in the presence or absence of one or the other.

Sex hormones, both natural and synthetic, are sometimes used therapeutically. For example, a woman who has had her ovaries removed may be given female hormones to compensate. Some of the earliest chemical compounds employed in cancer chemotherapy were sex hormones. For example, estrogens are one treatment option for prostate cancer because they block the release and activity of testosterone. Testosterone enhances prostate cancer growth. Sex hormones are also administered in preparation for sex-change operations, to

promote the development of the proper secondary sexual characteristics. Oral contraceptives are synthetic derivatives of the female sex hormones; they work by preventing ovulation.

Bile Salts

Bile is a yellowish green liquid (pH 7.8–8.6) produced in the liver. The most important constituents of bile are bile salts, which are sodium salts of amide like combinations of bile acids, such as cholic acid (part (a) of Figure 28.2w.) and an amine such as the amino acid glycine (part (b) of Figure 28.2w.). They are synthesized from cholesterol in the liver, stored in the gallbladder, and then secreted in bile into the small intestine. In the gallbladder, the composition of bile gradually changes as water is absorbed and the other components become more concentrated.

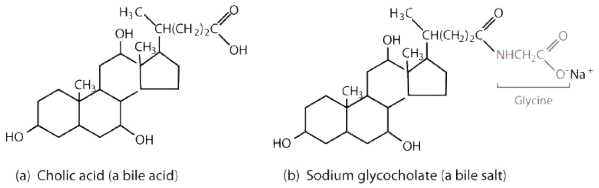


Figure 28.2w. Bile Acids. (a) Cholic acid is an example of a bile acid. (b) Sodium glycocholate is a bile salt synthesized from cholic acid and glycine. (Credit: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Because they contain both hydrophobic and hydrophilic groups, bile salts are highly effective detergents and emulsifying agents; they break down large fat globules into smaller ones and keep those smaller globules suspended in the aqueous digestive environment. Enzymes can then hydrolyze fat molecules more efficiently. Thus, the major function of bile salts is to aid in the digestion of dietary lipids.

Surgical removal is often advised for a gallbladder that becomes infected, inflamed, or perforated. This surgery does not seriously affect digestion because bile is still produced by the liver, but the liver's bile is more dilute and its secretion into the small intestine is not as closely tied to the arrival of food.

Attribution & References

Notes

1. Hydrophilic means water loving. In order to be soluble in water, a structure must be polar. The carboxylic acid ends of lipids are polar and thus will interact with water. Hydrophobic means water fearing. In order to be insoluble in water, a structure must be essentially non-polar. The hydrocarbon ends of lipids are non-polar and thus will not interact with water.

2. Cholesterol is not water soluble. Its structure is predominantly hydrocarbon based and non-polar resulting in very limited (if any) ability to interact with water.

28.3 AMINO ACIDS, PROTEINS, AND ENZYMES

Learning Objectives

By the end of this section, you will be able to:

- Recognize amino acids and classify them based on the characteristics of their side chains
- Explain how an amino acid can act as both an acid and a base
- Explain how a peptide is formed from individual amino acids
- Explain why the sequence of amino acids in a protein is important
- 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.
- Explain the functions of enzymes and how enzymes are classified and named
- Describe the interaction between an enzyme and its substrate

From Structural Integrity to Biological Function

Proteins may be defined as compounds of high molar mass consisting largely or entirely of chains of amino acids. Their masses range from several thousand to several million daltons (Da). In addition to carbon, hydrogen, and oxygen atoms, all proteins contain nitrogen and sulfur atoms, and many also contain phosphorus atoms and traces of other elements. Proteins serve a variety of roles in living organisms and are often classified by these biological roles. Muscle tissue is largely protein, as are skin and hair. Proteins are present in the blood, in the brain, and even in tooth enamel. Each type of cell in our bodies makes its own specialized proteins, as well as proteins common to all or most cells. We begin our study of proteins by looking at the properties and reactions of amino acids, which is followed by a discussion of how amino acids

link covalently to form peptides and proteins. We end the chapter with a discussion of enzymes—the proteins that act as catalysts in the body.

Spotlight on Everyday Chemistry: 1923 Nobel Prize

The 1923 Nobel Prize in Medicine or Physiology was awarded to Frederick Grant Banting and John James Richard Macleod for their discovery of the protein *insulin*. In 1958, the Nobel Prize in Chemistry was awarded to Frederick Sanger for his discoveries concerning the structure of proteins and, in particular, the structure of insulin. What is so important about insulin that two Nobel Prizes have been awarded for work on this protein?

Insulin is a hormone that is synthesized in the pancreas. Insulin stimulates the transport of glucose into cells throughout the body and the storage of glucose as glycogen. People with diabetes do not produce insulin or use it properly. The isolation of insulin in 1921 led to the first effective treatment for these individuals.

Figure 28.3a. Insulin pump, showing an infusion set loaded into spring-loaded insertion device. A reservoir is attached to the infusion set (shown here removed from the pump). (Photo by David-i98, PDM).

Amino Acids

The proteins in all living species, from bacteria to humans, are constructed from the same set of 20 amino acids, so called because each contains an amino group attached to a carboxylic acid. The amino acids in proteins are α-amino acids, which means the amino group is attached to the α-carbon of the carboxylic acid. Humans can synthesize only about half of the needed amino acids; the remainder must be obtained from the diet and are known as essential amino acids. However, two additional amino acids have been found in limited quantities in proteins: Selenocysteine was discovered in 1986, while pyrrolysine was discovered in 2002.

The amino acids are colourless, nonvolatile, crystalline solids, melting and decomposing at temperatures above 200°C. These melting temperatures are more like those of inorganic salts than those of amines or organic acids and indicate that the structures of the amino acids in the solid state and in neutral solution are best represented as having both a negatively charged group and a positively charged group. Such a species is known as a zwitterion (Figure 28.3b.).

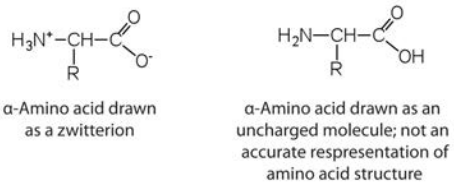


Figure 28.3b. Structural representations of an amino acid in zwitterion form (the more "realistic" state) and as an uncharged molecule (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Classification

In addition to the amino and carboxyl groups, amino acids have a side chain or R group attached to the α-carbon. Each amino acid has unique characteristics arising from the size, shape, solubility, and ionization properties of its R group. As a result, the side chains of amino acids exert a profound effect on the structure and biological activity of proteins. Although amino acids can be classified in various ways, one common approach is to classify them according to whether the functional group on the side chain at neutral pH is nonpolar, polar but uncharged, negatively charged, or positively charged. The structures and names of the 20 amino acids, their one- and three-letter abbreviations, and some of their distinctive features are given in Table 28.3a-d.

Table 28.3a. Common Amino Acids Found in Proteins – Nonpolar R Groups (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Common Name	Abbreviation	Structural Formula (at pH 6)	Molar Mass	Distinctive Feature
glycine	gly (G)		75	the only amino acid lacking a chiral carbon
alanine	ala (A)		89	—
valine	val (V)		117	a branched-chain amino acid
leucine	leu (L)		131	a branched-chain amino acid
isoleucine	ile (I)		131	an essential amino acid because most animals cannot synthesize branched-chain amino acids
phenylalanine	phe (F)		165	also classified as an aromatic amino acid
tryptophan	trp (W)		204	also classified as an aromatic amino acid
methionine	met (M)		149	side chain functions as a methyl group donor
proline	pro (P)		115	contains a secondary amine group; referred to as an *α-imino acid*

Table 28.3b. Common Amino Acids Found in Proteins – Polar but Neutral R Group (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Common Name	Abbreviation	Structural Formula (at pH 6)	Molar Mass	Distinctive Feature
serine	ser (S)		105	found at the active site of many enzymes
threonine	thr (T)		119	named for its similarity to the sugar threose
cysteine	cys (C)		121	oxidation of two cysteine molecules yields *cystine*
tyrosine	tyr (Y)		181	also classified as an aromatic amino acid
asparagine	asn (N)		132	the amide of aspartic acid
glutamine	gln (Q)		146	the amide of glutamic acid

Table 28.3c. Common Amino Acids Found in Proteins – Negatively Charged R Group (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Common Name	Abbreviation	Structural Formula (at pH 6)	Molar Mass	Distinctive Feature
aspartic acid	asp (D)		132	carboxyl groups are ionized at physiological pH; also known as aspartate
glutamic acid	glu (E)		146	carboxyl groups are ionized at physiological pH; also known as glutamate

Table 28.3d. Common Amino Acids Found in Proteins – Positively Charged R Group (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Common Name	Abbreviation	Structural Formula (at pH 6)	Molar Mass	Distinctive Feature
histidine	his (H)		155	the only amino acid whose R group has a pK_a (6.0) near physiological pH
lysine	lys (K)		147	—
arginine	arg (R)		175	almost as strong a base as sodium hydroxide

Link to Enhanced Learning

Review A Brief Guide to the Twenty Common Amino Acids on Compound Interest [New tab] (https://www.compoundchem.com/2014/09/16/aminoacids/)

Exercise 28.3a

Examine the structures in Table 28.3a.

a. For the "Amino acids with a nonpolar R group", identify the main functional groups in the R section and explain why each is non-polar.
b. For the "Amino acids with a polar but neutral R group", identify the main functional groups in the R section and explain why each is polar.

Check Your Answers:[1]

Source: Exercise 28.3a by Samantha Sullivan Sauer, licensed under CC BY-NC 4.0

The first amino acid to be isolated was asparagine in 1806. It was obtained from protein found in asparagus juice (hence the name). Glycine, the major amino acid found in gelatin, was named for its sweet taste (Greek *glykys*, meaning "sweet"). In some cases an amino acid found in a protein is actually a derivative of one of the common 20 amino acids (one such derivative is hydroxyproline, Figure 28.3c.). The modification occurs after the amino acid has been assembled into a protein.

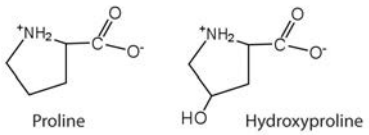


Figure 28.3c. Proline is one of the 20 common amino acids. When modified with the addition of a hydroxyl group, it becomes hydroxyproline. This modification step occurs after the amino acid has been assembled into a protein (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Configuration

Notice in Table 28.3a. that glycine is the only amino acid whose α-carbon is not chiral. Therefore, with the exception of glycine, the amino acids could theoretically exist in either the D- or the L-enantiomeric form and rotate plane-polarized light. As with sugars, chemists used L-glyceraldehyde as the reference compound for the assignment of absolute configuration to amino acids (Figure 28.3d.). Its structure closely resembles an amino acid structure except that in the latter, an amino group takes the place of the OH group on the chiral carbon of the L-glyceraldehyde and a carboxylic acid replaces the aldehyde. Modern stereochemistry assignments using the Cahn-Ingold-Prelog priority rules used ubiquitously in chemistry show that all of the naturally occurring chiral amino acids are S except cysteine which is R.

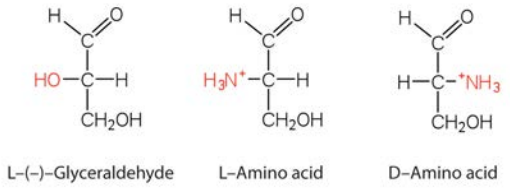


Figure 28.3d. L-glyceraldehyde is the reference compound for the assignment of absolute configuration of amino acids. An example of an L-amino acid and a D-amino acid are given (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

We learned that all naturally occurring sugars belong to the D series. It is interesting, therefore, that nearly all known plant and animal proteins are composed entirely of L-amino acids. However, certain bacteria contain D-amino acids in their cell walls, and several antibiotics (e.g., actinomycin D and the gramicidins) contain varying amounts of D-leucine, D-phenylalanine, and D-valine.

Reactions of Amino Acids

The structure of an amino acid allows it to act as both an acid and a base. An amino acid has this ability because at a certain pH value (different for each amino acid) nearly all the amino acid molecules exist as zwitterions. If acid is added to a solution containing the zwitterion, the carboxylate group captures a hydrogen (H^+) ion, and the amino acid becomes positively charged. If base is added, ion removal of the H^+ ion from the amino group of the zwitterion produces a negatively charged amino acid. In both circumstances (Figure 28.3e.), the amino acid acts to maintain the pH of the system—that is, to remove the added acid (H^+) or base (OH^-) from solution.

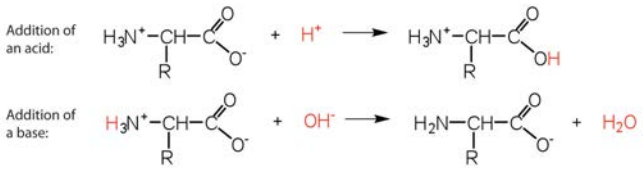


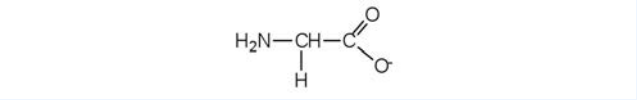
Figure 28.3e. Amino acids naturally exist as zwitterions. As a result, the addition of either an acid or a base allows the amino acid to itself act as a base or an acid, respectively (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Example 28.3a

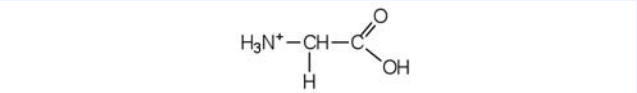
a. Draw the structure for the anion formed when glycine (at neutral pH) reacts with a base.
b. Draw the structure for the cation formed when glycine (at neutral pH) reacts with an acid.

Solution

a. The base removes H^+ from the protonated amine group.



b. The acid adds H^+ to the carboxylate group.



Example & image source: Introduction to Chemistry: GOB (v. 1.0), CC BY-NC-SA 3.0

The particular pH at which a given amino acid exists in solution as a zwitterion is called the isoelectric point (pI). At its pI, the positive and negative charges on the amino acid balance, and the molecule as a whole is electrically neutral. The amino acids whose side chains are always neutral have isoelectric points ranging from 5.0 to 6.5. The basic amino acids (which have positively charged side chains at neutral pH) have relatively high examples. Acidic amino acids (which have negatively charged side chains at neutral pH) have quite low examples (Table 28.3e.).

Table 28.3e. ExampIes of Some Representative Amino Acids

Amino Acid	Classification	pI
alanine	nonpolar	6.0
valine	nonpolar	6.0
serine	polar, uncharged	5.7
threonine	polar, uncharged	6.5
arginine	positively charged (basic)	10.8
histidine	positively charged (basic)	7.6
lysine	positively charged (basic)	9.8
aspartic acid	negatively charged (acidic)	3.0
glutamic acid	negatively charged (acidic)	3.2

Source: "18.2: Reactions of Amino Acids" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

Amino acids undergo reactions characteristic of carboxylic acids and amines. The reactivity of these functional groups is particularly important in linking amino acids together to form peptides and proteins, as you will see later in this chapter. Simple chemical tests that are used to detect amino acids take advantage of the reactivity of these functional groups. An example is the ninhydrin test in which the amine functional group of α-amino acids reacts with ninhydrin to form purple-coloured compounds. Ninhydrin is used to detect fingerprints because it reacts with amino acids from the proteins in skin cells transferred to the surface by the individual leaving the fingerprint (Figure 28.3f.).

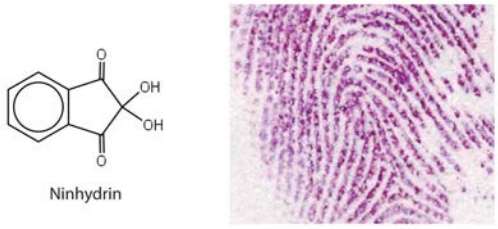


Figure 28.3f. The ninhydrin test is used to detect fingerprints on samples. Ninhydrin reacts with the amine functional group present in a amino acids in the proteins of skin cells to form a purple-coloured compound (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Peptides

Two or more amino acids can join together into chains called peptides. We have discussed the reaction

between ammonia and a carboxylic acid to form an amide. In a similar reaction, the amino group on one amino acid molecule reacts with the carboxyl group on another, releasing a molecule of water and forming an amide linkage, as seen in Figure 28.3g.

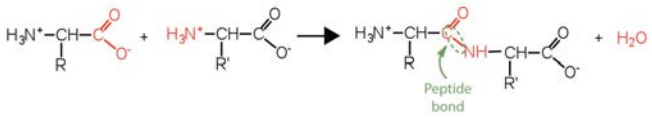


Figure 28.3g. When two amino acids react to form a peptide bond, this linkage occurs between the amino group of one amino acid and the carboxyl group of another, forming an amide linkage with the release of a molecule of water (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

An amide bond joining two amino acid units is called a peptide bond. Note that the product molecule still has a reactive amino group on the left and a reactive carboxyl group on the right. These can react with additional amino acids to lengthen the peptide. The process can continue until thousands of units have joined, resulting in large proteins, as seen in Figure 28.3h.

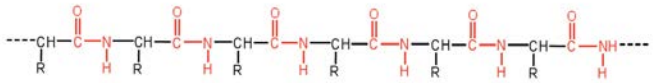


Figure 28.3h. The continued reaction between the carboxyl group of a growing chain of amino acids with the amino group of additional amino acids leads to many amino acids becoming linked, forming a peptide (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

A chain consisting of only two amino acid units is called a dipeptide; a chain consisting of three is a tripeptide. By convention, peptide and protein structures are depicted with the amino acid whose amino group is free (the N-terminal end) on the left and the amino acid with a free carboxyl group (the C-terminal end) to the right, as shown in Figure 28.3i.

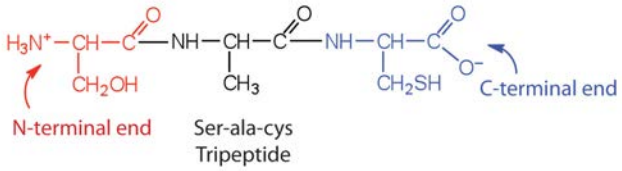


Figure 28.3i. A typical protein structure, in which the free amino group (N-terminal end) is depicted as being on the left and the free carboxyl group (C-terminal end) is depicted as being on the right (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The general term peptide refers to an amino acid chain of unspecified length. However, chains of about 50 amino acids or more are usually called proteins or polypeptides. In its physiologically active form, a protein

may be composed of one or more polypeptide chains. As an example, consider the 3-dimensional shape of the protein bradykinin, a protein important for vasodilation in the body (Figure 28.3j.):

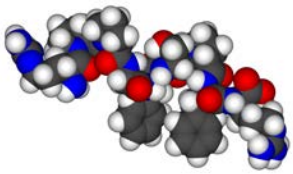


Figure 28.3j. Space-filling model of bradykinin. (Photo by Fvasconcellos, PDM)

For peptides and proteins to be physiologically active, it is not enough that they incorporate certain amounts of specific amino acids. The order, or sequence, in which the amino acids are connected is also of critical importance. Bradykinin is a nine-amino acid peptide (Figure 28.3j.) produced in the blood that has the following amino acid sequence:

arg-pro-pro-gly-phe-ser-pro-phe-arg

This peptide lowers blood pressure, stimulates smooth muscle tissue, increases capillary permeability, and causes pain. When the order of amino acids in bradykinin is reversed,

arg-phe-pro-ser-phe-gly-pro-pro-arg

the peptide resulting from this synthesis shows none of the activity of bradykinin.

Just as millions of different words are spelled with our 26-letter English alphabet, millions of different proteins are made with the 20 common amino acids. However, just as the English alphabet can be used to write gibberish, amino acids can be put together in the wrong sequence to produce nonfunctional proteins. Although the correct sequence is ordinarily of utmost importance, it is not always absolutely required. Just as you can sometimes make sense of incorrectly spelled English words, a protein with a small percentage of "incorrect" amino acids may continue to function. However, it rarely functions as well as a protein having the correct sequence. There are also instances in which seemingly minor errors of sequence have disastrous effects. For example, in some people, every molecule of hemoglobin (a protein in the blood that transports oxygen) has a single incorrect amino acid unit out of about 300 (a single valine replaces a glutamic acid). That "minor" error is responsible for sickle cell anemia, an inherited condition that usually is fatal.

Proteins

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.

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 283k.

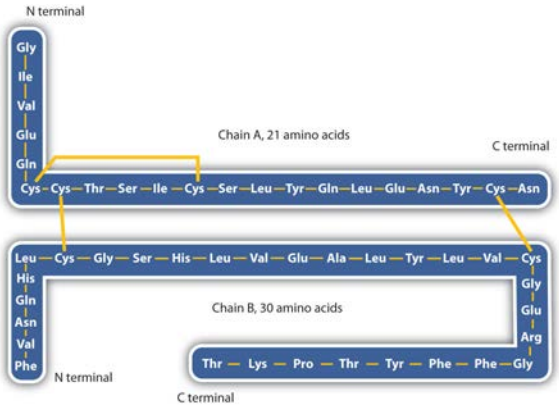


Figure 28.3k. Primary Structure of Human Insulin. Human insulin, whose amino acid sequence is shown here, is a hormone that is required for the proper metabolism of glucose (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

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 28.3l.). 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.

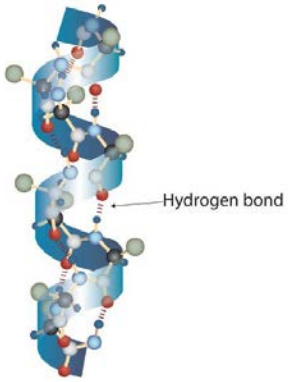


Figure 28.3l. A Ball-and-Stick Model of an α-Helix. This ball-and-stick model shows the intrachain hydrogen bonding between carbonyl oxygen atoms and amide hydrogen atoms. Each turn of the helix spans 3.6 amino acids. Note that the side chains (represented as green spheres) point out from the helix. (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*CC BY-NC-SA 4.0)

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 28.3m.). 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.

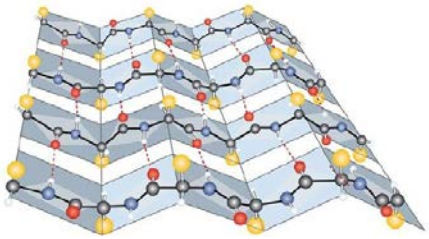


Figure 28.3m. A Ball-and-Stick Model of the β-Pleated Sheet Structure in Proteins. The side chains extend above or below the sheet and alternate along the chain. The protein chains are held together by interchain hydrogen bonding. (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

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 28.3n. shows a depiction of the three-dimensional structure of insulin.

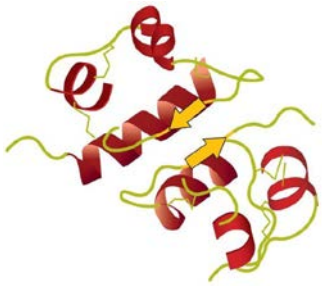


Figure 28.3n. A Ribbon Model of the Three-Dimensional Structure of Insulin. The spiral regions represent sections of the polypeptide chain that have an α-helical structure, while the broad arrows represent β-pleated sheet structures. (Credit: Introduction to Chemistry: GOB (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Four major types of attractive interactions determine the shape and stability of the tertiary structure of proteins. These are shown in Figure 28.3o.

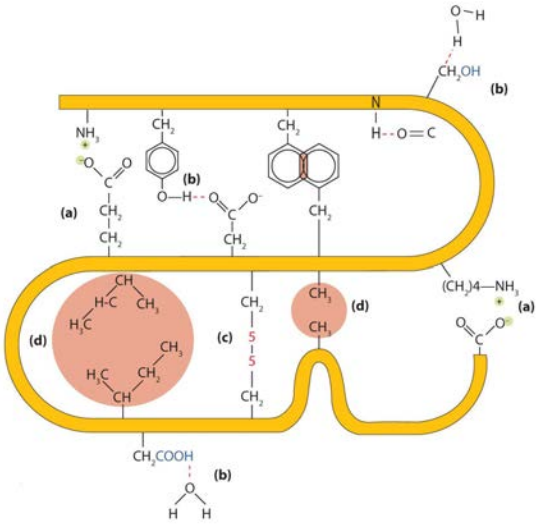


Figure 28.3o. Tertiary Protein Structure Interactions. Four interactions stabilize the tertiary structure of a protein: (a) ionic bonding, (b) hydrogen bonding, (c) disulfide linkages, and (d) dispersion forces (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

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 28.3o.).
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 28.3o.).
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 28.3o. and figure 28.3p.). Intrachain disulfide linkages are found in many proteins, including insulin (yellow bars in Figure 28.3k.) and have a strong stabilizing effect on the tertiary structure.

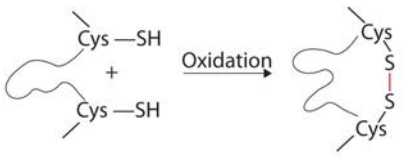


Figure 28.3p. Formation of disulfide linkages between two cystine amino acid units (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

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 neighbouring nonpolar atom. Dispersion forces are weak but can be important when other types of interactions are either missing or minimal (part (d) of Figure 28.3o.). 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.

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 28.3q.). The quaternary structure of a protein is produced and stabilized by the same kinds of interactions that produce and maintain the tertiary structure.

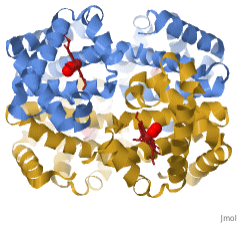


Figure 28.3q. The Quaternary Structure of Hemoglobin. Hemoglobin is a protein that transports oxygen throughout the body. (Image by Theislikerice, CC BY-SA 4.0)

A schematic representation of the four levels of protein structure is shown in Figure 28.3r. 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.

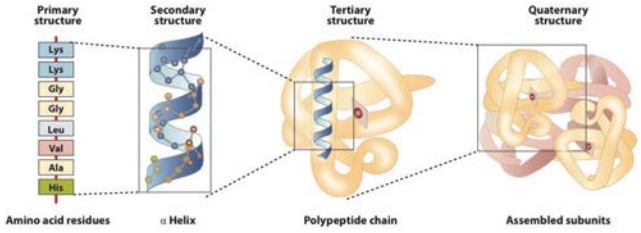


Figure 28.3r. Levels of Structure in Proteins.(Credit: Introduction to Chemistry: General, Organic, and Biological (v. 1.0), edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

Denaturation of Proteins

The highly organized structures of proteins are truly masterworks of chemical architecture. But highly organized structures tend to have a certain delicacy, and this is true of proteins. Denaturation is the term used for any change in the three-dimensional structure of a protein that renders it incapable of performing its assigned function. A denatured protein cannot do its job. (Sometimes denaturation is equated with the precipitation or coagulation of a protein; our definition is a bit broader.) A wide variety of reagents and conditions, such as heat, organic compounds, pH changes, and heavy metal ions can cause protein denaturation (Table 28.3f.).

Table 28.3f. 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.

Source: "18.4: Proteins" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

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 28.3s.). 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.

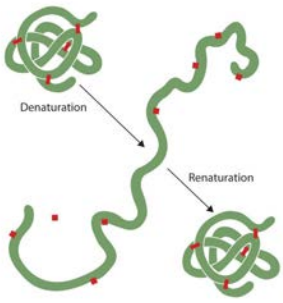


Figure 28.3s. Denaturation and Renaturation of a Protein. The denaturation (unfolding) and renaturation (refolding) of a protein is depicted. The red boxes represent stabilizing interactions, such as disulfide linkages, hydrogen bonding, and/or ionic bonds (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

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.

Enzymes

A catalyst is any substance that increases the rate or speed of a chemical reaction without being changed or consumed in the reaction. Enzymes are biological catalysts, and nearly all of them are proteins. The reaction rates attained by enzymes are truly amazing. In their presence, reactions occur at rates that are a million (10^6) or more times faster than would be attainable in their absence. What is even more amazing is that enzymes perform this function at body temperature (~37°C) and physiological pH (pH ~7), rather than at the conditions that are typically necessary to increase reaction rates (high temperature or pressure, the use of strong oxidizing or reducing agents or strong acids or bases, or a combination of any of these). In addition, enzymes are highly specific in their action; that is, each enzyme catalyzes only one type of reaction in only one compound or a group of structurally related compounds. The compound or compounds on which an enzyme acts are known as its substrates.

Hundreds of enzymes have been purified and studied in an effort to understand how they work so effectively and with such specificity. The resulting knowledge has been used to design drugs that inhibit or activate particular enzymes. An example is the intensive research to improve the treatment of or find a cure for acquired immunodeficiency syndrome (AIDS). AIDS is caused by the human immunodeficiency virus (HIV). Researchers are studying the enzymes produced by this virus and are developing drugs intended to

block the action of those enzymes without interfering with enzymes produced by the human body. Several of these drugs have now been approved for use by AIDS patients.

The first enzymes to be discovered were named according to their source or method of discovery. The enzyme *pepsin*, which aids in the hydrolysis of proteins, is found in the digestive juices of the stomach (Greek *pepsis*, meaning "digestion"). *Papain*, another enzyme that hydrolyzes protein (in fact, it is used in meat tenderizers), is isolated from papayas. As more enzymes were discovered, chemists recognized the need for a more systematic and chemically informative identification scheme. In the current numbering and naming scheme, under the oversight of the Nomenclature Commission of the International Union of Biochemistry, enzymes are arranged into six groups according to the general type of reaction they catalyze (Table 28.3g.), with subgroups and secondary subgroups that specify the reaction more precisely.

Table 28.3g. Classes of Enzymes

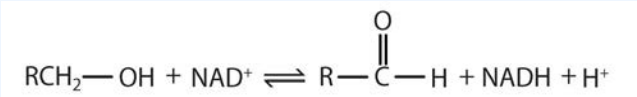
Class	Type of Reaction Catalyzed	Examples
oxidoreductases	oxidation-reduction reactions	Dehydrogenases catalyze oxidation-reduction whic
transferases	transfer reactions of groups, such as methyl, amino, and acetyl	Transaminases catalyze the transfer of amin
hydrolases	hydrolysis reactions	Lipases catalyze the hydrolysis of li
lyases	reactions in which groups are removed without hydrolysis or addition of groups to a double bond	Decarboxylases ca
isomerases	reactions in which a compound is converted to its isomer	Isomerases may catalyze the conversion of a functional group is transfe
ligases	reactions in which new bonds are formed between carbon and another atom; energy is required	Synthetases catalyze reactions in whi

Source: "18.5: Enzymes" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

Each enzyme is assigned a four-digit number, preceded by the prefix EC—for enzyme classification—that indicates its group, subgroup, and so forth. This is demonstrated in Table 28.19b. for alcohol dehydrogenase. Each enzyme is also given a name consisting of the root of the name of its substrate or substrates and the *–ase* suffix. Thus urease is the enzyme that catalyzes the hydrolysis of urea.

Example 28.3b

Alcohol dehydrogenase has an Enzyme Classification Number of EC 1.1.1.1. What does this mean?



(credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Solution:

The first digit indicates that this enzyme is an oxidoreductase; that is, an enzyme that catalyzes an oxidation-reduction reaction. The second digit indicates that this oxidoreductase catalyzes a reaction involving a primary or secondary alcohol. The third digit indicates that either the coenzyme NAD^+ or $NADP^+$ is required for this reaction. The fourth digit indicates that this was the first enzyme isolated, characterized, and named using this system of nomenclature. The systematic name for this enzyme is *alcohol:NAD^+ oxidoreductase*, while the recommended or common name is alcohol dehydrogenase.

Enzyme Action

Enzyme-catalyzed reactions occur in at least two steps. In the first step, an enzyme molecule (E) and the substrate molecule or molecules (S) collide and react to form an intermediate compound called the *enzyme-substrate* (E–S) *complex*. (This step is reversible because the complex can break apart into the original substrate or substrates and the free enzyme.) Once the E–S complex forms, the enzyme is able to catalyze the formation of product (P), which is then released from the enzyme surface:

$$S + E \rightarrow E\text{–}S \qquad (1)$$

$$E\text{–}S \rightarrow P + E \qquad (2)$$

Hydrogen bonding and other electrostatic interactions hold the enzyme and substrate together in the complex. The structural features or functional groups on the enzyme that participate in these interactions are located in a cleft or pocket on the enzyme surface. This pocket, where the enzyme combines with the substrate and transforms the substrate to product is called the active site of the enzyme (Figure 28.3t.).

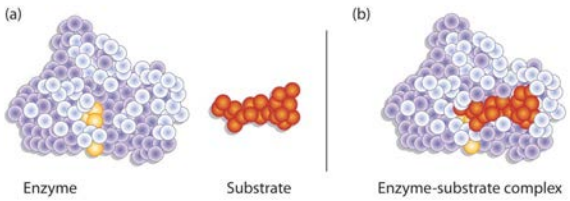


Figure 28.3t. Substrate Binding to the Active Site of an Enzyme. The enzyme dihydrofolate reductase is shown with one of its substrates: NADP+ (a) unbound and (b) bound. The NADP+ (shown in red) binds to a pocket that is complementary to it in shape and ionic properties (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The active site of an enzyme possesses a unique conformation (including correctly positioned bonding groups) that is complementary to the structure of the substrate, so that the enzyme and substrate molecules fit together in much the same manner as a key fits into a tumbler lock. In fact, an early model describing the formation of the enzyme-substrate complex was called the lock-and-key model (Figure 28.3u.). This model portrayed the enzyme as conformationally rigid and able to bond only to substrates that exactly fit the active site.

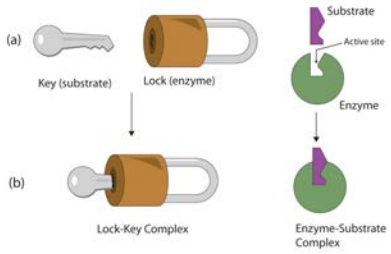


Figure 28.3u. The Lock-and-Key Model of Enzyme Action. (a) Because the substrate and the active site of the enzyme have complementary structures and bonding groups, they fit together as a key fits a lock. (b) The catalytic reaction occurs while the two are bonded together in the enzyme-substrate complex (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Working out the precise three-dimensional structures of numerous enzymes has enabled chemists to refine the original lock-and-key model of enzyme actions. They discovered that the binding of a substrate often leads to a large conformational change in the enzyme, as well as to changes in the structure of the substrate or substrates. The current theory, known as the induced-fit model, says that enzymes can undergo a change in conformation when they bind substrate molecules, and the active site has a shape complementary to that of the substrate only after the substrate is bound, as shown for hexokinase in Figure 28.3v. After catalysis, the enzyme resumes its original structure.

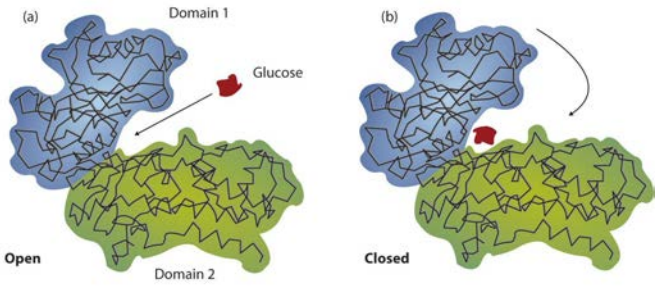


Figure 28.3v. The Induced-Fit Model of Enzyme Action. (a) The enzyme hexokinase without its substrate (glucose, shown in red) is bound to the active site. (b) The enzyme conformation changes dramatically when the substrate binds to it, resulting in additional interactions between hexokinase and glucose (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The structural changes that occur when an enzyme and a substrate join together bring specific parts of a substrate into alignment with specific parts of the enzyme's active site. Amino acid side chains in or near the binding site can then act as acid or base catalysts, provide binding sites for the transfer of functional groups from one substrate to another or aid in the rearrangement of a substrate. The participating amino acids, which are usually widely separated in the primary sequence of the protein, are brought close together in the active site as a result of the folding and bending of the polypeptide chain or chains when the protein acquires its tertiary and quaternary structure. Binding to enzymes brings reactants close to each other and aligns them properly, which has the same effect as increasing the concentration of the reacting compounds.

Example 28.3c

a. What type of interaction would occur between an OH group present on a substrate molecule and a functional group in the active site of an enzyme?
b. Suggest an amino acid whose side chain might be in the active site of an enzyme and form the type of interaction you just identified.

Solution

a. An OH group would most likely engage in hydrogen bonding with an appropriate functional group present in the active site of an enzyme.
b. Several amino acid side chains would be able to engage in hydrogen bonding with an OH group. One example would be asparagine, which has an amide functional group.

One characteristic that distinguishes an enzyme from all other types of catalysts is its substrate specificity. An inorganic acid such as sulfuric acid can be used to increase the reaction rates of many different reactions, such

as the hydrolysis of disaccharides, polysaccharides, lipids, and proteins, with complete impartiality. In contrast, enzymes are much more specific. Some enzymes act on a single substrate, while other enzymes act on any of a group of related molecules containing a similar functional group or chemical bond. Some enzymes even distinguish between D- and L-stereoisomers, binding one stereoisomer but not the other. Urease, for example, is an enzyme that catalyzes the hydrolysis of a single substrate—urea—but not the closely related compounds methyl urea, thiourea, or biuret (Figure 28.3w.). The enzyme carboxypeptidase, on the other hand, is far less specific. It catalyzes the removal of nearly any amino acid from the carboxyl end of any peptide or protein.

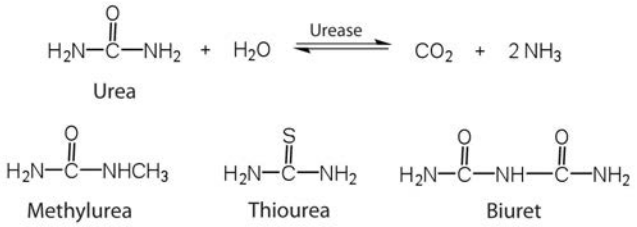


Figure 28.3w. Urease catalyzes the hydrolysis of urea into carbon dioxide and ammonia. Urease is specific to urea; it will not catalyze the hydrolysis of closely-related molecules such as methylurea, thiourea, nor biuret (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Enzyme specificity results from the uniqueness of the active site in each different enzyme because of the identity, charge, and spatial orientation of the functional groups located there. It regulates cell chemistry so that the proper reactions occur in the proper place at the proper time. Clearly, it is crucial to the proper functioning of the living cell.

Attribution & References

Except where otherwise noted, portions of this page were written by Gregory A. Anderson while others were adapted by Gregory A. Anderson and Samantha Sullivan Sauer from "18.0: Prelude to Amino Acids, Proteins, and Enzymes", "18.1: Properties of Amino Acids", "18.2: Reactions of Amino Acids", "18.3: Peptides", "18.4: Proteins", "18.5: Enzymes", & "18.6: Enzyme Action" *Basics of General, Organic, and Biological Chemistry (Ball et al.)* by David W. Ball, John W. Hill, and Rhonda J. Scott via LibreTexts, CC BY-NC-SA 4.0./ A LibreTexts version of *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0.

Notes

1. a. glycine, alanine, valine, leucine, and isoleucine all have alkane based R groups which are non-polar due to the carbon-hydrogen components. Phenylalanine is aromatic and thus non-polar. Tryptophane and proline have amine and hydrocarbon R groups which

are non-polar. Methionine has a S-C based R group and is non-polar as the electronegativity of S is the same as C.

b. Serine, threonine and tyrosine have alcohol functional groups in their R groups. Oxygen is electronegative resulting in a polar region. Cysteine has a thiol group and S is electronegative compared to H. Asparagine and glutamine have carbonyl groups and the O makes the region polar.

28.4 NUCLEIC ACIDS AND DNA

Learning Objectives

By the end of this section, you will be able to:

- Identify the different molecules that combine to form nucleotides
- Identify the two types of nucleic acids and the function of each type
- Describe how nucleotides are linked together to form nucleic acids
- Describe the secondary structure of DNA and the importance of complementary base pairing
- Describe how a new copy of DNA is synthesized
- Describe how RNA is synthesized from DNA
- Identify the different types of RNA and the function of each type of RNA
- Describe the characteristics of the genetic code
- Describe how a protein is synthesized from mRNA
- Describe the causes of genetic mutations and how they lead to genetic diseases

The Key to Heredity

The blueprint for the reproduction and the maintenance of each organism is found in the nuclei of its cells, concentrated in elongated, threadlike structures called chromosomes. These complex structures, consisting of DNA and proteins, contain the basic units of heredity, called genes. The number of chromosomes (and genes) varies with each species. Human body cells have 23 pairs of chromosomes having 20,000–40,000 different genes.

Sperm and egg cells contain only a single copy of each chromosome; that is, they contain only one member of each chromosome pair. Thus, in sexual reproduction, the entire complement of chromosomes is achieved only when an egg and sperm combine. A new individual receives half its hereditary material from each parent. Calling the unit of heredity a "gene" merely gives it a name. But what really are genes and how is the information they contain expressed? One definition of a gene is that it is a segment of DNA that constitutes the code for a specific polypeptide. If genes are segments of DNA, we need to learn more about the structure

and physiological function of DNA. We begin by looking at the small molecules needed to form DNA and RNA (ribonucleic acid)—the nucleotides.

Spotlight on Everyday Chemistry: The Birth of Genetic Engineering

Figure 28.4a. A vial of insulin. It has been given a trade name, Actrapid, by the manufacturer. (Photo by Mr Hyde, PDM)

Following the initial isolation of insulin in 1921, diabetic patients could be treated with insulin obtained from the pancreases of cattle and pigs. Unfortunately, some patients developed an allergic reaction to this insulin because its amino acid sequence was not identical to that of human insulin. In the 1970s, an intense research effort began that eventually led to the production of genetically engineered human insulin—the first genetically engineered product to be approved for medical use. To accomplish this feat, researchers first had to determine how insulin is made in the body and then find a way of causing the same process to occur in nonhuman organisms, such as bacteria or yeast cells.

Nucleotides

The repeating, or monomer, units that are linked together to form nucleic acids are known as nucleotides. The deoxyribonucleic acid (DNA) of a typical mammalian cell contains about 3×10^9 nucleotides. Nucleotides can be further broken down to phosphoric acid (H_3PO_4), a pentose sugar (a sugar with five carbon atoms), and a nitrogenous base (a base containing nitrogen atoms).

$$\text{nucleic acids} \xrightarrow[\text{down into}]{\text{can be broken}} \text{nucleotides} \xrightarrow[\text{down into}]{\text{can be broken}} H_3PO_4 + \text{nitrogen base} + \text{pentose sugar}$$

If the pentose sugar is ribose, the nucleotide is more specifically referred to as a *ribonucleotide*, and the resulting nucleic acid is ribonucleic acid (RNA). If the sugar is 2-deoxyribose, the nucleotide is a *deoxyribonucleotide*, and the nucleic acid is DNA, as shown in Figure 28.4b.

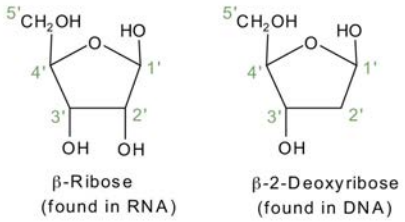


Figure 28.4b. Backbone structure of both ribose and deoxyribose (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The nitrogenous bases found in nucleotides are classified as pyrimidines or purines. Pyrimidines are heterocyclic amines with two nitrogen atoms in a six-member ring and include uracil, thymine, and cytosine. Purines are heterocyclic amines consisting of a pyrimidine ring fused to a five-member ring with two nitrogen atoms. Adenine and guanine are the major purines found in nucleic acids (Figure 28.4c.).

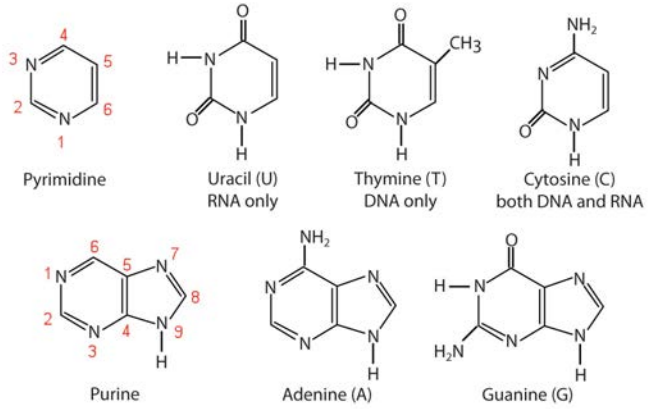


Figure 28.4c. The nitrogenous bases found in DNA and RNA (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The formation of a bond between C1′ of the pentose sugar and N1 of the pyrimidine base or N9 of the purine base joins the pentose sugar to the nitrogenous base. In the formation of this bond, a molecule of water is removed. Table 28.4a. summarizes the similarities and differences in the composition of nucleotides in DNA and RNA. The numbering convention is that primed numbers designate the atoms of the pentose ring, and unprimed numbers designate the atoms of the purine or pyrimidine ring.

Table 28.4a. Composition of Nucleotides in DNA and RNA

Composition	DNA	RNA
purine bases	adenine and guanine	adenine and guanine
pyrimidine bases	cytosine and thymine	cytosine and uracil
pentose sugar	2-deoxyribose	ribose
inorganic acid	phosphoric acid (H_3PO_4)	H_3PO_4

Source: "19.1: Nucleotides" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

The names and structures of the major ribonucleotides and one of the deoxyribonucleotides are given in Figure 28.4d.

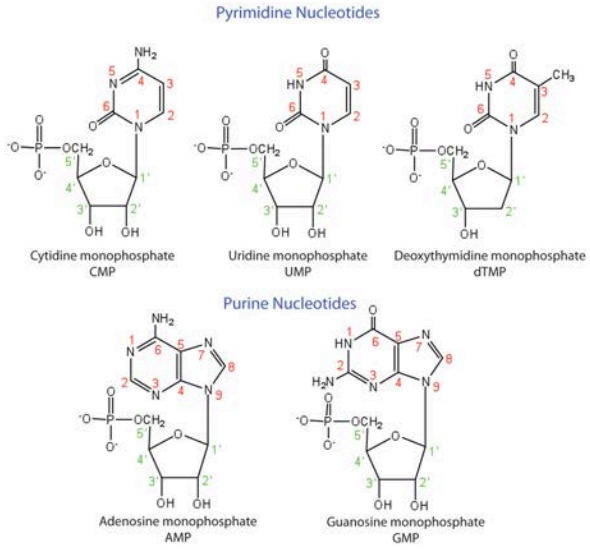


Figure 28.4d. The Pyrimidine and Purine Nucleotides (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Exercise 28.4a

Identify some of the main functional groups found in the structures of Figure 28.4d.

Check Your Answers:[1]

Source: Exercise 28.4a by Samantha Sullivan Sauer, licensed under CC BY-NC 4.0

Apart from being the monomer units of DNA and RNA, the nucleotides and some of their derivatives have other functions as well. Adenosine diphosphate (ADP) and adenosine triphosphate (ATP), shown in Figure 28.4e., have a role in cell metabolism. Moreover, a number of coenzymes, including flavin adenine dinucleotide (FAD), nicotinamide adenine dinucleotide (NAD^+), and coenzyme A, contain adenine nucleotides as structural components.

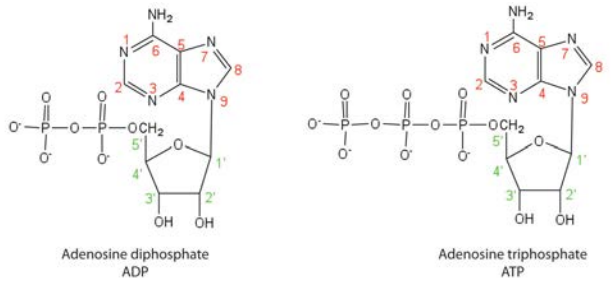


Figure 28.4e. Structures of Two Important Adenine-Containing Nucleotides (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Nucleic Acid Structure

Nucleic acids are large polymers formed by linking nucleotides together and are found in every cell. Deoxyribonucleic acid (DNA) is the nucleic acid that stores genetic information. If all the DNA in a typical mammalian cell were stretched out end to end, it would extend more than 2 m. Ribonucleic acid (RNA) is the nucleic acid responsible for using the genetic information encoded in DNA to produce the thousands of proteins found in living organisms.

Primary Structure of Nucleic Acids

Nucleotides are joined together through the phosphate group of one nucleotide connecting in an ester linkage to the OH group on the third carbon atom of the sugar unit of a second nucleotide. This unit joins to a third nucleotide, and the process is repeated to produce a long nucleic acid chain (Figure 28.4f.). The backbone of the chain consists of alternating phosphate and sugar units (2-deoxyribose in DNA and ribose in RNA). The purine and pyrimidine bases branch off this backbone. Each phosphate group has one acidic hydrogen atom that is ionized at physiological pH. This is why these compounds are known as nucleic acids.

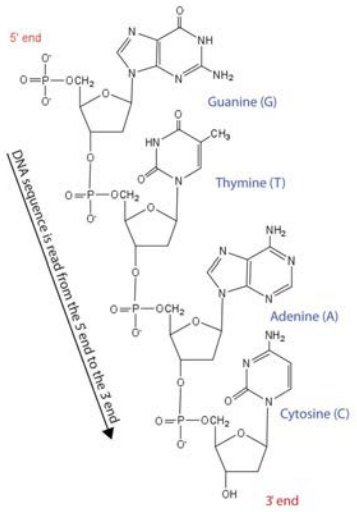


Figure 28.4f. Structure of a Segment of DNA. A similar segment of RNA would have OH groups on each C2', and uracil would replace thymine (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Like proteins, nucleic acids have a primary structure that is defined as the sequence of their nucleotides. Unlike proteins, which have 20 different kinds of amino acids, there are only 4 different kinds of nucleotides in nucleic acids. For amino acid sequences in proteins, the convention is to write the amino acids in order starting with the N-terminal amino acid. In writing nucleotide sequences for nucleic acids, the convention is to write the nucleotides (usually using the one-letter abbreviations for the bases, shown in Figure 28.4f.) starting with the nucleotide having a free phosphate group, which is known as the 5′ end, and indicate the nucleotides in order. For DNA, a lowercase *d* is often written in front of the sequence to indicate that the monomers are deoxyribonucleotides. The final nucleotide has a free OH group on the 3′ carbon atom and is called the *3′ end*. The sequence of nucleotides in the DNA segment shown in Figure 28.4f. would be written 5′-dG-dT-dA-dC-3′, which is often further abbreviated to dGTAC or just GTAC.

Secondary Structure of DNA

The three-dimensional structure of DNA was the subject of an intensive research effort in the late 1940s to early 1950s. Initial work revealed that the polymer had a regular repeating structure. In 1950, Erwin Chargaff of Columbia University showed that the molar amount of adenine (A) in DNA was always equal to that of thymine (T). Similarly, he showed that the molar amount of guanine (G) was the same as that of cytosine (C). Chargaff drew no conclusions from his work, but others soon did.

At Cambridge University in 1953, James D. Watson and Francis Crick announced that they had a model for the secondary structure of DNA. Using the information from Chargaff's experiments (as well as other experiments) and data from the X ray studies of Rosalind Franklin (which involved sophisticated chemistry, physics, and mathematics), Watson and Crick worked with models that were not unlike a child's construction

set and finally concluded that DNA is composed of two nucleic acid chains running antiparallel to one another—that is, side-by-side with the 5′ end of one chain next to the 3′ end of the other. Moreover, as their model showed, the two chains are twisted to form a double helix—a structure that can be compared to a spiral staircase, with the phosphate and sugar groups (the backbone of the nucleic acid polymer) representing the outside edges of the staircase. The purine and pyrimidine bases face the inside of the helix, with guanine always opposite cytosine and adenine always opposite thymine. These specific base pairs, referred to as complementary bases, are the steps, or treads, in our staircase analogy (Figure 28.4g.).

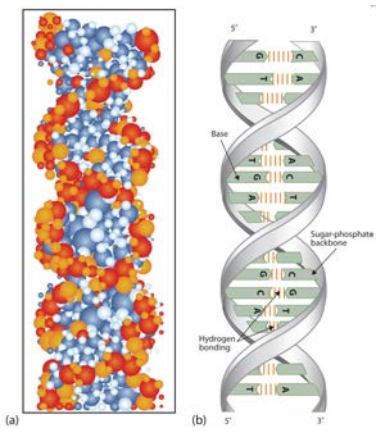


Figure 28.4g. DNA Double Helix. (a) This represents a computer-generated model of the DNA double helix. (b) This represents a schematic representation of the double helix, showing the complementary bases. (Credit: *Introduction to Chemistry: GOB (v. 1.0)*, edited by *(Ball et al.)*, CC BY-NC-SA 4.0)

The structure proposed by Watson and Crick provided clues to the mechanisms by which cells are able to divide into two identical, functioning daughter cells; how genetic data are passed to new generations; and even how proteins are built to required specifications. All these abilities depend on the pairing of complementary bases. Figure 28.4h. shows the two sets of base pairs and illustrates two things. First, a pyrimidine is paired with a purine in each case, so that the long dimensions of both pairs are identical (1.08 nm).

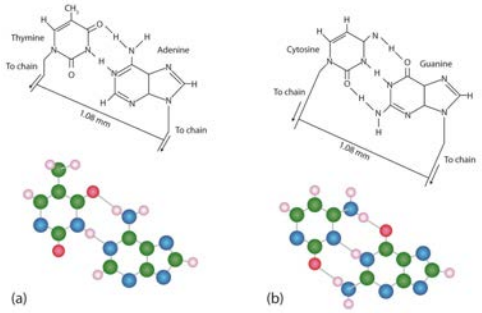


Figure 28.4h. Complementary Base Pairing. Complementary bases engage in hydrogen bonding with one another: (a) thymine and adenine; (b) cytosine and guanine (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

If two pyrimidines were paired or two purines were paired, the two pyrimidines would take up less space than a purine and a pyrimidine, and the two purines would take up more space, as illustrated in Figure 28.4i. If these pairings were ever to occur, the structure of DNA would be like a staircase made with stairs of different widths. For the two strands of the double helix to fit neatly, a pyrimidine must always be paired with a purine. The second thing you should notice in Figure 28.4i. is that the correct pairing enables formation of three instances of hydrogen bonding between guanine and cytosine and two between adenine and thymine. The additive contribution of this hydrogen bonding imparts great stability to the DNA double helix.

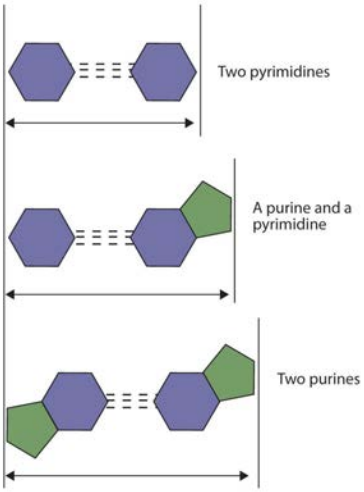
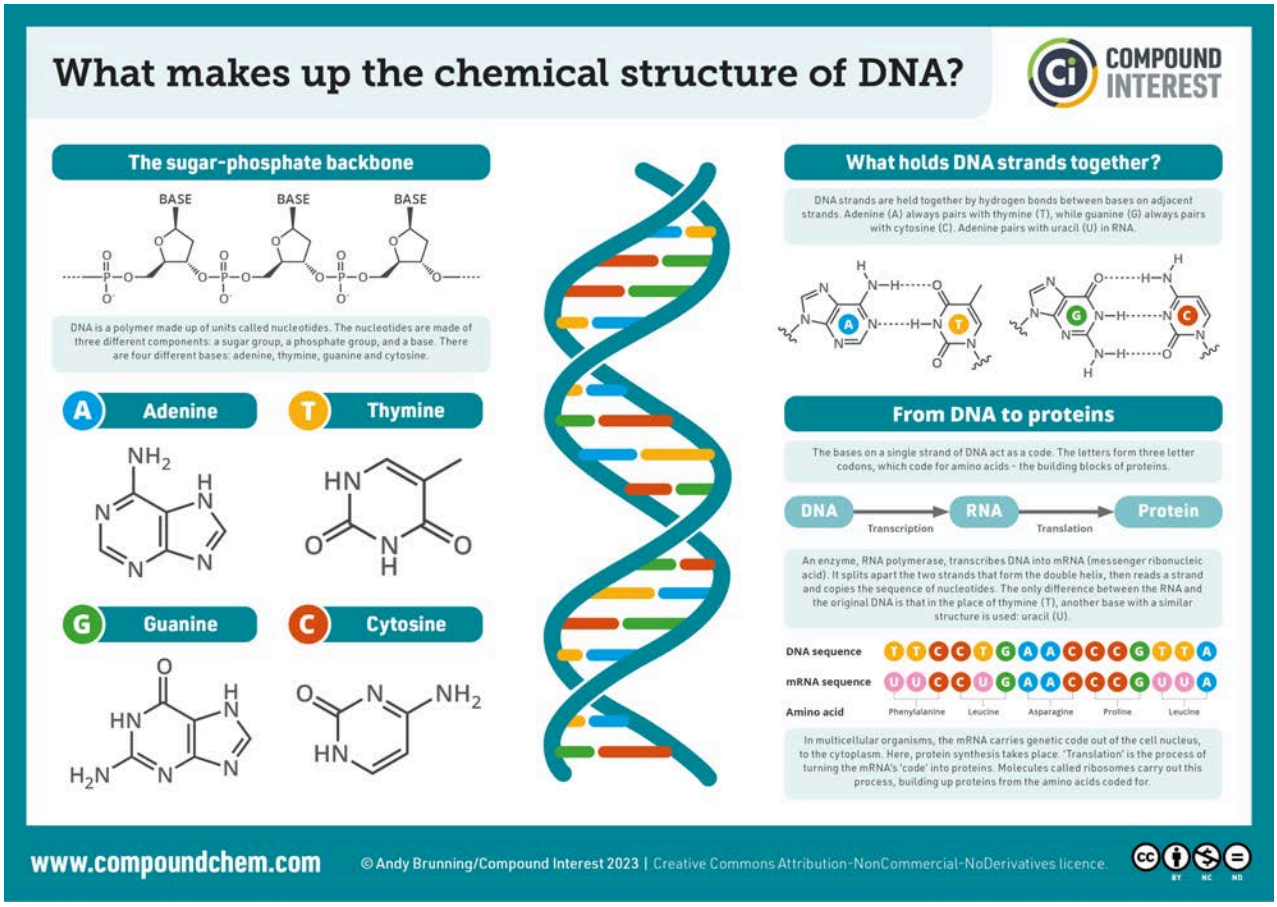


Figure 28.4i. Difference in Widths of Possible Base Pairs (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

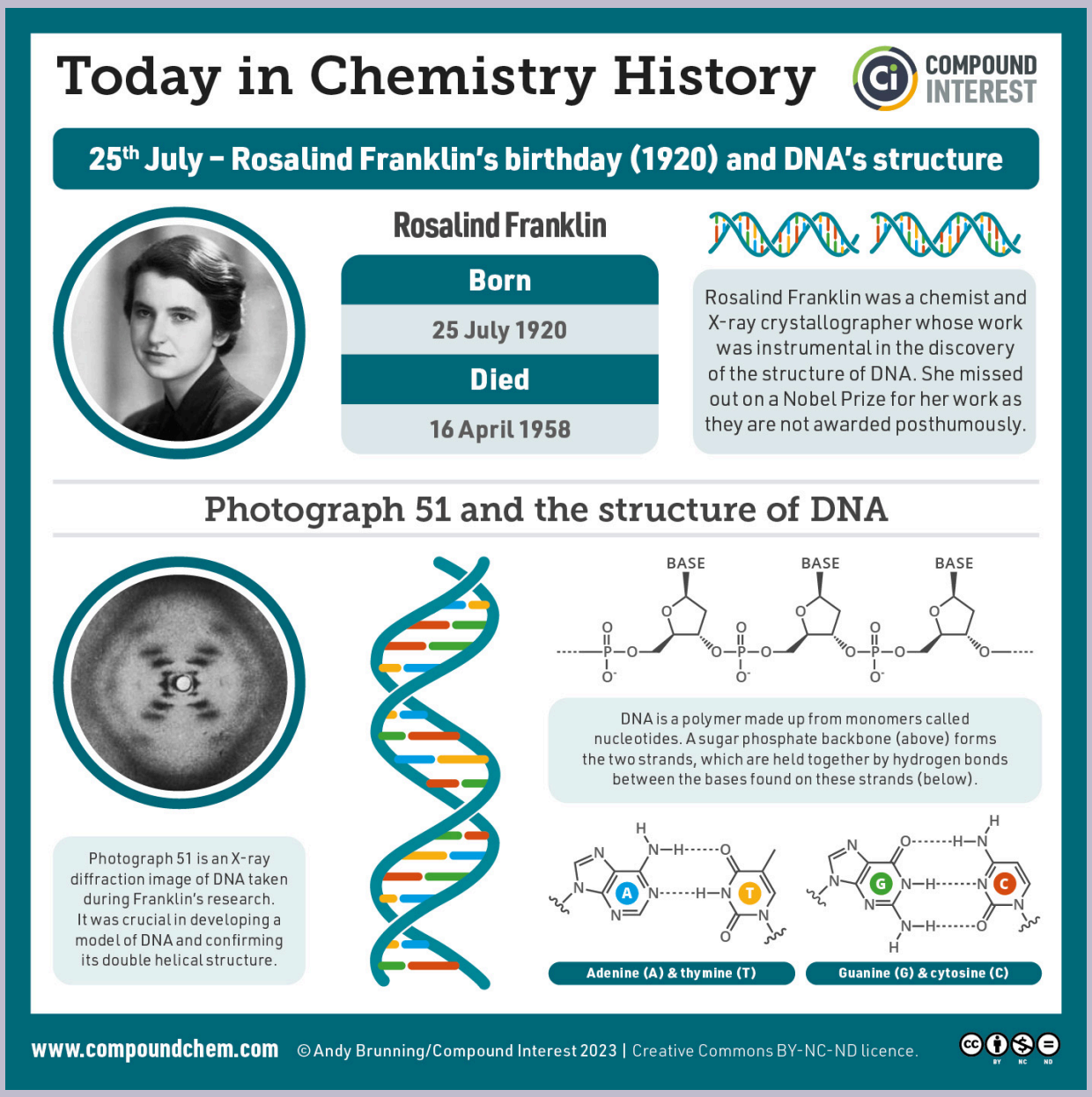
Infographic 28.4a. summarizes the chemical structure of DNA including the backbone, bases, hydrogen bonding, and formation of proteins from DNA and RNA.



Infographic 28.4a. Read more about "What makes up the Chemical Structure of DNA? (https://www.compoundchem.com/2015/03/24/dna/)" by Andy Brunning / Compound Interest, CC BY-NC-ND, or access a text-based summary of infographic 28.4a [New tab].

Spotlight on Everyday Chemistry: Scientist Rosalind Franklin

Rosalind Franklin was instrumental in determining the structure of DNA. Read more about her and this discovery.



Infographic 28.4b. Read more about "Today in Chemistry History – Rosalind Franklin and the structure of DNA (https://www.compoundchem.com/2017/07/25/franklin/)" by Andy Brunning / Compound Interest, CC BY-NC-ND, or access a text-based summary of infographic 28.4b [New tab].

Expressing Genetic Information

We previously stated that deoxyribonucleic acid (DNA) stores genetic information, while ribonucleic acid (RNA) is responsible for transmitting or expressing genetic information by directing the synthesis of

thousands of proteins found in living organisms. But how do the nucleic acids perform these functions? Three processes are required: (1) replication, in which new copies of DNA are made; (2) transcription, in which a segment of DNA is used to produce RNA; and (3) translation, in which the information in RNA is translated into a protein sequence.

Replication

New cells are continuously forming in the body through the process of cell division. For this to happen, the DNA in a dividing cell must be copied in a process known as replication. The complementary base pairing of the double helix provides a ready model for how genetic replication occurs. If the two chains of the double helix are pulled apart, disrupting the hydrogen bonding between base pairs, each chain can act as a template, or pattern, for the synthesis of a new complementary DNA chain.

The nucleus contains all the necessary enzymes, proteins, and nucleotides required for this synthesis. A short segment of DNA is "unzipped," so that the two strands in the segment are separated to serve as templates for new DNA. DNA polymerase, an enzyme, recognizes each base in a template strand and matches it to the complementary base in a free nucleotide. The enzyme then catalyzes the formation of an ester bond between the 5′ phosphate group of the nucleotide and the 3′ OH end of the new, growing DNA chain. In this way, each strand of the original DNA molecule is used to produce a duplicate of its former partner (Figure 28.4j.). Whatever information was encoded in the original DNA double helix is now contained in each replicate helix. When the cell divides, each daughter cell gets one of these replicates and thus all of the information that was originally possessed by the parent cell.

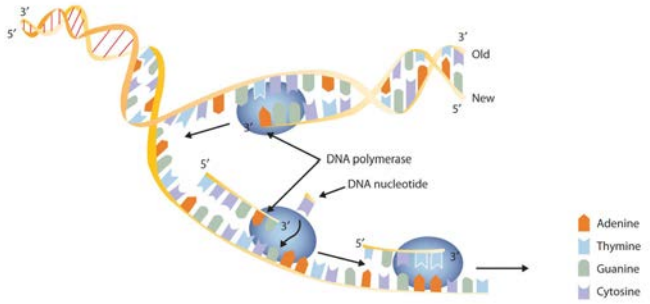


Figure 28.4j. A Schematic Diagram of DNA Replication. DNA replication occurs by the sequential unzipping of segments of the double helix. Each new nucleotide is brought into position by DNA polymerase and is added to the growing strand by the formation of a phosphate ester bond. Thus, two double helixes form from one, and each consists of one old strand and one new strand, an outcome called semiconservative replications. (This representation is simplified; many more proteins are involved in replication (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Example 28.4a

A segment of one strand from a DNA molecule has the sequence 5'-TCCATGAGTTGA-3'. What is the sequence of nucleotides in the opposite, or complementary, DNA chain?

Solution

Knowing that the two strands are antiparallel and that T base pairs with A, while C base pairs with G, the sequence of the complementary strand will be 3'-AGGTACTCAACT-5' (can also be written as TCAACTCATGGA).

What do we mean when we say information is encoded in the DNA molecule? An organism's DNA can be compared to a book containing directions for assembling a model airplane or for knitting a sweater. Letters of the alphabet are arranged into words, and these words direct the individual to perform certain operations with specific materials. If all the directions are followed correctly, a model airplane or sweater is produced.

In DNA, the particular sequences of nucleotides along the chains encode the directions for building an organism. Just as *saw* means one thing in English and *was* means another, the sequence of bases CGT means one thing, and TGC means something different. Although there are only four letters—the four nucleotides—in the genetic code of DNA, their sequencing along the DNA strands can vary so widely that information storage is essentially unlimited.

Transcription

For the hereditary information in DNA to be useful, it must be "expressed," that is, used to direct the growth and functioning of an organism. The first step in the processes that constitute DNA expression is the synthesis of RNA, by a template mechanism that is in many ways analogous to DNA replication. Because the RNA that is synthesized is a complimentary copy of information contained in DNA, RNA synthesis is referred to as transcription. There are three key differences between replication and transcription:

1. RNA molecules are much shorter than DNA molecules; only a portion of one DNA strand is copied or transcribed to make an RNA molecule.
2. RNA is built from ribonucleotides rather than deoxyribonucleotides.
3. The newly synthesized RNA strand does not remain associated with the DNA sequence it was transcribed from.

The DNA sequence that is transcribed to make RNA is called the *template strand*, while the complementary sequence on the other DNA strand is called the *coding* or *informational strand*. To initiate RNA synthesis, the two DNA strands unwind at specific sites along the DNA molecule. Ribonucleotides are attracted to the uncoiling region of the DNA molecule, beginning at the 3′ end of the template strand, according to the rules of base pairing. Thymine in DNA calls for adenine in RNA, cytosine specifies guanine, guanine calls for cytosine, and adenine requires uracil. RNA polymerase—an enzyme—binds the complementary ribonucleotide and catalyzes the formation of the ester linkage between ribonucleotides, a reaction very similar to that catalyzed by DNA polymerase (figure 28.4k). Synthesis of the RNA strand takes place in the 5′ to 3′ direction, antiparallel to the template strand. Only a short segment of the RNA molecule is hydrogen-bonded to the template strand at any time during transcription. When transcription is completed, the RNA is released, and the DNA helix reforms. The nucleotide sequence of the RNA strand formed during transcription is identical to that of the corresponding coding strand of the DNA, except that U replaces T.

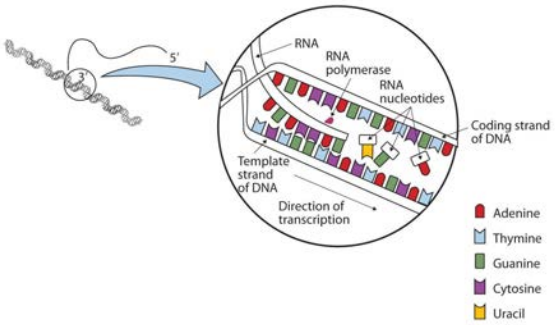


Figure 28.4k. A Schematic Diagram of RNA Transcription from a DNA Template. The representation of RNA polymerase is proportionately much smaller than the actual molecule, which encompasses about 50 nucleotides at a time (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Example 28.4b

A portion of the template strand of a gene has the sequence 5'-TCCATGAGTTGA-3'. What is the sequence of nucleotides in the RNA that is formed from this template?

Solution

Four things must be remembered in answering this question: (1) the DNA strand and the RNA strand being synthesized are antiparallel; (2) RNA is synthesized in a 5' to 3' direction, so transcription begins at the 3' end of the template strand; (3) ribonucleotides are used in place of deoxyribonucleotides; and (4) thymine (T) base pairs with adenine (A), A base pairs with uracil (U; in RNA), and cytosine (C) base

pairs with guanine (G). The sequence is determined to be 3'-AGGUACUCAACU-5' (can also be written as 5'-UCAACUCAUGGA-3').

Three types of RNA are formed during transcription: *messenger RNA* (mRNA), *ribosomal RNA* (rRNA), and *transfer RNA* (tRNA). These three types of RNA differ in function, size, and percentage of the total cell RNA (Table 28.4b.). mRNA makes up only a small percent of the total amount of RNA within the cell, primarily because each molecule of mRNA exists for a relatively short time; it is continuously being degraded and resynthesized. The molecular dimensions of the mRNA molecule vary according to the amount of genetic information a given molecule contains. After transcription, which takes place in the nucleus, the mRNA passes into the cytoplasm, carrying the genetic message from DNA to the ribosomes, the sites of protein synthesis.

Table 28.4b. Properties of Cellular RNA in *Escherichia coli*

Type	Function	Approximate Number of Nucleotides	Percentage of Total Cell RNA
mRNA	codes for proteins	100–6,000	~3
rRNA	component of ribosomes	120–2900	83
tRNA	adapter molecule that brings the amino acid to the ribosome	75–90	14

Source: "19.3: Replication and Expression of Genetic Information" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

Ribosomes are cellular substructures where proteins are synthesized. They contain about 65% rRNA and 35% protein, held together by numerous noncovalent interactions, such as hydrogen bonding, in an overall structure consisting of two globular particles of unequal size.

Molecules of tRNA, which bring amino acids (one at a time) to the ribosomes for the construction of proteins, differ from one another in the kinds of amino acid each is specifically designed to carry (Figure 28.4l.). A set of three nucleotides, known as a codon, on the mRNA determines which kind of tRNA will add its amino acid to the growing chain. Each of the 20 amino acids found in proteins has at least one corresponding kind of tRNA, and most amino acids have more than one.

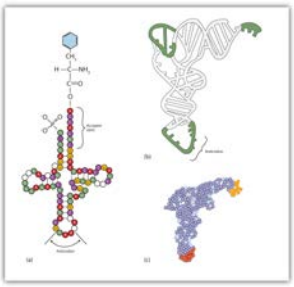
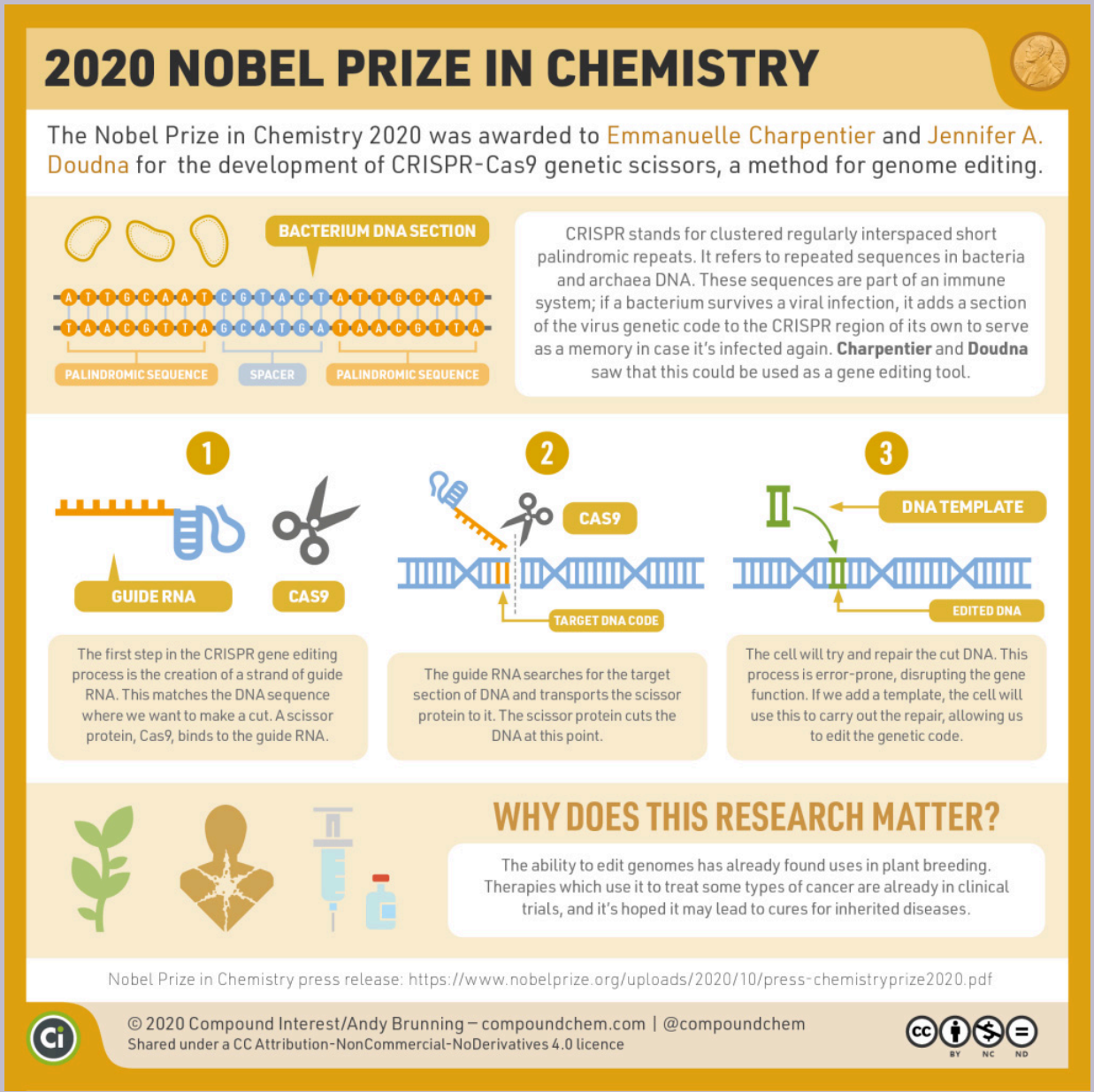


Figure 28.4l. Transfer RNA. (a) In the two-dimensional structure of a yeast tRNA molecule for phenylalanine, the amino acid binds to the acceptor stem located at the 3' end of the tRNA primary sequence. (The nucleotides that are not specifically identified here are slightly altered analogs of the four common ribonucleotides A, U, C, and G.) (b) In the three-dimensional structure of yeast phenylalanine tRNA, note that the anticodon loop is at the bottom and the acceptor stem is at the top right. (c) This shows a space-filling model of the tRNA (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The two-dimensional structure of a tRNA molecule has three distinctive loops, reminiscent of a cloverleaf (Figure 28.4l.). On one loop is a sequence of three nucleotides that varies for each kind of tRNA. This triplet, called the anticodon, is complementary to and pairs with the codon on the mRNA. At the opposite end of the molecule is the acceptor stem, where the amino acid is attached.

Spotlight on Everyday Chemistry: Genome Editing

The 2020 Nobel Prize in Chemistry was awarded to scientists who developed a method of genome editing.



Infographic 28.4c. Read more about "The 2020 Nobel Prize in Chemistry: Using genetic scissors to edit the genome (https://www.compoundchem.com/2020/10/07/2020nobelchemistry/)" by Andy Brunning / Compound Interest, CC BY-NC-ND, or access a text-based summary of infographic 28.4c [New tab].

Protein Synthesis

One of the definitions of a gene is as follows: a segment of deoxyribonucleic acid (DNA) carrying the code for a specific polypeptide. Each molecule of messenger RNA (mRNA) is a transcribed copy of a gene that is used

by a cell for synthesizing a polypeptide chain. If a protein contains two or more different polypeptide chains, each chain is coded by a different gene. We turn now to the question of how the sequence of nucleotides in a molecule of ribonucleic acid (RNA) is translated into an amino acid sequence.

How can a molecule containing just 4 different nucleotides specify the sequence of the 20 amino acids that occur in proteins? If each nucleotide coded for 1 amino acid, then obviously the nucleic acids could code for only 4 amino acids. What if amino acids were coded for by groups of 2 nucleotides? There are 4^2, or 16, different combinations of 2 nucleotides (AA, AU, AC, AG, UU, and so on). Such a code is more extensive but still not adequate to code for 20 amino acids. However, if the nucleotides are arranged in groups of 3, the number of different possible combinations is 4^3, or 64. Here we have a code that is extensive enough to direct the synthesis of the primary structure of a protein molecule.

Watch Translation (mRNA to protein) | Biomolecules | MCAT | Khan Academy on YouTube (14 mins) (https://youtu.be/ocAAkB32Hqs)

Before an amino acid can be incorporated into a polypeptide chain, it must be attached to its unique tRNA. This crucial process requires an enzyme known as aminoacyl-tRNA synthetase (Figure 28.4m.). There is a specific aminoacyl-tRNA synthetase for each amino acid. This high degree of specificity is vital to the incorporation of the correct amino acid into a protein.

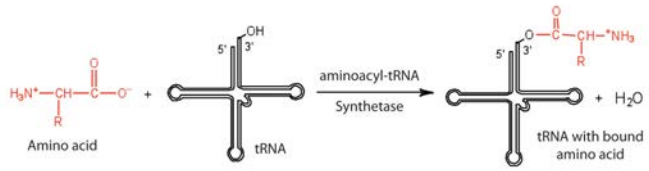


Figure 28.4m. Binding of an Amino Acid to Its tRNA (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

After the amino acid molecule has been bound to its tRNA carrier, protein synthesis can take place. Figures 28.4n-q depicts a schematic stepwise representation of this all-important process.

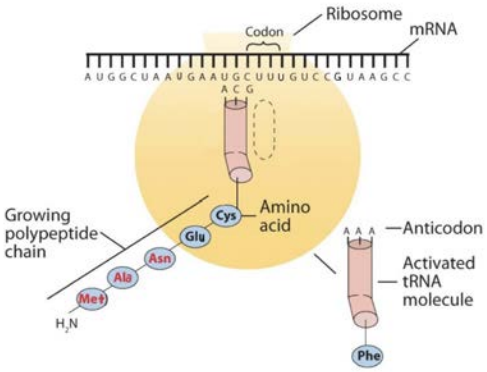


Figure 28.4n. The Elongation Steps in Protein Synthesis – Protein synthesis is already in progress at the ribosome. The growing polypeptide chain is attached to the tRNA that brought in the previous amino acid (in this illustration, Cys) (Credit: *Introduction to Chemistry: General, Organic, and Biological (v. 1.0)*, edited by *(Ball et al.)* CC BY-NC-SA 4.0).

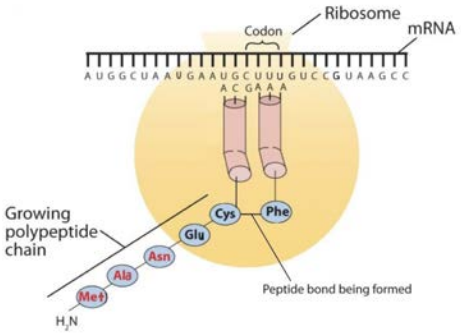


Figure 28.4o. The Elongation Steps in Protein Synthesis – An activated tRNA, which has the anticodon AAA, binds to the ribosome next to the previous bound tRNA and interacts with the mRNA molecule though base pairing of the codon and anticodon. The amino acid Phe is being incorporated into the polypeptide chain by the formation of a peptide linkage between the carboxyl group of Cys and the amino acid group of the Phe. This reaction is catalyzed by the enzyme peptidyl transferase, a component of the ribosome. (Credit: *Introduction to Chemistry: General, Organic, and Biological (v. 1.0)*, edited by *(Ball et al.)* CC BY-NC-SA 4.0)

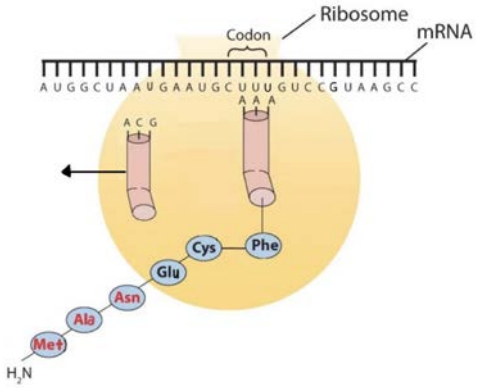


Figure 28.4p. The Elongation Steps in Protein Synthesis – The Cys-Phe linkage is now complete, and the growing polypeptide chain remains attached to the tRNA for Phe. (Credit: *Introduction to Chemistry: General, Organic, and Biological (v. 1.0)*, edited by *(Ball et al.)* CC BY-NC-SA 4.0)

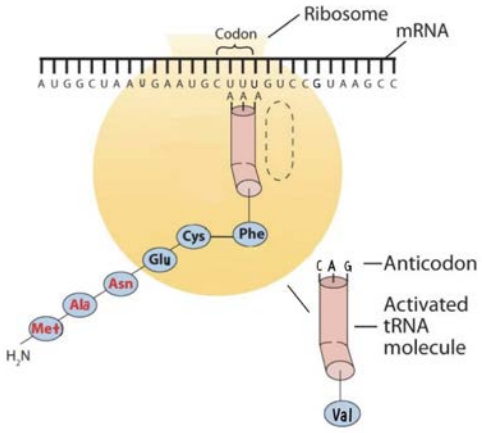


Figure 28.4q. The Elongation Steps in Protein Synthesis – The ribosome moves to the right along the mRNA strand. This shift brings the next codon, GUC, into its correct position on the surface of the ribosome. Note that an activated tRNA molecule, containing the next amino acid to be attached to the chain is moving to the ribosome. Steps (b)-(d) will be repeated until the ribosome reaches a stop codon. (Credit: *Introduction to Chemistry: General, Organic, and Biological (v. 1.0)*, edited by *(Ball et al.)* CC BY-NC-SA 4.0)

Early experimenters were faced with the task of determining which of the 64 possible codons stood for each of the 20 amino acids. The cracking of the genetic code was the joint accomplishment of several well-known geneticists—notably Har Khorana, Marshall Nirenberg, Philip Leder, and Severo Ochoa—from 1961 to 1964. The genetic dictionary they compiled, summarized in figure 28.4r, shows that 61 codons code for amino acids, and 3 codons serve as signals for the termination of polypeptide synthesis (much like the period at the end of a sentence). Notice that only methionine (AUG) and tryptophan (UGG) have single codons. All other amino acids have two or more codons.

First base	Second base: U	Second base: C	Second base: A	Second base: G	Third base
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Figure 28.4r. The Genetic Code (Credit: *Introduction to Chemistry: General, Organic, and Biological (v. 1.0)*, edited by *(Ball et al.)* CC BY-NC-SA 4.0)

Example 28.4c

A portion of an mRNA molecule has the sequence 5'-AUGCCACGAGUUGAC-3'. What amino acid sequence does this code for?

Solution

Use Figure 28.4r to determine what amino acid each set of three nucleotides (codon) codes for. Remember that the sequence is read starting from the 5' end and that a protein is synthesized starting with the N-terminal amino acid. The sequence 5'-AUGCCACGAGUUGAC-3' codes for met-pro-arg-val-asp.

1. The code is virtually universal; animal, plant, and bacterial cells use the same codons to specify each amino acid (with a few exceptions).
2. The code is "degenerate"; in all but two cases (methionine and tryptophan), more than one triplet codes for a given amino acid.
3. The first two bases of each codon are most significant; the third base often varies. This suggests that a change in the third base by a mutation may still permit the correct incorporation of a given amino acid into a protein. The third base is sometimes called the "wobble" base.
4. The code is continuous and nonoverlapping; there are no nucleotides between codons, and adjacent codons do not overlap.

5. The three termination codons are read by special proteins called release factors, which signal the end of the translation process.
6. The codon AUG codes for methionine and is also the initiation codon. Thus methionine is the first amino acid in each newly synthesized polypeptide. This first amino acid is usually removed enzymatically before the polypeptide chain is completed; the vast majority of polypeptides do not begin with methionine.

Mutations and Genetic Diseases

We have seen that the sequence of nucleotides in a cell's deoxyribonucleic acid (DNA) is what ultimately determines the sequence of amino acids in proteins made by the cell and thus is critical for the proper functioning of the cell. On rare occasions, however, the nucleotide sequence in DNA may be modified either spontaneously (by errors during replication, occurring approximately once for every 10 billion nucleotides) or from exposure to heat, radiation, or certain chemicals. Any chemical or physical change that alters the nucleotide sequence in DNA is called a mutation. When a mutation occurs in an egg or sperm cell that then produces a living organism, it will be inherited by all the offspring of that organism.

Common types of mutations include substitution (a different nucleotide is substituted), insertion (the addition of a new nucleotide), and deletion (the loss of a nucleotide). These changes within DNA are called point mutations because only one nucleotide is substituted, added, or deleted (Figure 28.4s.). Because an insertion or deletion results in a frame-shift that changes the reading of subsequent codons and, therefore, alters the entire amino acid sequence that follows the mutation, insertions and deletions are usually more harmful than a substitution in which only a single amino acid is altered.

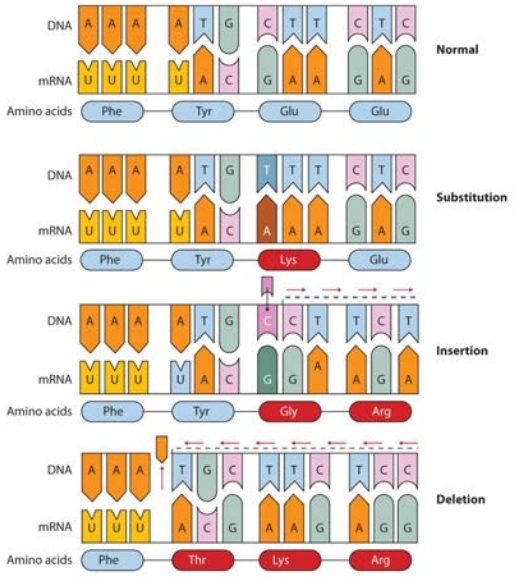


Figure 28.4s. Three Types of Point Mutations (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

The chemical or physical agents that cause mutations are called mutagens. Examples of physical mutagens are ultraviolet (UV) and gamma radiation. Radiation exerts its mutagenic effect either directly or by creating free radicals that in turn have mutagenic effects. Radiation and free radicals can lead to the formation of bonds between nitrogenous bases in DNA. For example, exposure to UV light can result in the formation of a covalent bond between two adjacent thymines on a DNA strand, producing a thymine dimer (Figure 28.4t.). If not repaired, the dimer prevents the formation of the double helix at the point where it occurs. The genetic disease *xeroderma pigmentosum* is caused by a lack of the enzyme that cuts out the thymine dimers in damaged DNA. Individuals affected by this condition are abnormally sensitive to light and are more prone to skin cancer than normal individuals.

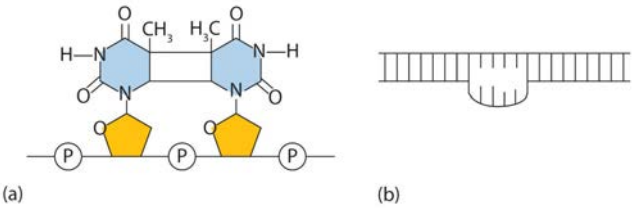


Figure 28.4t. An Example of Radiation Damage to DNA. (a) The thymine dimer is formed by the action of UV light. (b) When a defect in the double strand is produced by the thymine dimer, this defect temporarily stops DNA replication, but the dimer can be removed, and the region can be repaired by an enzyme repair system (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Sometimes gene mutations are beneficial, but most of them are detrimental. For example, if a point mutation occurs at a crucial position in a DNA sequence, the affected protein will lack biological activity, perhaps

resulting in the death of a cell. In such cases the altered DNA sequence is lost and will not be copied into daughter cells. Nonlethal mutations in an egg or sperm cell may lead to metabolic abnormalities or hereditary diseases. Such diseases are called *inborn errors of metabolism* or genetic diseases. A partial listing of genetic diseases is presented in Table 28.4c., and two specific diseases are discussed in the following sections. In most cases, the defective gene results in a failure to synthesize a particular enzyme.

Table 28.4c. Some Representative Genetic Diseases in Humans and the Protein or Enzyme Responsible

Disease	Responsible Protein or Enzyme
alkaptonuria	homogentisic acid oxidase
galactosemia	galactose 1-phosphate uridyl transferase, galactokinase, or UDP galactose epimerase
Gaucher disease	glucocerebrosidase
gout and Lesch-Nyhan syndrome	hypoxanthine-guanine phosphoribosyl transferase
hemophilia	antihemophilic factor (factor VIII) or Christmas factor (factor IX)
homocystinuria	cystathionine synthetase
maple syrup urine disease	branched chain α-keto acid dehydrogenase complex
McArdle syndrome	muscle phosphorylase
Niemann-Pick disease	sphingomyelinase
phenylketonuria (PKU)	phenylalanine hydroxylase
sickle cell anemia	hemoglobin
Tay-Sachs disease	hexosaminidase A
tyrosinemia	fumarylacetoacetate hydrolase or tyrosine aminotransferase
von Gierke disease	glucose 6-phosphatase
Wilson disease	Wilson disease protein

Source: "19.5: Mutations and Genetic Diseases" In *Basics of GOB Chemistry (Ball et al.)*, CC BY-NC-SA 4.0.

Phenylketonuria (PKU), as seen in the table above, results from the absence of the enzyme phenylalanine hydroxylase. Without this enzyme, a person cannot convert phenylalanine to tyrosine, which is the precursor of the neurotransmitters dopamine and norepinephrine as well as the skin pigment melanin (Figure 28.4u.).

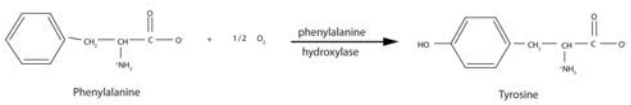


Figure 28.4u. Normally, phenylalanine is converted to tyrosine by the enzyme phenylalanine hydroxylase. People with PKU lack this enzyme, and thus cannot process phenylalanine, which is necessary for the production of neurotransmitters such as dopamine and norepenephrine and the pigment melanin (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

When this reaction cannot occur, phenylalanine accumulates and is then converted to higher than normal quantities of phenylpyruvate. The disease acquired its name from the high levels of phenylpyruvate (a phenyl ketone) in urine. Excessive amounts of phenylpyruvate impair normal brain development, which causes severe mental retardation (Figure 28.4v.).

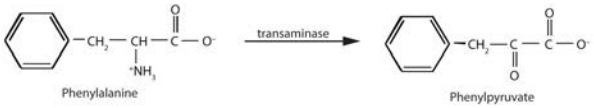


Figure 28.4v. A buildup of phenylalanine in those with PKU results in the accumulation of phenylpyruvate, which inhibits normal brain development (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

PKU may be diagnosed by assaying a sample of blood or urine for phenylalanine or one of its metabolites. Medical authorities recommend testing every newborn's blood for phenylalanine within 24 h to 3 weeks after birth. If the condition is detected, mental retardation can be prevented by immediately placing the infant on a diet containing little or no phenylalanine. Because phenylalanine is plentiful in naturally produced proteins, the low-phenylalanine diet depends on a synthetic protein substitute plus very small measured amounts of naturally produced foods. Before dietary treatment was introduced in the early 1960s, severe mental retardation was a common outcome for children with PKU. Prior to the 1960s, 85% of patients with PKU had an intelligence quotient (IQ) less than 40, and 37% had IQ scores below 10. Since the introduction of dietary treatments, however, over 95% of children with PKU have developed normal or near-normal intelligence. The incidence of PKU in newborns is about 1 in 12,000 in North America. Every state in the United States has mandated that screening for PKU be provided to all newborns.

Several genetic diseases are collectively categorized as lipid-storage diseases. Lipids are constantly being synthesized and broken down in the body, so if the enzymes that catalyze lipid degradation are missing, the lipids tend to accumulate and cause a variety of medical problems. When a genetic mutation occurs in the gene for the enzyme hexosaminidase A, for example, gangliosides cannot be degraded but accumulate in brain tissue, causing the ganglion cells of the brain to become greatly enlarged and nonfunctional. This genetic disease, known as Tay-Sachs disease, leads to a regression in development, dementia, paralysis, and blindness, with death usually occurring before the age of three. There is currently no treatment, but Tay-Sachs disease can be diagnosed in a fetus by assaying the amniotic fluid (amniocentesis) for hexosaminidase A. A blood test can identify Tay-Sachs carriers—people who inherit a defective gene from only one rather than both parents—because they produce only half the normal amount of hexosaminidase A, although they do not exhibit symptoms of the disease.

Recombinant DNA Technology

More than 3,000 human diseases have been shown to have a genetic component, caused or in some way modulated by the person's genetic composition. Moreover, in the last decade or so, researchers have succeeded

in identifying many of the genes and even mutations that are responsible for specific genetic diseases. Now scientists have found ways of identifying and isolating genes that have specific biological functions and placing those genes in another organism, such as a bacterium, which can be easily grown in culture. With these techniques, known as recombinant DNA technology, the ability to cure many serious genetic diseases appears to be within our grasp.

Isolating the specific gene or genes that cause a particular genetic disease is a monumental task. One reason for the difficulty is the enormous amount of a cell's DNA, only a minute portion of which contains the gene sequence. Thus, the first task is to obtain smaller pieces of DNA that can be more easily handled. Fortunately, researchers are able to use restriction enzymes (also known as restriction endonucleases), discovered in 1970, which are enzymes that cut DNA at specific, known nucleotide sequences, yielding DNA fragments of shorter length. For example, the restriction enzyme *EcoRI* recognizes the nucleotide sequence shown here and cuts both DNA strands as indicated in Figure 28.4w.

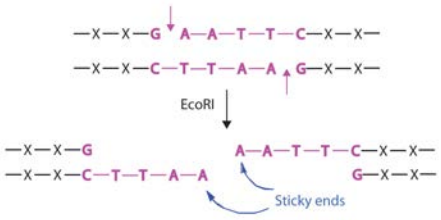


Figure 28.4w. EcoRI is a restriction endonuclease that always cuts DNA at a specific genetic sequence, as shown here (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Once a DNA strand has been fragmented, it must be cloned; that is, multiple identical copies of each DNA fragment are produced to make sure there are sufficient amounts of each to detect and manipulate in the laboratory. Cloning is accomplished by inserting the individual DNA fragments into phages (bacterial viruses) that can enter bacterial cells and be replicated. When a bacterial cell infected by the modified phage is placed in an appropriate culture medium, it forms a colony of cells, all containing copies of the original DNA fragment. This technique is used to produce many bacterial colonies, each containing a different DNA fragment. The result is a DNA library, a collection of bacterial colonies that together contain the entire genome of a particular organism.

The next task is to screen the DNA library to determine which bacterial colony (or colonies) has incorporated the DNA fragment containing the desired gene. A short piece of DNA, known as a hybridization probe, which has a nucleotide sequence complementary to a known sequence in the gene, is synthesized, and a radioactive phosphate group is added to it as a "tag." You might be wondering how researchers are able to prepare such a probe if the gene has not yet been isolated. One way is to use a segment of the desired gene isolated from another organism. An alternative method depends on knowing all or part of the amino acid sequence of the protein produced by the gene of interest: the amino acid sequence is used to produce an approximate genetic code for the gene, and this nucleotide sequence is then produced

synthetically. (The amino acid sequence used is carefully chosen to include, if possible, many amino acids such as methionine and tryptophan, which have only a single codon each.)

After a probe identifies a colony containing the desired gene, the DNA fragment is clipped out, again using restriction enzymes, and spliced into another replicating entity, usually a plasmid. Plasmids are tiny mini-chromosomes found in many bacteria, such as *Escherichia coli* (*E. coli*). A recombined plasmid would then be inserted into the host organism (usually the bacterium *E. coli*), where it would go to work to produce the desired protein (Figure 28.4x.).

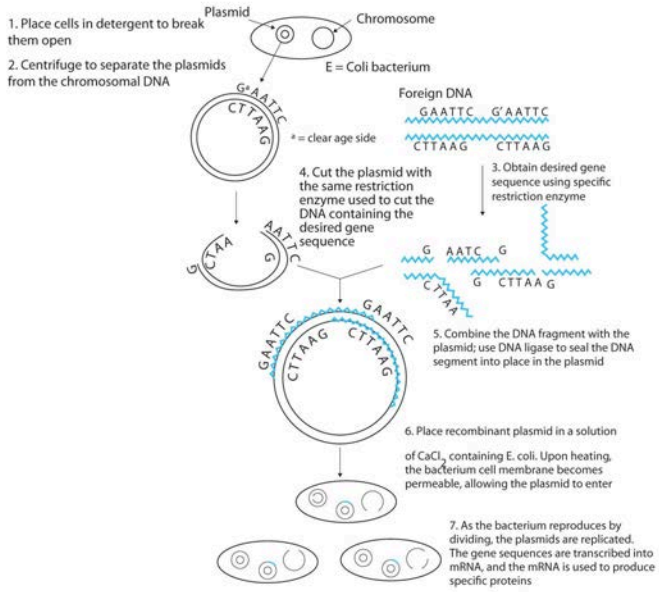


Figure 28.4x. Flow chart diagram outlining the process of cloning (credit: *Intro Chem: GOB (v. 1.0)*, CC BY-NC-SA 3.0).

Proponents of recombinant DNA research are excited about its great potential benefits. An example is the production of human growth hormone, which is used to treat children who fail to grow properly. Formerly, human growth hormone was available only in tiny amounts obtained from cadavers. Now it is readily available through recombinant DNA technology. Another gene that has been cloned is the gene for epidermal growth factor, which stimulates the growth of skin cells and can be used to speed the healing of burns and other skin wounds. Recombinant techniques are also a powerful research tool, providing enormous aid to scientists as they map and sequence genes and determine the functions of different segments of an organism's DNA.

In addition to advancements in the ongoing treatment of genetic diseases, recombinant DNA technology may actually lead to cures. When appropriate genes are successfully inserted into *E. coli*, the bacteria can become miniature pharmaceutical factories, producing great quantities of insulin for people with diabetes, clotting factor for people with hemophilia, missing enzymes, hormones, vitamins, antibodies, vaccines, and so on. Recent accomplishments include the production in *E. coli* of recombinant DNA molecules containing

synthetic genes for tissue plasminogen activator, a clot-dissolving enzyme that can rescue heart attack victims, as well as the production of vaccines against hepatitis B (humans) and hoof-and-mouth disease (cattle).

Scientists have used other bacteria besides *E. coli* in gene-splicing experiments and also yeast and fungi. Plant molecular biologists use a bacterial plasmid to introduce genes for several foreign proteins (including animal proteins) into plants. The bacterium is *Agrobacterium tumefaciens*, which can cause tumors in many plants, but which can be treated so that its tumor-causing ability is eliminated. One practical application of its plasmids would be to enhance a plant's nutritional value by transferring into it the gene necessary for the synthesis of an amino acid in which the plant is normally deficient (for example, transferring the gene for methionine synthesis into pinto beans, which normally do not synthesize high levels of methionine).

Restriction enzymes have been isolated from a number of bacteria and are named after the bacterium of origin. *EcoRI* is a restriction enzyme obtained from the R strain of *E. coli*. The roman numeral I indicates that it was the first restriction enzyme obtained from this strain of bacteria.

Attribution & References

Except where otherwise noted, portions of this page were written by Gregory A. Anderson, while others were adapted by Gregory A. Anderson and Samantha Sullivan Sauer from "19: Nucleic Acids", "19.1: Nucleotides", "19.2: Nucleic Acid Structure","19.3: Replication and Expression of Genetic Information", "19.4: Protein Synthesis and the Genetic Code", and "19.5: Mutations and Genetic Diseases" In *Basics of General, Organic, and Biological Chemistry (Ball et al.)* by David W. Ball, John W. Hill, and Rhonda J. Scott via LibreTexts, CC BY-NC-SA 4.0./ A LibreTexts version of *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0. / Pages were combined and content edited to improve flow and student understanding.

Notes

1. In addition to the alkane and phosphate components, CMP has alcohol, ether, amide and amine groups. UMP has alcohol, ether and amides groups. dTMP has alcohol, ether and amides groups. AMP has alcohol, ether, alkene, and amine groups. GMP has alcohol, ether, alkene, amine and amide groups.

28.5 VITAMINS

Learning Objectives

By the end of this section, you will be able to:

- Explain why vitamins are necessary in the diet
- Explain why some vitamins are water soluble and some are lipid soluble.
- Describe the functional role, intake recommendations and sources of vitamins.
- Explain the role of vitamins as antioxidants and as coenzymes.

Vitamins are essential to human health and can be obtained in our diet from different types of food.

The Vitamins: Vital, but Not All are Amines

In 1747, the Scottish surgeon James Lind discovered that citrus foods helped prevent scurvy, a particularly deadly disease in which collagen is not properly formed, causing poor wound healing, bleeding of the gums, severe pain, and death. In 1753, Lind published his *Treatise on the Scurvy*, which recommended using lemons and limes to avoid scurvy, which was adopted by the British Royal Navy. This led to the nickname *limey* for British sailors.

In East Asia, where polished white rice was the common staple food of the middle class, beriberi resulting from lack of vitamin B_1 was endemic. In 1884, Takaki Kanehiro, a British-trained medical doctor of the Imperial Japanese Navy, observed that beriberi was endemic among low-ranking crew who often ate nothing but rice, but not among officers who consumed a Western-style diet. This convinced Takaki and the Japanese Navy that diet was the cause of beriberi, but they mistakenly believed that sufficient amounts of protein prevented it. That diseases could result from some dietary deficiencies was further investigated by Christiaan Eijkman, who in 1897 discovered that feeding unpolished rice instead of the polished variety to chickens helped to prevent beriberi in the chickens. The following year, Frederick Hopkins postulated that some foods contained "accessory factors" — in addition to proteins, carbohydrates, fats *etc.* — that are necessary for the

functions of the human body. Hopkins and Eijkman were awarded the Nobel Prize for Physiology or Medicine in 1929 for their discoveries.

In 1910, the first vitamin complex was isolated by Japanese scientist Umetaro Suzuki, who succeeded in extracting a water-soluble complex of micronutrients from rice bran and named it aberic acid (later *Orizanin*). He published this discovery in a Japanese scientific journal. When the article was translated into German, the translation failed to state that it was a newly discovered nutrient, a claim made in the original Japanese article, and hence his discovery failed to gain publicity. In 1912 Polish-born biochemist Casimir Funk, working in London, isolated the same complex of micronutrients and proposed the complex be named "vitamine". It was later to be known as vitamin B_3 (niacin), though he described it as "anti-beri-beri-factor" (which would today be called thiamine or vitamin B_1). Funk proposed the hypothesis that other diseases, such as rickets, pellagra, coeliac disease, and scurvy could also be cured by vitamins.

Vitamine to Vitamin

Max Nierenstein a friend and reader of Biochemistry at Bristol University reportedly suggested the "vitamine" name (from "vital amine"). The name soon became synonymous with Hopkins' "accessory factors", and, by the time it was shown that not all vitamins are amines, the word was already ubiquitous. In 1920, Jack Cecil Drummond proposed that the final "e" be dropped to deemphasize the "amine" reference, after researchers began to suspect that not all "vitamines" (in particular, vitamin A) have an amine component (Figure 28.5a.).

LIX. THE NOMENCLATURE OF THE SO-CALLED ACCESSORY FOOD FACTORS (VITAMINS).

By JACK CECIL DRUMMOND.

From the Institute of Physiology, University College, London.

(*Received August 12th, 1920.*)

In 1912 Hopkins published his classical paper in which he described the important influence of certain dietary constituents on the processes of growth and nutrition. These substances he termed the "accessory factors of the diet." At about the same time Funk, who was working on the subject of experimental beriberi, coined the name "Vitamine" for the same class of substances. Since then the literature has been a good deal confused by the great variety of names which have been utilised to denote these or similar dietary constituents (auximones, Bottomley; nutramines, Abderhalden, etc.). The criticism usually raised against Funk's word Vitamine is that the termination "-ine" is one strictly employed in chemical nomenclature to denote substances of a basic character, whereas there is no evidence which supports his original idea that these indispensable dietary constituents are amines. The word has, however, been widely adopted, and therefore until we know more about the actual nature of the substances themselves, it would be difficult and perhaps unwise to eliminate it altogether. The suggestion is now advanced that the final "-e" be dropped, so that the resulting word Vitamin is acceptable under the standard scheme of nomenclature adopted by the Chemical Society, which permits a neutral substance of undefined composition to bear a name ending in "-in." If this suggestion is adopted, it is recommended that the somewhat cumbrous nomenclature introduced by McCollum (Fat-soluble A, Water-soluble B), be dropped, and that the substances be spoken of as Vitamin A, B, C, etc. This simplified scheme should be quite sufficient until such time as the factors are isolated, and their true nature identified.

Figure 28.5a. Jack Drummond's single-paragraph article in 1920 which provided structure and nomenclature used today for vitamins (credit: *Chemistry for Changing Times (Hill & McCreary)*, CC BY-NC-SA 4.0).

Vitamins are organic compounds found in foods and are a necessary part of the biochemical reactions in the

body (Figure 28.5b.). They are involved in a number of processes, including mineral and bone metabolism, and cell and tissue growth, and they act as cofactors for energy metabolism.

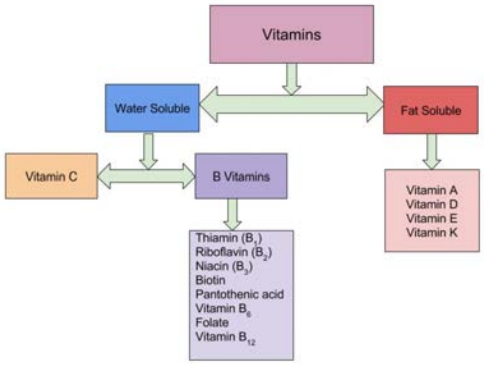


Figure 28.5b. The Vitamins (credit: Image by Allison Calabrese, CC BY 4.0).

You get most of your vitamins through your diet, although some can be formed from the precursors absorbed during digestion. For example, the body synthesizes vitamin A from the β-carotene in orange vegetables like carrots and sweet potatoes. Vitamins are either fat-soluble or water-soluble. Fat-soluble vitamins A, D, E, and K, are absorbed through the intestinal tract with lipids in chylomicrons. Vitamin D is also synthesized in the skin through exposure to sunlight. Because they are carried in lipids, fat-soluble vitamins can accumulate in the lipids stored in the body. If excess vitamins are retained in the lipid stores in the body, hypervitaminosis can result.

Water-soluble vitamins, including the eight B vitamins and vitamin C, are absorbed with water in the gastrointestinal tract. These vitamins move easily through bodily fluids, which are water based, so they are not stored in the body. Excess water-soluble vitamins are excreted in the urine. Therefore, hypervitaminosis of water-soluble vitamins rarely occurs, except with an excess of vitamin supplements. The B vitamins play the largest role of any vitamins in metabolism (Table 28.5a. and Table 28.5b.).

All fat-soluble vitamins contain a high proportion of hydrocarbon structural components. There are one or two oxygen atoms present, but the compounds as a whole are nonpolar. In contrast, water-soluble vitamins contain large numbers of electronegative oxygen and nitrogen atoms, which can engage in hydrogen bonding with water. Most water-soluble vitamins act as coenzymes or are required for the synthesis of coenzymes. The fat-soluble vitamins are important for a variety of physiological functions. A coenzyme is an organic molecule that is necessary for an enzyme's proper functioning. It is one type of cofactor; another type is inorganic ions.

Fat Soluble Vitamins

From the structures in Figure 28.5c., it should be clear that these compounds have more than a solubility connection with lipids. Vitamin A is a terpene, and vitamins E and K have long terpene chains attached to an

aromatic moiety. The structure of vitamin D can be described as a steroid in which ring B is cut open and the remaining three rings remain unchanged. The precursors of vitamins A and D have been identified as the tetraterpene beta-carotene and the steroid ergosterol, respectively. Table 28.5a. lists the different fat-soluble vitamins and its function.

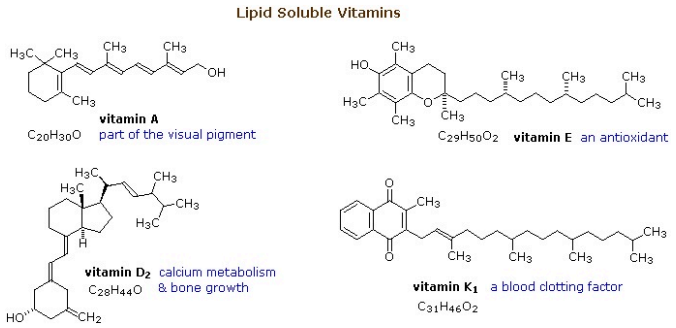


Figure 28.5c. Lipid soluble vitamins of vitamin A, E, D_2 and K_1 (credit: *Chemistry for Changing Times (Hill & McCreary)*, CC BY-NC-SA 4.0).

Table 28.5a. Fat Soluble Vitamins and Their Function (credit: *Chemistry for Changing Times (Hill & McCreary)*, CC BY-NC-SA 4.0).

Vitamin and alternative name	Sources	Recommended daily allowance	Function	Problems associated with deficiency
A retinal or β-carotene	Yellow and orange fruits and vegetables, dark green leafy vegetables, eggs, milk, liver	700–900 μg	Eye and bone development, immune function	Night blindness, epithelial changes, immune system deficiency
D cholecalciferol	Dairy products, egg yolks; also synthesized in the skin from exposure to sunlight	5–15 μg	Aids in calcium absorption, promoting bone growth	Rickets, bone pain, muscle weakness, increased risk of death from cardiovascular disease, cognitive impairment, asthma in children, cancer
E tocopherols	Seeds, nuts, vegetable oils, avocados, wheat germ	15 mg	Antioxidant	Anemia
K phylloquinone	Dark green leafy vegetables, broccoli, Brussels sprouts, cabbage	90–120 μg	Blood clotting, bone health	Hemorrhagic disease of newborn in infants; uncommon in adults

Links to Enhanced Learning

More detailed information on the different fat-soluble vitamins can be found at 9.2: Fat-Soluble Vitamins – Medicine LibreTexts (https://med.libretexts.org/Under_Construction/Purgatory/Book%3A_Human_Nutrition_1e_(University_of_Hawaii)/09%3A_Vitamins/9.02%3A_Fat-Soluble_Vitamins).

Water Soluble Vitamins

All water-soluble vitamins (Table 28.5b.) play a different kind of role in energy metabolism; they are required as functional parts of enzymes involved in energy release and storage. Vitamins and minerals that make up part of enzymes are referred to as coenzymes and cofactors**,** respectively. Coenzymes and cofactors are required by enzymes to catalyze a specific reaction. They assist in converting a substrate to an end-product. Coenzymes and cofactors are essential in catabolic pathways and play a role in many anabolic pathways too. In addition to being essential for metabolism, many vitamins and minerals are required for blood renewal and function. At insufficient levels in the diet these vitamins and minerals impair the health of blood and consequently the delivery of nutrients in and wastes out, amongst its many other functions.

Table 28.5b. Water Soluble Vitamins and Their Function (credit: *Chemistry for Changing Times (Hill & McCreary)*, CC BY-NC-SA 4.0).

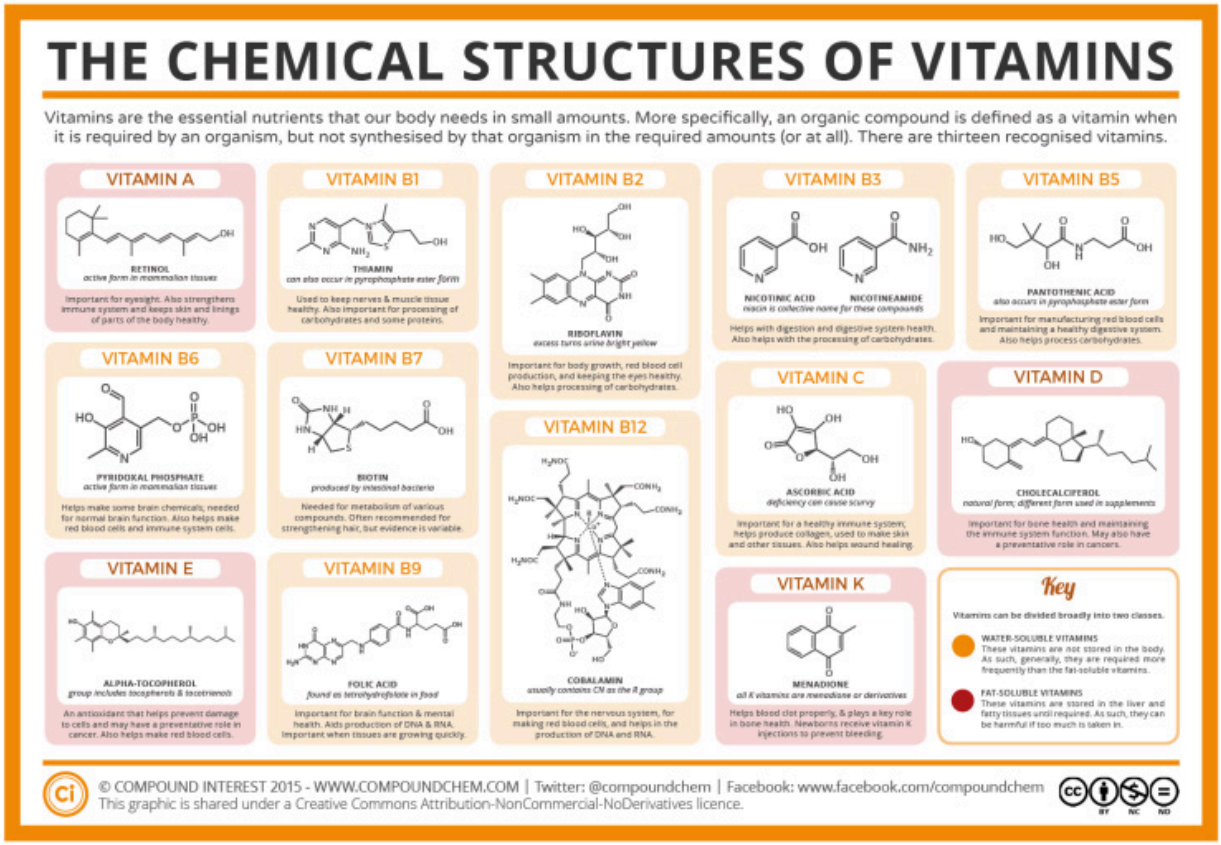
Vitamin and alternative name	Sources	Recommended daily allowance	Function	Problems associated with deficiency
B_1 thiamine	Whole grains, enriched bread and cereals, milk, meat	1.1–1.2 mg	Carbohydrate metabolism	Beriberi, Wernicke-Korsikoff syndrome
B_2 riboflavin	Brewer's yeast, almonds, milk, organ meats, legumes, enriched breads and cereals, broccoli, asparagus	1.1–1.3 mg	Synthesis of FAD for metabolism, production of red blood cells	Fatigue, slowed growth, digestive problems, light sensitivity, epithelial problems like cracks in the corners of the mouth
B_3 niacin	Meat, fish, poultry, enriched breads and cereals, peanuts	14–16 mg	Synthesis of NAD, nerve function, cholesterol production	Cracked, scaly skin; dementia; diarrhea; also known as pellagra
B_5 pantothenic acid	Meat, poultry, potatoes, oats, enriched breads and cereals, tomatoes	5 mg	Synthesis of coenzyme A in fatty acid metabolism	Rare: symptoms may include fatigue, insomnia, depression, irritability
B_6 pyridoxine	Potatoes, bananas, beans, seeds, nuts, meat, poultry, fish, eggs, dark green leafy vegetables, soy, organ meats	1.3–1.5 mg	Sodium and potassium balance, red blood cell synthesis, protein metabolism	Confusion, irritability, depression, mouth and tongue sores
B_7 biotin	Liver, fruits, meats	30 μg	Cell growth, metabolism of fatty acids, production of blood cells	Rare in developed countries; symptoms include dermatitis, hair loss, loss of muscular coordination
B_9 folic acid	Liver, legumes, dark green leafy vegetables, enriched breads and cereals, citrus fruits	400 μg	DNA/protein synthesis	Poor growth, gingivitis, appetite loss, shortness of breath, gastrointestinal problems, mental deficits
B_{12} cyanocobalamin	Fish, meat, poultry, dairy products, eggs	2.4 μg	Fatty acid oxidation, nerve cell function, red blood cell production	Pernicious anemia, leading to nerve cell damage
C ascorbic acid	Citrus fruits, red berries, peppers, tomatoes, broccoli, dark green leafy vegetables	75–90 mg	Necessary to produce collagen for formation of connective tissue and teeth, and for wound healing	Dry hair, gingivitis, bleeding gums, dry and scaly skin, slow wound healing, easy bruising, compromised immunity; can lead to scurvy

Links to Enhanced Learning

More detailed information on the different water-soluble vitamins can be found at 9.3: Water-Soluble Vitamins – Medicine LibreTexts (https://med.libretexts.org/Under_Construction/Purgatory/Book%3A_Human_Nutrition_1e_(University_of_Hawaii)/09%3A_Vitamins/9.03%3A_Water-Soluble_Vitamins).

Exercise 28.5a

Using Infographic 28.5a., identify the functional groups that make the water-soluble vitamins soluble.



Infographic 28.5a. Read more about "The Chemical Structures of Vitamins (https://www.compoundchem.com/2015/01/13/vitamins/)" by Andy Brunning / Compound Interest, CC BY-NC-ND, or access a text-based summary of infographic 28.5a [New tab].

Solution:

Functional Groups that make water-soluble vitamins soluble

Water soluble vitamin	Functional groups responsible for water solubility
B1	Amine, alcohol, sulfur group
B2	Alcohol, amide, amine
B3	(left) carboxylic acid, amine (right) amide, amine
B5	Alcohol, carboxylic acid, amide
B6	Aldehyde, alcohol, amine, phosphate group
B7	Amide, carboxylic acid, sulfur group
B9	Amide, amine, carboxylic acid
B12	Amide, alcohol, amine, phosphate group
C	Ester, alcohol

Indigenous Perspectives: Inuit Nutrition

Sources of vitamins are dependent on access to specific types of food. In the Arctic, food sources are limited. The Inuit traditional diet has been studied to understand the source of key vitamins available in Arctic conditions (Naqitarvik et al, 2022):

Figure 28.5d. Polar bear (credit: Image by Alan Wilson, CC BY-SA 3.0).

- Vitamin A: obtained from liver of animals such as polar bears. Typically eaten raw.
- Vitamin D: obtained from liver of animals.
- Vitamin C: obtained from liver of animals (raw only), berries, raw fish eggs, raw whale skin (maktaaq/mattak)
- Vitamin K1: obtained from rhubarb

- Omega fatty acids: obtained from seal blubber (uqsuq) and oil

For more details, read this article: Living on the Edge | Chem 13 News Magazine | University of Waterloo (uwaterloo.ca) (https://uwaterloo.ca/chem13-news-magazine/fall-2022-special-edition/feature/living-edge)

Vitamins as Antioxidants

The "big three" vitamin antioxidants are vitamins E, A, and C, although it may be that they are called the "big three" only because they are the most studied. Antioxidants prevent damage from free radicals, which are molecules that are highly reactive because they have unpaired electrons. Free radicals are formed not only through metabolic reactions involving oxygen but also by such environmental factors as radiation and pollution. Free radicals react most commonly with lipoproteins and unsaturated fatty acids in cell membranes, removing an electron from those molecules and thus generating a new free radical. The process becomes a chain reaction that finally leads to the oxidative degradation of the affected compounds. Antioxidants react with free radicals to stop these chain reactions by forming a more stable molecule or, in the case of vitamin E, a free radical that is much less reactive (vitamin E is converted back to its original form through interaction with vitamin C). A simplified diagram on the role of antioxidants with DNA is shown in Figure 28.5e. Here, antioxidants neutralize free radicals to prevent DNA damage. Preventing DNA damage helps maintain our genetic code. Other antioxidants obtained from the diet are given in Table 28.5c.

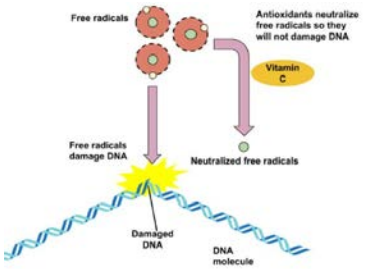


Figure 28.5e. Antioxidants Role (credit: Image by Allison Calabrese, CC BY 4.0).

Table 28.5c. Some Antioxidants Obtained from Diet and Their Related Functions.

Antioxidant	Functions Attributed to Antioxidant Capacity
Vitamin A	Protects cellular membranes, prevents glutathione depletion, maintains free radical detoxifying enzyme systems, reduces inflammation
Vitamin E	Protects cellular membranes, prevents glutathione depletion
Vitamin C	Protects DNA, RNA, proteins, and lipids, aids in regenerating vitamin E
Carotenoids	Free radical scavengers
Lipoic acid	Free radical scavenger, aids in regeneration of vitamins C and E
Phenolic acids	Free radical scavengers, protect cellular membranes

Table source: Wikipedia, *Chemistry for Changing Times (Hill & McCreary)*, CC BY-NC-SA 4.0.

Effects of Cooking

The USDA has conducted extensive studies on the percentage losses of various nutrients from different food types and cooking methods. Some vitamins may become more "bio-available" – that is, usable by the body – when foods are cooked. Table 28.5d. shows whether various vitamins are susceptible to loss from heat—such as heat from boiling, steaming, frying, etc. The effect of cutting vegetables can be seen from exposure to air and light. Water-soluble vitamins such as B and C dissolve into the water when a vegetable is boiled and are then lost when the water is discarded.

Table 28.5d. Vitamin Stability Upon Air, Light and Heat Exposure.

Vitamin	Soluble in Water	Stable to Air Exposure	Stable to Light Exposure	Stable to Heat Exposure
Vitamin A	no	partially	partially	relatively stable
Vitamin C	very unstable	yes	no	no
Vitamin D	no	no	no	no
Vitamin E	no	yes	yes	no
Vitamin K	no	no	yes	no
Thiamine (B_1)	highly	no	?	> 100 °C
Riboflavin (B_2)	slightly	no	in solution	no
Niacin (B_3)	yes	no	no	no
Pantothenic Acid (B_5)	quite stable	no	no	yes
Vitamin B_6	yes	?	yes	?
Biotin (B_7)	somewhat	?	?	no
Folic Acid (B_9)	yes	?	when dry	at high temp
Cobalamin (B_{12})	yes	?	yes	no

Table source: Wikipedia, *Chemistry for Changing Times (Hill & McCreary)*, CC BY-NC-SA 4.0.

Attribution & References

Except where otherwise noted, this page has been adapted by Samantha Sullivan Sauer from

- "17.4: Minerals, Vitamins, and Other Essentials" In *Chemistry for Changing Times (Hill & McCreary)* by LibreTexts, licensed under CC BY-NC-SA 4.0. Contributors from original source:
  - OpenStaxAP
  - UofHawaiiNutrition
  - Wikipedia
  - Marisa Alviar-Agnew (Sacramento City College)
- "18.9: Enzyme Cofactors and Vitamins" In *Basics of General, Organic, and Biological Chemistry (Ball et al.)* by David W. Ball, John W. Hill, and Rhonda J. Scott via LibreTexts, CC BY-NC-SA 4.0./ A LibreTexts version of *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0.

Modifications: combined the two sources, removed mention of minerals and fibre from material

References cited in-text

Naqitarvik, R., Anderson, C. C., & Rayner-Canham, G. (2022, Fall). Living on the edge: Some chemistry of the Inuit diet (https://uwaterloo.ca/chem13-news-magazine/fall-2022-special-edition/feature/living-edge). *Chem 13 News Magazine*.

CHAPTER 28 - SUMMARY

28.1 Carbohydrates

Carbohydrates, a large group of biological compounds containing carbon, hydrogen, and oxygen atoms, include sugars, starch, glycogen, and cellulose. All carbohydrates contain alcohol functional groups, and either an aldehyde or a ketone group (or a functional group that can be converted to an aldehyde or ketone). The simplest carbohydrates are monosaccharides. Those with two monosaccharide units are disaccharides, and those with many monosaccharide units are polysaccharides. Most sugars are either monosaccharides or disaccharides. Cellulose, glycogen, and starch are polysaccharides.

Many carbohydrates exist as stereoisomers, in which the three-dimensional spatial arrangement of the atoms in space is the only difference between the isomers. These particular stereoisomers contain at least one chiral carbon, a carbon atom that has four different groups bonded to it. A molecule containing a chiral carbon is nonsuperimposable on its mirror image, and two molecules that are nonsuperimposable mirror images of each other are a special type of stereoisomer called enantiomers. Enantiomers have the same physical properties, such as melting point, but differ in the direction they rotate polarized light.

A sugar is designated as being a D sugar or an L sugar according to how, in a Fischer projection of the molecule, the hydrogen atom and OH group are attached to the *penultimate* carbon atom, which is the carbon atom immediately before the terminal alcohol carbon atom. If the structure at this carbon atom is the same as that of D-glyceraldehyde (OH to the right), the sugar is a D sugar; if the configuration is the same as that of L-glyceraldehyde (OH to the left), the sugar is an L sugar.

Monosaccharides of five or more carbons atoms readily form cyclic structures when the carbonyl carbon atom reacts with an OH group on a carbon atom three or four carbon atoms distant. Consequently, glucose in solution exists as an equilibrium mixture of three forms, two of them cyclic (α- and β-) and one open chain. In Haworth projections, the *alpha* form is drawn with the OH group on the "former" carbonyl carbon atom (anomeric carbon) pointing downward; the *beta* form, with the OH group pointing upward; these two compounds are stereoisomers and are given the more specific term of anomers. Any solid sugar can be all alpha or all beta. Once the sample is dissolved in water, however, the ring opens up into the open-chain structure and then closes to form either the α- or the β-anomer. These interconversions occur back and forth until a dynamic equilibrium mixture is achieved in a process called mutarotation.

The carbonyl group present in monosaccharides is easily oxidized by Tollens' or Benedict's reagents (as well as others). Any mono- or disaccharide containing a free anomeric carbon is a reducing sugar. The disaccharide *maltose* contains two glucose units joined in an α-1,4-glycosidic linkage. The disaccharide *lactose* contains a galactose unit and a glucose unit joined by a β-1,4-glycosidic linkage. Both maltose and lactose contain a free

anomeric carbon that can convert to an aldehyde functional group, so they are reducing sugars; they also undergo mutarotation. Many adults, and some children, have a deficiency of the enzyme lactase (which is needed to break down lactose) and are said to be lactose intolerant. A more serious problem is the genetic disease galactosemia, which results from the absence of an enzyme needed to convert galactose to glucose.

The disaccharide *sucrose* (table sugar) consists of a glucose unit and a fructose unit joined by a glycosidic linkage. The linkage is designated as an α-1,β-2-glycosidic linkage because it involves the OH group on the first carbon atom of glucose and the OH group on the second carbon atom of fructose. Sucrose is not a reducing sugar because it has no anomeric carbon that can reform a carbonyl group, and it cannot undergo mutarotation because of the restrictions imposed by this linkage.

Starch, the principal carbohydrate of plants, is composed of the polysaccharides amylose (10%–30%) and amylopectin (70%–90%). When ingested by humans and other animals, starch is hydrolyzed to glucose and becomes the body's energy source. *Glycogen* is the polysaccharide animals use to store excess carbohydrates from their diets. Similar in structure to amylopectin, glycogen is hydrolyzed to glucose whenever an animal needs energy for a metabolic process. The polysaccharide *cellulose* provides structure for plant cells. It is a linear polymer of glucose units joined by β-1,4-glycosidic linkages. It is indigestible in the human body but digestible by many microorganisms, including microorganisms found in the digestive tracts of many herbivores.

28.2 Lipids

Lipids, found in the body tissues of all organisms, are compounds that are more soluble in organic solvents than in water. Many of them contain fatty acids, which are carboxylic acids that generally contain an even number of 4–20 carbon atoms in an unbranched chain. Saturated fatty acids have no carbon-to-carbon double bonds. Monounsaturated fatty acids have a single carbon-to-carbon double bond, while polyunsaturated fatty acids have more than one carbon-to-carbon double bond. Linoleic and linolenic acid are known as essential fatty acids because the human body cannot synthesize these polyunsaturated fatty acids. The lipids known as fats and oils are triacylglycerols, more commonly called triglycerides—esters composed of three fatty acids joined to the trihydroxy alcohol glycerol. Fats are triglycerides that are solid at room temperature, and oils are triglycerides that are liquid at room temperature. Fats are found mainly in animals, and oils found mainly in plants. *Saturated triglycerides* are those containing a higher proportion of saturated fatty acid chains (fewer carbon-to-carbon double bonds); *unsaturated triglycerides* contain a higher proportion of unsaturated fatty acid chains.

Saponification is the hydrolysis of a triglyceride in a basic solution to form glycerol and three carboxylate anions or soap molecules. Other important reactions are the hydrogenation and oxidation of double bonds in unsaturated fats and oils.

Phospholipids are lipids containing phosphorus. In phosphoglycerides, the phosphorus is joined to an amino alcohol unit. Some phosphoglycerides, like lecithins, are used to stabilize an emulsion—a dispersion of

two liquids that do not normally mix, such as oil and water. Sphingolipids are lipids for which the precursor is the amino alcohol sphingosine, rather than glycerol. A glycolipid has a sugar substituted at one of the OH groups of either glycerol or sphingosine. All are highly polar lipids found in cell membranes.

Polar lipids have dual characteristics: one part of the molecule is ionic and dissolves in water; the rest has a hydrocarbon structure and dissolves in nonpolar substances. Often, the ionic part is referred to as hydrophilic (literally, "water loving") and the nonpolar part as hydrophobic ("water fearing"). When placed in water, polar lipids disperse into any one of three arrangements: *micelles*, *monolayers*, and *bilayers*. Micelles are aggregations of molecules in which the hydrocarbon tails of the lipids, being hydrophobic, are directed inward (away from the surrounding water), and the hydrophilic heads that are directed outward into the water. Bilayers are double layers arranged so that the hydrophobic tails are sandwiched between the two layers of hydrophilic heads, which remain in contact with the water.

Every living cell is enclosed by a *cell membrane* composed of a lipid bilayer. In animal cells, the bilayer consists mainly of phospholipids, glycolipids, and the steroid cholesterol. Embedded in the bilayer are integral proteins, and peripheral proteins are loosely associated with the surface of the bilayer. Everything between the cell membrane and the membrane of the cell nucleus is called the cytoplasm.

Most lipids can be saponified, but some, such as steroids, cannot be saponified. The steroid cholesterol is found in animal cells but never in plant cells. It is a main component of all cell membranes and a precursor for hormones, vitamin D, and bile salts. Bile salts are the most important constituents of bile, which is a yellowish-green liquid secreted by the gallbladder into the small intestine and is needed for the proper digestion of lipids.

28.3 Amino Acids, Proteins, and Enzymes

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 neighbouring chain; (2) ionic bonding between one positively charged side chain and one negatively charged side chain; (3) disulfide linkages between cysteine units; and (4) dispersion forces between nonpolar side chains.

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

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

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

$$E + S \rightarrow E\text{–}S \rightarrow E + P$$

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.

28.4 Nucleic Acids and DNA

A cell's hereditary information is encoded in chromosomes in the cell's nucleus. Each chromosome is composed of proteins and deoxyribonucleic acid (DNA). The chromosomes contain smaller hereditary units called genes, which are relatively short segments of DNA. The hereditary information is expressed or used through the synthesis of ribonucleic acid (RNA). Both nucleic acids—DNA and RNA—are polymers composed of monomers known as nucleotides, which in turn consist of phosphoric acid (H_3PO_4), a nitrogenous base, and a pentose sugar.

The two types of nitrogenous bases most important in nucleic acids are purines—adenine (A) and guanine

(G)—and pyrimidines—cytosine (C), thymine (T), and uracil (U). DNA contains the nitrogenous bases adenine, cytosine, guanine, and thymine, while the bases in RNA are adenine, cytosine, guanine, and uracil. The sugar in the nucleotides of RNA is ribose; the one in DNA is 2-deoxyribose. The sequence of nucleotides in a nucleic acid defines the primary structure of the molecule.

RNA is a single-chain nucleic acid, whereas DNA possesses two nucleic-acid chains intertwined in a secondary structure called a double helix. The sugar-phosphate backbone forms the outside the double helix, with the purine and pyrimidine bases tucked inside. Hydrogen bonding between complementary bases holds the two strands of the double helix together; A always pairs with T and C always pairs with G.

Cell growth requires replication, or reproduction of the cell's DNA. The double helix unwinds, and hydrogen bonding between complementary bases breaks so that there are two single strands of DNA, and each strand is a template for the synthesis of a new strand. For protein synthesis, three types of RNA are needed: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). All are made from a DNA template by a process called transcription. The double helix uncoils, and ribonucleotides base-pair to the deoxyribonucleotides on one DNA strand; however, RNA is produced using uracil rather than thymine. Once the RNA is formed, it dissociates from the template and leaves the nucleus, and the DNA double helix reforms.

Translation is the process in which proteins are synthesized from the information in mRNA. It occurs at structures called ribosomes, which are located outside the nucleus and are composed of rRNA and protein. The 64 possible three-nucleotide combinations of the 4 nucleotides of DNA constitute the genetic code that dictates the sequence in which amino acids are joined to make proteins. Each three-nucleotide sequence on mRNA is a codon. Each kind of tRNA molecule binds a specific amino acid and has a site containing a three-nucleotide sequence called an anticodon.

The general term for any change in the genetic code in an organism's DNA is mutation. A change in which a single base is substituted, inserted, or deleted is a point mutation. The chemical and/or physical agents that cause mutations are called mutagens. Diseases that occur due to mutations in critical DNA sequences are referred to as genetic diseases.

28.5 Vitamins

Vitamins are essential parts of the diet. They are needed for the proper function of metabolic pathways in the body. Vitamins are not stored in the body, so they must be obtained from the diet or synthesized from precursors available in the diet. Some vitamins are water soluble such as the B vitamins. Some vitamins are lipid soluble such as Vitamin A. Cooking can significantly impact the stability of vitamins. Some vitamins act as antioxidants by neutralizing free radicals.

Attribution & References

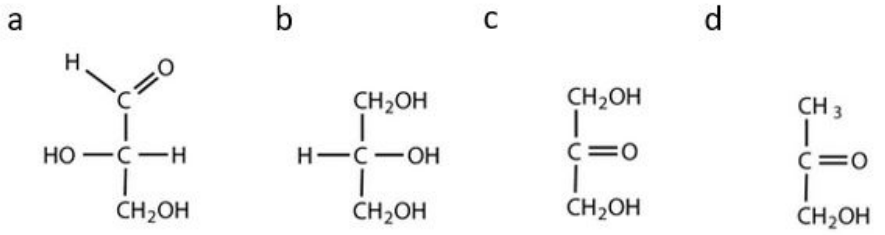
Except where otherwise noted, this page is adapted by Gregory A. Anderson and Samantha Sullivan Sauer from

- "16.S: Carbohydrates (Summary)", "17.S: Lipids (Summary)", "18.S: Amino Acids, Proteins, and Enzymes (Summary)", & "19.S: Nucleic Acids (Summary)" In *Basics of General, Organic, and Biological Chemistry (Ball et al.)* by David W. Ball, John W. Hill, and Rhonda J. Scott via LibreTexts, CC BY-NC-SA 4.0./ A LibreTexts version of *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0.
- "17.4: Minerals, Vitamins, and Other Essentials" In *Chemistry for Changing Times (Hill & McCreary)* by LibreTexts, licensed under CC BY-NC-SA 4.0. Contributors from original source:
  - OpenStaxAP
  - UofHawaiiNutrition
  - Wikipedia
  - Marisa Alviar-Agnew (http://www.scc.losrios.edu/chemistry/faculty/) (Sacramento City College (http://www.scc.losrios.edu/chemistry/))

CHAPTER 28 - REVIEW

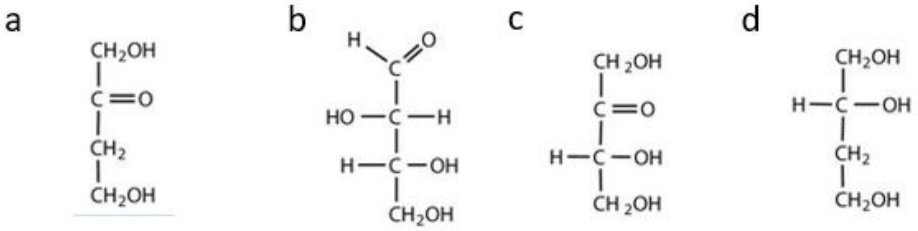
28.1 Carbohydrates

1. Which compounds would be classified as carbohydrates? **Check answer**[1]



(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

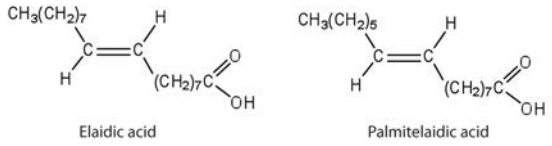
2. Which compounds would be classified as carbohydrates? **Check answer**[2]



(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

28.2 Lipids

1. The melting point of elaidic acid is 52°C.
   a. What trend is observed when comparing the melting points of elaidic acid, oleic acid, and stearic acid? Explain.**Check answer**[3]
   b. Would you expect the melting point of palmitelaidic acid to be lower or higher than that of elaidic acid? Explain. **Check answer**[4]



(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

2. Examine the labels on two brands of margarine and two brands of shortening and list the oils used in the various brands.
3. In cerebrosides (Figure 28.2s), is the linkage between the fatty acid and sphingosine an amide bond or an ester bond? Justify your answer.
4. Explain whether each compound would be expected to diffuse through the lipid bilayer of a cell membrane.
    a. potassium chloride
    b. $CH_3CH_2CH_2CH_2CH_2CH_3$
    c. fructose
5. Identify the role of each steroid hormone in the body.
    a. progesterone **Check answer**[5]
    b. aldosterone **Check answer**[6]
    c. testosterone **Check answer**[7]
    d. cortisol **Check answer**[8]
6. What fatty acid is the precursor for the prostaglandins? **Check answer**[9]
7. Identify three biological effects of prostaglandins. **Check answer**[10]
8. Why is it important to determine the ratio of LDLs to HDLs, rather than just the concentration of serum cholesterol?

28.3 Amino Acids, Proteins, and Enzymes

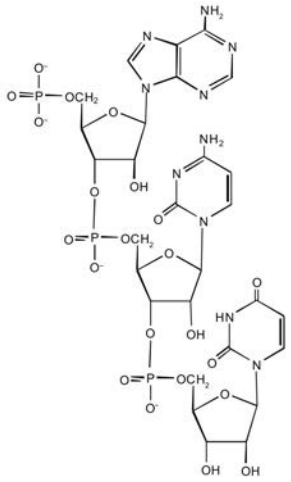
1. What is the general structure of an α-amino acid? **Check answer**[11]
2. Identify the amino acid that fits each description.
    a. also known as aspartate **Check answer**[12]
    b. almost as strong a base as sodium hydroxide **Check answer**[13]
    c. does not have a chiral carbon **Check answer**[14]
3. Write the side chain of each amino acid.
    a. serine **Check answer**[15]
    b. arginine **Check answer**[16]
    c. phenylalanine **Check answer**[17]
4. Write the side chain of each amino acid.
    a. aspartic acid **Check answer**[18]
    b. methionine **Check answer**[19]
    c. valine **Check answer**[20]
5. Identify an amino acid whose side chain contains a(n)
    a. amide functional group.
    b. aromatic ring.

   c. carboxyl group.
6. Identify an amino acid whose side chain contains a(n)
   a. OH group
   b. branched chain
   c. amino group
7. Define each term.
   a. zwitterion **Check answer**[21]
   b. isoelectric point **Check answer**[22]
8. Draw the structure for the anion formed when alanine (at neutral pH) reacts with a base. **Check answer**[23]
9. Draw the structure for the cation formed when alanine (at neutral pH) reacts with an acid. **Check answer**[24]
10. Distinguish between the N-terminal amino acid and the C-terminal amino acid of a peptide or protein. **Check answer**[25]
11. Describe the difference between an amino acid and a peptide. **Check answer**[26]
12. Amino acid units in a protein are connected by peptide bonds. What is another name for the functional group linking the amino acids? **Check answer**[27]
13. Draw the structure for each peptide.
   a. gly-val **Check answer**[28]
   b. val-gly **Check answer**[29]
14. Identify the C- and N-terminal amino acids for the peptide lys-val-phe-gly-arg-cys. **Check answer**[30]
15. Identify the C- and N-terminal amino acids for the peptide asp-arg-val-tyr-ile-his-pro-phe.
16. What is the predominant attractive force that stabilizes the formation of secondary structure in proteins? **Check answer**[31]
17. Distinguish between the tertiary and quaternary levels of protein structure. **Check answer**[32]
18. Briefly describe four ways in which a protein could be denatured. **Check answer**[33]
19. What name is given to the predominant secondary structure found in silk? **Check answer**[34]
20. What name is given to the predominant secondary structure found in wool protein?
21. A protein has a tertiary structure formed by interactions between the side chains of the following pairs of amino acids. For each pair, identify the strongest type of interaction between these amino acids.
   a. aspartic acid and lysine **Check answer**[35]
   b. phenylalanine and alanine **Check answer**[36]
   c. serine and lysine **Check answer**[37]
   d. two cysteines **Check answer**[38]
22. What level(s) of protein structure is(are) ordinarily disrupted in denaturation? What level(s) is(are) not? **Check answer**[39]
23. Distinguish between the lock-and-key model and induced-fit model of enzyme action. **Check answer**[40]

28.4 Nucleic Acids and DNA

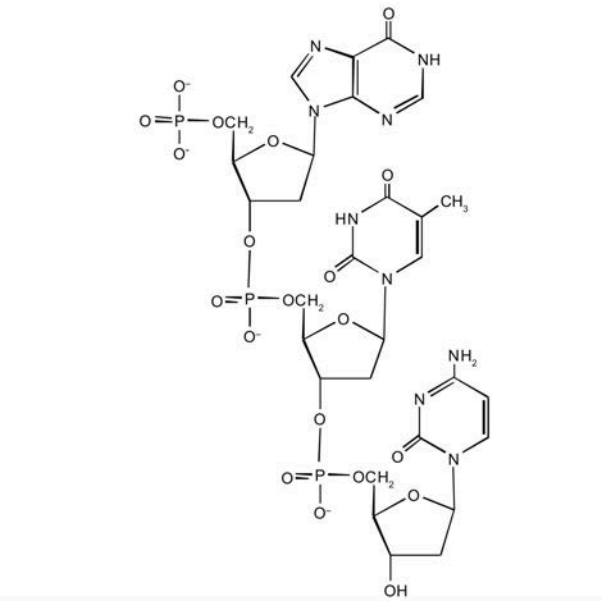
1. Name the two kinds of nucleic acids. **Check answer**[41]
2. Which type of nucleic acid stores genetic information in the cell? **Check answer**[42]
3. What are complementary bases? **Check answer**[43]
4. Why is it structurally important that a purine base always pair with a pyrimidine base in the DNA double helix? **Check answer**[44]
5. For this short RNA segment,
   a. identify the 5′ end and the 3′ end of the molecule.
   b. circle the atoms that comprise the backbone of the nucleic acid chain.
   c. write the nucleotide sequence of this RNA segment.

   Check answer[45]



(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

6. For this short DNA segment,
   a. identify the 5′ end and the 3′ end of the molecule.
   b. circle the atoms that comprise the backbone of the nucleic acid chain
   c. write the nucleotide sequence of this DNA segment.

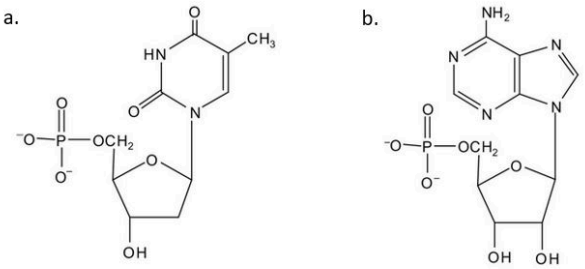


(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

7. Which nitrogenous base in DNA pairs with each nitrogenous base?
   a. cytosine **Check answer**[46]
   b. adenine **Check answer**[47]
   c. guanine **Check answer**[48]
   d. thymine **Check answer**[49]
8. Which nitrogenous base in RNA pairs with each nitrogenous base?
   a. cytosine
   b. adenine
   c. guanine
   d. thymine
9. How many hydrogen bonds can form between the two strands in the short DNA segment shown below?
   5′ ATGCGACTA 3′ 3′ TACGCTGAT 5′ **Check answer**[50]
10. How many hydrogen bonds can form between the two strands in the short DNA segment shown below?
    5′ CGATGAGCC 3′ 3′ GCTACTCGG 5′
11. Identify the three molecules needed to form the nucleotides in each nucleic acid.
    a. DNA **Check answer**[51]
    b. RNA **Check answer**[52]
12. Classify each compound as a pentose sugar, a purine, or a pyrimidine.
    a. adenine **Check answer**[53]
    b. guanine **Check answer**[54]

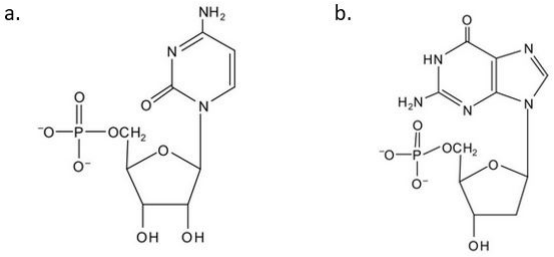
c. deoxyribose **Check answer**[55]
d. thymine **Check answer**[56]
e. ribose **Check answer**[57]
f. cytosine **Check answer**[58]

13. What is the sugar unit in each nucleic acid?
    a. RNA **Check answer**[59]
    b. DNA **Check answer**[60]
14. For each structure, circle the sugar unit and identify the nucleotide as a ribonucleotide or a deoxyribonucleotide. **Check answer**[61]



(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

15. For each structure, circle the nitrogenous base and identify it as a purine or pyrimidine. **Check answer**[62]



(credit: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0).

16. In DNA replication, a parent DNA molecule produces two daughter molecules. What is the fate of each strand of the parent DNA double helix? **Check answer**[63]
17. What is the role of DNA in transcription? What is produced in transcription? **Check answer**[64]
18. Describe how replication and transcription are similar. **Check answer**[65]
19. Describe how replication and transcription differ.
20. A portion of the coding strand for a given gene has the sequence

5′-ATGAGCGACTTTGCGGGATTA-3′.

a. What is the sequence of complementary template strand? **Check answer**[66]
b. What is the sequence of the mRNA that would be produced during transcription from this segment of DNA?**Check answer**[67]

21. A portion of the coding strand for a given gene has the sequence 5′-ATGGCAATCCTCAAACGCTGT-3′.
    a. What is the sequence of complementary template strand?
    b. What is the sequence of the mRNA that would be produced during transcription from this segment of DNA?
22. What are the roles of mRNA and tRNA in protein synthesis? **Check answer**[68]
23. What is the initiation codon? **Check answer**[69]
24. What are the termination codons and how are they recognized? **Check answer**[70]
25. Write the anticodon on tRNA that would pair with each mRNA codon.
    a. 5′-UUU-3′ **Check answer**[71]
    b. 5′-CAU-3′ **Check answer**[72]
    c. 5′-AGC-3′ **Check answer**[73]
    d. 5′-CCG-3′ **Check answer**[74]
26. Write the codon on mRNA that would pair with each tRNA anticodon.
    a. 5′-UUG-3′
    b. 5′-GAA-3′
    c. 5′-UCC-3′
    d. 5′-CAC-3′
27. The peptide hormone oxytocin contains 9 amino acid units. What is the minimum number of nucleotides needed to code for this peptide? **Check answer**[75]
28. Myoglobin, a protein that stores oxygen in muscle cells, has been purified from a number of organisms. The protein from a sperm whale is composed of 153 amino acid units. What is the minimum number of nucleotides that must be present in the mRNA that codes for this protein?
29. Use Figure 28.4r to determine the amino acid sequence produced from this mRNA sequence: 5′-AUGAGCGACUUUGCGGGAUUA-3′. **Check answer**[76]
30. Use Figure 28.4r to determine the amino acid sequence produced from this mRNA sequence: 5′-AUGGCAAUCCUCAAACGCUGU-3′
31. What effect can UV radiation have on DNA? **Check answer**[77]
32. What causes PKU? **Check answer**[78]
33. How is PKU detected and treated? **Check answer**[79]
34. A portion of the coding strand of a gene was found to have the sequence 5′-ATGAGCGACTTTCGCCCATTA-3′. A mutation occurred in the gene, making the sequence 5′-ATGAGCGACCTTCGCCCATTA-3′.

   a. Identify the mutation as a substitution, an insertion, or a deletion. **Check answer**[80]
   b. What effect would the mutation have on the amino acid sequence of the protein obtained from this mutated gene? **Check answer**[81]

35. A portion of the coding strand of a gene was found to have the sequence 5′-ATGGCAATCCTCAAACGCTGT-3′. A mutation occurred in the gene, making the sequence 5′-ATGGCAATCCTCAACGCTGT-3′.
   a. Identify the mutation as a substitution, an insertion, or a deletion.
   b. What effect would the mutation have on the amino acid sequence of the protein obtained from this mutated gene?
36. What is a mutagen? **Check answer**[82]
37. Give two examples of mutagens. **Check answer**[83]

28.5 Vitamins

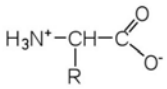
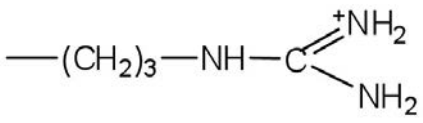
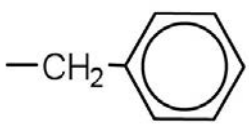
1. Identify each vitamin as water soluble or fat soluble. a) vitamin D b) vitamin C c) vitamin B12 **Check answer**[84]
2. Identify each vitamin as water soluble or fat soluble. a) niacin b) cholecalciferol c) biotin
3. What is the function of each vitamin? a) vitamin A b) biotin c) vitamin K **Check answer**[85]

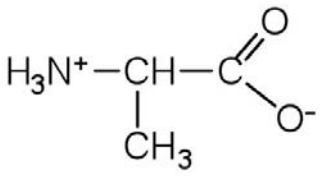
Attribution & References

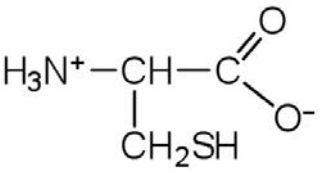
Except where otherwise noted, this page (including images in solutions) is adapted by Gregory A. Anderson and Samantha Sullivan Sauer from

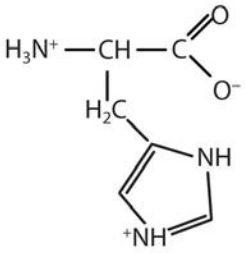
- “17.E: Exercises“, “18.E: Amino Acids, Proteins, and Enzymes (Exercises)” & “19.E: Nucleic Acids” (Exercises) In *Basics of General, Organic, and Biological Chemistry (Ball et al.)* by David W. Ball, John W. Hill, and Rhonda J. Scott via LibreTexts, CC BY-NC-SA 4.0./ A LibreTexts version of *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0.
- “18.1 Properties of Amino Acids” & “18.3 Peptides” In *Introduction to Chemistry: GOB (v. 1.0),* CC BY-NC 3.0
- Images in solutions are from the original source, except:
  - 28.3 Question 1, 3, 4: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0
  - 28.3 Question 8, 9: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0
  - 28.3 Question 13: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0
  - 28.4 Question 5: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0
  - 28.5 Question 14, 15: *Intro Chem: GOB (V. 1.0).*, CC BY-NC-SA 3.0

Notes

1. a. This is a carbohydrate because the molecule contains an aldehyde functional group with OH groups on the other two carbon atoms. b. This is not a carbohydrate because the molecule does not contain an aldehyde or a ketone functional group. c. This is a carbohydrate because the molecule contains a ketone functional group with OH groups on the other two carbon atoms. d. This is not a carbohydrate; although it has a ketone functional group, one of the other carbons atoms does not have an OH group attached.
2. **a.** This is not a carbohydrate; although it has a ketone functional group, one of the other carbons atoms does not have an OH group attached. **b.** This is a carbohydrate because the molecule contains an aldehyde functional group with OH groups on the other two carbon atoms. **c.** This is a carbohydrate because the molecule contains a ketone functional group with OH groups on the other two carbon atoms. **d.** This is not a carbohydrate; it does not contain an aldehyde group, nor does it contain a ketone group.
3. Stearic acid has the highest melting point, followed by elaidic acid, and then oleic acid with the lowest melting point. Elaidic acid is a trans fatty acid, and the carbon chains can pack together almost as tightly as those of the saturated stearic acid. Oleic acid is a cis fatty acid, and the bend in the hydrocarbon chain keeps these carbon chains from packing as closely together; fewer interactions lead to a much lower melting point.
4. The melting point of palmitelaidic acid should be lower than that of elaidic acid because it has a shorter carbon chain (16, as compared to 18 for elaidic acid). The shorter the carbon chain, the lower the melting point due to a decrease in intermolecular interactions.
5. regulates the menstrual cycle and maintains pregnancy
6. regulates salt metabolism by stimulating the kidneys to retain sodium and excrete potassium
7. stimulates and maintains male sex characteristics
8. stimulates the conversion of proteins to carbohydrates
9. arachidonic acid
10. induce smooth muscle contraction, lower blood pressure, and contribute to the inflammatory response
11. 
12. aspartic acid
13. arginine
14. glycine
15. CH_2OH^-
16. 
17. 

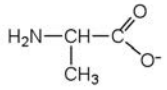
18. 

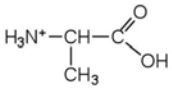
19. 

20. 

21. an electrically neutral compound that contains both negatively and positively charged groups

22. the pH at which a given amino acid exists in solution as a zwitterion

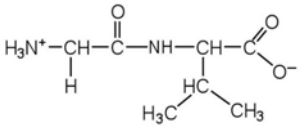
23. 

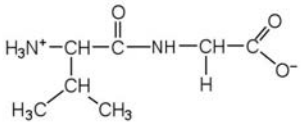
24. 

25. The N-terminal end is the end of a peptide or protein whose amino group is free (not involved in the formation of a peptide bond), while the C-terminal end has a free carboxyl group.

26. A peptide is composed of two or more amino acids. Amino acids are the building blocks of peptides.

27. amide bond

28. 

29. 

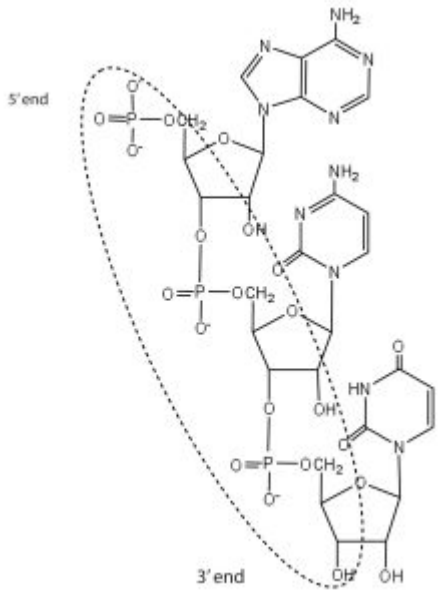
30. C-terminal amino acid: cys; N-terminal amino acid: lys

31. hydrogen bonding

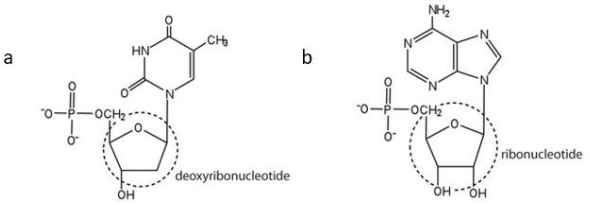
32. Tertiary structure refers to the unique three-dimensional shape of a single polypeptide chain, while quaternary structure describes the interaction between multiple polypeptide chains for proteins that have more than one

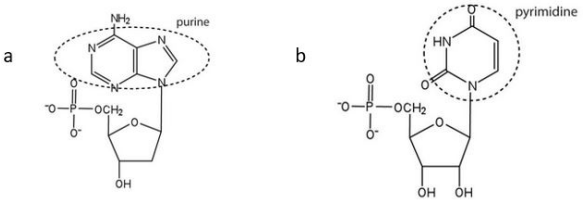
polypeptide chain.

33. (1) heat a protein above 50°C or expose it to UV radiation; (2) add organic solvents, such as ethyl alcohol, to a protein solution; (3) add salts of heavy metal ions, such as mercury, silver, or lead; and (4) add alkaloid reagents such as tannic acid
34. β-pleated sheet
35. ionic bonding
36. dispersion forces
37. dispersion forces
38. disulfide linkage
39. Protein denaturation disrupts the secondary, tertiary, and quaternary levels of structure. Only primary structure is unaffected by denaturation.
40. The lock-and-key model portrays an enzyme as conformationally rigid and able to bond only to substrates that exactly fit the active site. The induced fit model portrays the enzyme structure as more flexible and is complementary to the substrate only after the substrate is bound.
41. deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)
42. DNA
43. the specific base pairings in the DNA double helix in which guanine is paired with cytosine and adenine is paired with thymine
44. The width of the DNA double helix is kept at a constant width, rather than narrowing (if two pyrimidines were across from each other) or widening (if two purines were across from each other).
45. ACU



46. guanine
47. thymine
48. cytosine
49. adenine
50. 22 (2 between each AT base pair and 3 between each GC base pair)
51. nitrogenous base (adenine, guanine, cytosine, and thymine), 2-deoxyribose, and H_3PO_4

52. nitrogenous base (adenine, guanine, cytosine, and uracil), ribose, and H_3PO_4
53. purine
54. purine
55. pentose sugar
56. pyrimidine
57. pentose sugar
58. pyrimidine
59. ribose
60. deoxyribose
61. 

62. 

63. Each strand of the parent DNA double helix remains associated with the newly synthesized DNA strand.
64. DNA serves as a template for the synthesis of an RNA strand (the product of transcription).
65. Both processes require a template from which a complementary strand is synthesized.
66. 3′-TACTCGCTGAAACGCCCTAAT-5′
67. 5′-AUGAGCGACUUUGCGGGAUUA-3′
68. mRNA provides the code that determines the order of amino acids in the protein; tRNA transports the amino acids to the ribosome to incorporate into the growing protein chain.
69. AUG
70. UAA, UAG, and UGA; they are recognized by special proteins called release factors, which signal the end of the translation process.
71. 3′-AAA-5′
72. 3′-GUA-5′
73. 3′-UCG-5′
74. 3′-GGC-5′
75. 27 nucleotides (3 nucleotides/codon)
76. met-ser-asp-phe-ala-gly-leu
77. It can lead to the formation of a covalent bond between two adjacent thymines on a DNA strand, producing a thymine dimer.
78. the absence of the enzyme phenylalanine hydroxylase

79. PKU is diagnosed by assaying a sample of blood or urine for phenylalanine or one of its metabolites; treatment calls for an individual to be placed on a diet containing little or no phenylalanine.
80. substitution
81. Phenylalanine (UUU) would be replaced with leucine (CUU).
82. a chemical or physical agent that can cause a mutation
83. UV radiation and gamma radiation (answers will vary)
84. a) fat soluble, b) water soluble, c) water soluble
85. a) needed for the formation of vision pigments, b) needed for cell growth and metabolism c) needed for blood clotting

CHAPTER 28 - INFOGRAPHIC DESCRIPTIONS

Infographics used in Chapter 28

- 28.4a What makes up the Chemical Structure of DNA?
- 28.4b Today in Chemistry History – Rosalind Franklin and the structure of DNA
- 28.4c The 2020 Nobel Prize in Chemistry: Using genetic scissors to edit the genome
- 28.5a The Chemical Structures of Vitamins

28.4a What makes up the Chemical Structure of DNA?

The sugar-phosphate backbone. DNA is a polymer made up of units called nucleotides. The nucleotides are made of three different components: a sugar group, a phosphate group, and a base. There are four different bases: adenine thymine and cytosine. DNA strands are held together by hydrogen bonds between bases of adjacent strands. Adenine (A) always pairs with thymine (T), while guanine (G) always pairs with cytosine (C). Adenine pairs with uracil (U) in RNA.

$$\text{DNA} \xrightarrow[\textit{Transcription}]{} \text{RNA} \xrightarrow[\textit{Transcription}]{} \text{Protein}$$

The bases on a single strand of DNA act as a code. The letters form three letter codons, which code for amino acids – the building blocks of proteins.

An enzyme, RNA polymerase, transcribes DNA into mRNA (messenger ribonucleic acid). It splits apart the two strands that form the double helix, then reads a strand and copies the sequence of nucleotides. The only difference between the RNA and the original DNA is that in the place of thymine (T), another base with a similar structure is used: uracil (U).

- **DNA sequence:** T-T-C-C-T-G-A-A-C-C-C-G-T-T-A
- **mRNA sequence:** U-U-C-C-U-G-A-A-C-C-C-G-U-U-A
  - **Phenylalanine:** U-U-C
  - **Leucine:** C-U-G
  - **Asparagine:** A-A-C

- **Proline**: C-C-G
- **Leucine:** U-U-A

In multicellular organisms, the mRNA carries genetic code out of the cell nucleus, to the cytoplasm. Here, protein synthesis takes place. 'Translation' is the process of turning the mRNA's 'code' into proteins. Molecules called ribosomes carry out this process, building up proteins from the amino acids coded for.

Read more about "What makes up the Chemical Structure of DNA? [New tab] (https://www.compoundchem.com/2015/03/24/dna/)" by Andy Brunning / Compound Interest, CC BY-NC-ND

28.4b Today in Chemistry History – Rosalind Franklin and the structure of DNA

Franklin was born 25 July 1920 and died 16 April 1958.

Rosalind Franklin was a chemist and X-ray crystallographer whose work was instrumental in the discovery of the structure of DNA. She missed out on a Nobel Prize for her work as they are not awarded posthumously.

DNA is a polymer made up from monomers called nucleotides. A sugar phosphate backbone forms the two strands, which are held together by hydrogen bonds between the bases found on these strands. Photograph 51 is an X-ray diffraction image of DNA taken during Franklin's research. It was crucial in developing a model of DNA and confirming its double helical structure.

Read more about "Today in Chemistry History – Rosalind Franklin and the structure of DNA [New tab] (https://www.compoundchem.com/2017/07/25/franklin/)" by Andy Brunning / Compound Interest, CC BY-NC-ND

28.4c The 2020 Nobel Prize in Chemistry: Using genetic scissors to edit the genome

The 2020 Noble Prize in Chemistry was awarded to Emmanuelle Charpentier and Jennifer A. Doudna for the development if CRISPR-Cas9 genetic scissors, a method for genome editing.

CRISPR stands for clustered regularly interspaced short palindromic repeats. It refers to repeated sequences in bacteria and archaea DNA. The sequences are part of an immune system; if a bacterium survives a viral infection, it adds a section of the virus genetic code to the CRISPR region of its own to serve as a memory in case it's infected again. Charpentier and Doudna saw this could be used for gene editing.

CRISPR gene editing process is as follows:

1. Create a strand of guide RNA matching the DNA sequence where we want to make a cut and use a scissor protein, Cas9, to bind to the guide RNA.

2. The guide RNA searches for the target section of the DNA, transports the scissor protein to it, and the scissor protein cuts the DNA at this point.
3. The cell will try and repair the cut DNA, but this process is error-prone – disrupting the gene function. If we add a template, the cell will use this to carry out the repair, allowing us to edit the genetic code.

The ability to edit genomes has been used in plant breeding. Clinical tries underway to use this in therapies to treat some cancers and hopefully will lead to cures for some inherited diseases.

Read more about "The 2020 Nobel Prize in Chemistry: Using genetic scissors to edit the genome [New tab] (https://www.compoundchem.com/2020/10/07/2020nobelchemistry/)" by Andy Brunning / Compound Interest, CC BY-NC-ND

28.5a The Chemical Structures of Vitamins

Vitamins are essential nutrients that our body needs in small amounts. An organic compound is defined as a vitamin when it is required by an organism, but not synthesized by that organism in the required amounts (or at all). There are 13 recognized vitamins, these can be divided broadly into two classes: water-soluble vitamins and fat-soluble vitamins.

Water-soluble vitamin are not stored in the body and are generally required more frequently than the fat-soluble vitamins.

Fat-soluble vitamins are stored in the liver and fatty tissues until required, therefore they can be harmful if too much is taken in.

Listed below are the water-soluble vitamins:

1. Vitamin B1 (Thiamin): Can also occur in pyrophosphate ester form. Used to keep nerve and muscle tissue healthy, and important for processing of carbohydrates and some proteins.
2. Vitamin B2 (Riboflavin): Excess turns urine bright yellow. Important for body growth, red blood cell production, keeping the eyes healthy, and helps processing of carbohydrates.
3. Vitamin B3 (Nicotinic acid and Nicotineamide): Niacin is collective name for these compounds. Helps with digestion, digestive systems health and helps with the processing of carbohydrates.
4. Vitamin B5 (Pantothenic acid): Also occurs in pyrophosphate ester form. Important for manufacturing red blood cells, maintaining a healthy digestive system, helps process carbohydrates.
5. Vitamin B6 (Pyridoxal phosphate): Active form in mammalian tissue. Helps make some brain chemicals needed for normal brain function and also helps make red blood cells and immune system cells.
6. Vitamin B7 (Biotin): Produced by intestinal bacteria. Needed for metabolism of various compounds, often recommended for strengthening hair though evidence varies.
7. Vitamin B9 (Folic acid): Found as tetrahydrofolate in food. Important for brain function and mental health, aids production of DNA and RNA, important when tissues are growing quickly.

8. Vitamin B12 (Cobalamin): Usually contains CN as the R group. Important for the nervous system, for making red blood cells, and helps in the production of DNA and RNA.
9. Vitamin C (Ascorbic acid): Deficiency causes scurvy. Important for a healthy immune system, helps produce collagen (used to make skin and other tissues), and helps wound healing.

Listed below are the fat-soluble vitamins:

1. Vitamin A (Retinol): Is an active form in mammalian tissues. It is important for eyesight, strengthens the immune system, keeps skin and lining pf [arts of the body healthy.
2. Vitamin D (Cholecalciferol): There is a difference between natural form and form used in supplements. Important for bone health and maintaining the immune system function, may also have a preventative role in cancers.
3. Vitamin E (Alpha-tocopherol): Group includes tocopherols and tocotrienols. An antioxidant that helps prevent damage to cells and may have a preventative role in cancer, also helps to make red blood cells.
4. Vitamin K (Menadione): All K vitamins are menadione or derivatives. Helps blood clot properly, plays key role in bone health, newborns receive vitamin K injections to prevent bleeding.

Read more about "The Chemical Structures of Vitamins [New tab] (https://www.compoundchem.com/2015/01/13/vitamins/)" by Andy Brunning / Compound Interest, CC BY-NC-ND

Attribution & References

Compound Interest infographics are created by Andy Brunning and licensed under CC BY-NC-ND

Except where otherwise noted, content on this page has been created as a textual summary of the infographics used within our OER. Please refer to the original website (noted below each description) for further details about the image.